

VOLUME 115, 15 FEBRUARY 2017

ISSN 1046-2023

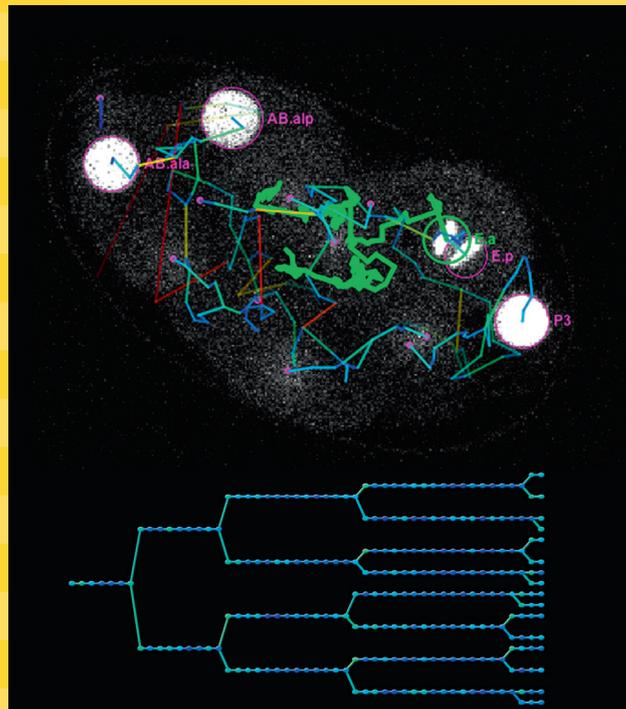
# METHODS

FOCUSING ON RAPIDLY DEVELOPING TECHNIQUES

## *Image Processing for Biologists*

Editors

Richard Butler and Alex Sossick



# METHODS

## EDITOR-IN-CHIEF

**Kenneth W. Adolph**

*University of Minnesota Medical School  
Minneapolis, Minnesota, USA*

## EDITORS

**Sun Kim**

*Seoul National University (SNU)  
Seoul, South Korea*

**Christophe Lavelle**

*National Museum of Natural History, Paris, France*

**Jia Li**

*Imperial College London, London, UK*

**Michelle Peckham**

*University of Leeds  
Leeds, UK*

**Nils Walter**

*University of Michigan, Ann Arbor, Michigan, USA*

## EDITORIAL BOARD

Karen Adelman  
Jurg Bahler  
Vytas Bankaitis  
Toni Cathomen  
Jonathan Chaires  
David Corey  
Nancy Cox  
Jack E. Dixon  
Erica Golemis

Allan Jacobson  
Ulf Landegren  
David M.J. Lilley  
James Manley  
Lynne Maquat  
Joachim Messing  
Tom Misteli  
Stephen Neidle  
Michael Pfaffl

Cenk Sahinalp  
Thomas Schmittgen  
Marc H.V. Van Regenmortel  
Johannes Walter  
Wei Wang  
Limsoon Wong  
Jerry Workman



## Editorial

## Image Processing and Analysis for Biologists



Over the last 30 years, biological image analysis methods have progressed from direct observation to obtaining and quantifying high content data using a wide range of mathematical approaches. This rapid technological development has made computational analysis essential for bio-imaging experiments in order to provide insights into biological systems that are not possible in any other way.

Few biologists would disagree with the idea that scientific investigation requires that the methods used can be described, shared and repeated. However, despite this philosophy, the use of closed-source commercial software for image analysis remains common, leading to published results obtained using unknown methods that can only be replicated by purchasing expensive licences for black box solutions. In addition, the lack of flexibility in closed applications often does not allow methods to be adapted and developed to meet the constantly changing requirements for novel biological image analysis software. For these reasons the papers in this special issue focus on open, scientific tools.

Advances in both super-resolution and high-throughput imaging have led to huge volumes of image data being produced, presenting new computational and algorithmic challenges for extraction of meaningful biological information from these datasets. Even for analysis of small datasets it is advantageous to minimise the requirement for input from human operators, which unavoidably leads to unconscious bias in the results. This has prompted the adoption of many automated and semi-automated approaches, the most common of which is to segment images and detect objects for measurement [1], which often requires methods specifically tailored to the object of interest and the type of images used [2]. When acquiring images for quantification, resolution is a key factor influencing the amount of information available from biological samples, giving rise to continuing work on methods to acquire images at high resolution [3] and to improve image resolution by deconvolution [4].

Colocalisation analysis is commonly used to detect localisation of proteins of interest to particular subcellular compartments [5] or the spatial relationships between labelled objects [6]. In addition to multi-channel acquisition in 2D or 3D, the increase in available methods for imaging live samples has led to the requirement for computational methods allowing quantification of large time series [7]. This includes continually developing algorithms for tracking of objects over time [8] and identification of events such as cell division [9].

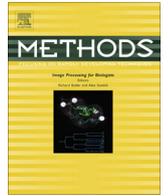
As new methods are developed and equipment for gathering large image datasets becomes widely available, data informatics

and statistical approaches become increasingly important for classification [10] and interpretation of biological results [11] as well as evaluation of their validity [12]. This includes data captured using techniques such as nuclear magnetic resonance and ultrasound imaging which are used for both scientific and medical purposes, producing digital images with different properties but requiring related methods for processing [13].

Biologists should ideally have access to image analysis specialists who can develop and implement appropriate tools, and these tools must be available in an open, user-friendly form. It is unrealistic to expect biologists to be experts in their field as well as in microscopy, mathematics, image processing, statistical analysis and programming. This special issue therefore attempts to present novel algorithms and reviews of existing methods in a way that is accessible to non-specialist biologists and provides insight into the wide range of techniques available for image analysis.

## References

- [1] X. Descombes, Multiple objects detection in biological images using a Marked Point Process Framework, *Methods* 115 (2017) 2–8.
- [2] A.D. Staszowska, P. Fox-Roberts, E. Foxall, G.E. Jones, S. Cox, Investigation of podosome ring protein arrangement using localization microscopy images, *Methods* 115 (2017) 9–16.
- [3] F. Lam, D. Cladière, C. Guillaume, K. Wassmann, S. Bolte, Super-resolution for everybody: an image processing workflow to obtain high-resolution images with a standard confocal microscope, *Methods* 115 (2017) 17–27.
- [4] D. Sage, L. Donati, F. Soulez, D. Fortun, G. Schmit, A. Seitz, R. Guet, C. Vonesch, M. Unser, DeconvolutionLab2: an open-source software for deconvolution microscopy, *Methods* 115 (2017) 28–41.
- [5] J. Pike, I. Styles, J. Rappoport, J. Heath, Quantifying receptor trafficking and colocalization with confocal microscopy, *Methods* 115 (2017) 42–54.
- [6] J.-F. Gilles, M. Dos Santos, T. Boudier, S. Bolte, N. Heck, DiAna, an ImageJ tool for object-based 3D co-localization and distance analysis, *Methods* 115 (2017) 55–64.
- [7] J. Rittscher, T. Nketia, H. Sailem, G. Rohde, R. Machiraju, Analysis of live cell images: methods, tools and opportunities, *Methods* 115 (2017) 65–79.
- [8] J.-Y. Tinevez, N. Perry, J. Schindelin, G.M. Hoopes, G.D. Reynolds, E. Laplantine, S.L. Shorte, K.W. Eliceiri, TrackMate: an open and extensible platform for single-particle tracking, *Methods* 115 (2017) 80–90.
- [9] J.S. Grah, J.A. Harrington, S.B. Koh, J.A. Pike, A. Schreiner, M. Burger, C.-B. Schönlieb, S. Reichelt, Mathematical imaging methods for mitosis analysis in live-cell phase contrast microscopy, *Methods* 115 (2017) 91–99.
- [10] Z. Li, D. Metaxas, A. Lu, S. Zhang, Interactive exploration for continuously expanding neuron databases, *Methods* 115 (2017) 100–109.
- [11] A. Chessel, An overview of data science uses in bioimage informatics, *Methods* 115 (2017) 110–118.
- [12] J. Funke, J. Klein, F. Moreno-Noguer, A. Cardona, M. Cook, TED: a tolerant edit distance for segmentation evaluation, *Methods* 115 (2017) 119–127.
- [13] C. Che, T.S. Mathai, J. Galeotti, Ultrasound registration: a review, *Methods* 115 (2017) 128–143.



## Analysis of live cell images: Methods, tools and opportunities



Thomas A. Nketia<sup>a</sup>, Heba Sailem<sup>a</sup>, Gustavo Rohde<sup>e</sup>, Raghu Machiraju<sup>b</sup>, Jens Rittscher<sup>a,c,d,\*</sup>

<sup>a</sup> Institute of Biomedical Engineering, Department of Engineering Science, Old Road Campus Research Building, University of Oxford, Headington, Oxford OX3 7DQ, United Kingdom

<sup>b</sup> Computer Science and Engineering, The Ohio State University, 2015 Neil Ave, Columbus, OH 43210, United States

<sup>c</sup> Ludwig Institute for Cancer Research, University of Oxford, Nuffield Department of Medicine, Old Road Campus Research Building, Oxford OX3 7DQ, United Kingdom

<sup>d</sup> Target Discovery Institute, NDM Research Building, University of Oxford, Old Road Campus, Headington OX3 7FZ, United Kingdom

<sup>e</sup> Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22908, United States

### ARTICLE INFO

#### Article history:

Received 4 October 2016

Received in revised form 20 February 2017

Accepted 21 February 2017

Available online 27 February 2017

#### Keywords:

Biological image analysis  
Quantitative biological imaging  
Live cell imaging  
Cell segmentation  
Cell tracking  
Machine learning

### ABSTRACT

Advances in optical microscopy, biosensors and cell culturing technologies have transformed live cell imaging. Thanks to these advances live cell imaging plays an increasingly important role in basic biology research as well as at all stages of drug development. Image analysis methods are needed to extract quantitative information from these vast and complex data sets. The aim of this review is to provide an overview of available image analysis methods for live cell imaging, in particular required preprocessing image segmentation, cell tracking and data visualisation methods. The potential opportunities recent advances in machine learning, especially deep learning, and computer vision provide are being discussed. This review includes overview of the different available software packages and toolkits.

© 2017 Published by Elsevier Inc.

### Contents

1. Introduction	66
1.1. Imaging the cell	66
1.2. Large-scale imaging of the cell	66
1.3. Imaging-based phenotyping	67
1.4. Purpose and outline	67
2. Image preprocessing methods	68
2.1. Denoising and enhancing signals	68
2.2. Label-free imaging	68
3. Delineating objects of interest	68
3.1. Traditional approaches	69
3.2. Segmentation through clustering	69
3.3. Probabilistic segmentation	70
3.4. Learning models from image data	70
3.5. Open challenges	70
4. Cell tracking	71
4.1. Learning-based cell tracking	72
4.2. Global object association tracking	72
4.3. Open challenges	72
5. Software tools	72
5.1. Supporting end users	72

\* Corresponding authors at: Institute of Biomedical Engineering, Department of Engineering Science, Old Road Campus Research Building, University of Oxford, Oxford OX3 7DQ, United Kingdom.

E-mail address: [jens.rittischer@eng.ox.ac.uk](mailto:jens.rittischer@eng.ox.ac.uk) (J. Rittscher).

URL: <http://www.ibme.ox.ac.uk> (J. Rittscher).

5.2. Supporting algorithm developers .....	73
5.3. Creating integrated environments .....	73
5.4. Open challenges .....	73
6. Information visualisation .....	73
6.1. Viewing complex imaging data sets .....	74
6.2. Shape and appearance variation .....	74
6.3. Visualising phenotypic structures .....	74
7. Conclusion & future directions .....	75
Acknowledgements .....	75
References .....	75

## 1. Introduction

Starting with Antonie van Leeuwenhoek, who published his discovery of bacteria, blood cells and muscle fibres in a number of letters to London's Royal Society in the late 17th century, optical microscopy has become an indispensable tool for scientific discovery. Imaging cells and tissues using methods such as automated high-content and high-throughput microscopy have offered new insights into biology. This review clarifies the role imaging can play in gaining insights into biology and describes the extraction of quantitative information that will aid in this task. Quantitative imaging will be required at all anatomical scales ranging from the sub-cellular to the organ level, where structure, form, organization, and most importantly function, are essential for a complete characterization of the organism.

In this review we will focus on the sub-cellular and the cellular scales of living organisms. The characteristic structure and function at each scale is associated with cues manifest as biomolecules which are made to express near visible or visible electromagnetic energy endogenously, or are tagged appropriately to respond to optical stimuli. Optical instrumentation will then convert the expressed energy into a digital signal that will be analysed to glean structure and function for the specimen. Imaging reveals both structure and function of a single cell or a collection of cells.

### 1.1. Imaging the cell

The nucleus was the first cellular organelle to be discovered by a microscope. Leeuwenhoek observed a "lumen", the nucleus, in the red blood cells of salmon. Today, tagging or binding specific molecules is now the preferred way to delineate the nucleus. Although, histology has leveraged the use of hematoxylin and eosin stains to great advantage for capturing tissue architecture, that specific labelling technology is of limited value for live-cell imaging. Fluorescence microscopy is resorted to localize the binding location in the chromosome of a fluorescent probe; excitation through the leverage of fluorescence phenomena, in turn, delineates the nucleus. Selected examples are shown in Fig. 1. Thus, DAPI (4',6-diamidino-2-phenylindole) is a popular fluorescent stain that binds strongly to A-T nucleotide rich regions in the DNA. A variegated and sampled landscape can be obtained by using fluorescent in situ hybridization (FISH) probes which localize specific DNA and RNA level expression in the cell. The cytoskeleton composed of the water-rich cytoplasm and various filaments is imaged through the use of fluorescence. Phalloidin is often used by labelling it with fluorescent analogs and then subsequently staining actin filaments. Further emblematic of the adoption of this approach has been the discovery of the green fluorescent protein [1] and the wide-spread use of microscopy methods such as confocal and two-photon optics that have dramatically transformed live cell imaging. Given the leverage that the optical spectrum offers in

capturing multiple structures and functionalities (cellular mechanisms and signalling), the colour revolution has spurred the wide-spread use of imaging both for high-content and high-throughput [2].

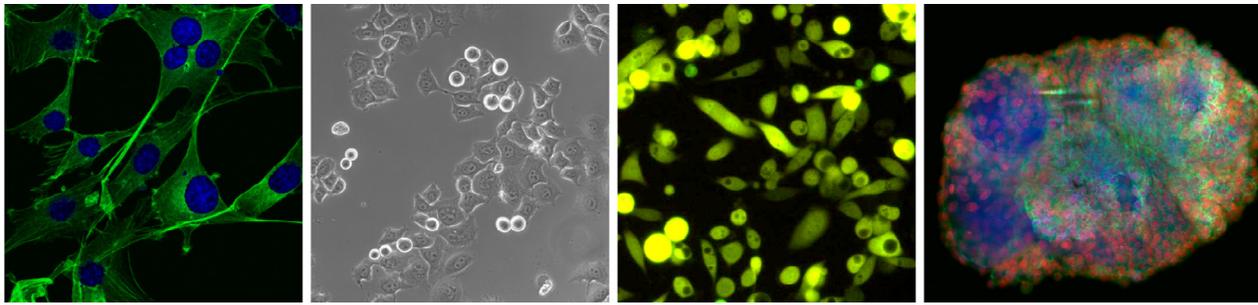
The desire to understand intercellular and intracellular processes has driven the development of super-resolved fluorescent microscopy [3,4]. To monitor organ formation or indeed the development of an entire organism in 4D (3D + t), the concept of *in toto* imaging [5,6] has been developed. With the help of advanced biosensors [7] it is now possible to report the activation states (e.g. conformation and phosphorylation) of endogenous proteins with minimal perturbation. In 2005 three different groups established Channelrhodopsin-2 (ChR2) as a tool for genetically targeted optical remote control [8–10]. Rapidly developing optogenetics techniques [11] now allow the fast and specific excitation and inhibition of proteins in complex cellular systems. However, high-content time lapse microscopy of living cells is still confined to the laboratory and its use is limited to gleaning cellular structure and function. There is a need to scale imaging experiments and methodologies.

### 1.2. Large-scale imaging of the cell

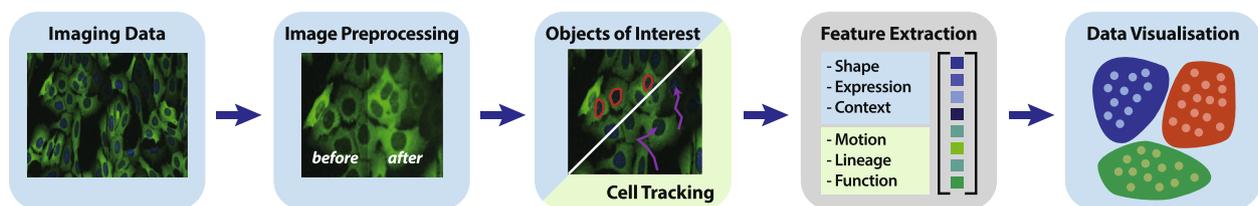
The concept of high-throughput screening was invented to address the needs of industrial and academic drug development efforts. Fairly basic cellular model systems were used to investigate specific molecular hypotheses. To overcome the limitations of this approach, the concept of phenotypic screening [12] was developed. New emerging developments in (patient derived) *ex vivo* cultures, induced pluripotent stem cells (iPSC) technology, three dimensional (3D) co-culture and organotypic systems hold the potential of designing more disease relevant model systems that will eventually replace traditional cell based models. Horvath et al. [13] recently published a comprehensive review that sets out the principles to facilitate the definition and development of disease relevant assays.

Optical microscopy platforms have evolved to support the growing demand for conducting more complex *in vitro* experiments. Advanced high-throughput microscopy platforms as for example the Cell Voyager 7000S (Yokogawa), Opera Phenix (Perkin Elmer) or the IN Cell Analyzer 6000 (GE Healthcare Life Sciences) offer the capability of acquiring three dimensional (3D) imagery over time. As a result it is now possible to acquire time lapse three dimensional data (3D + T) at scale. Sophisticated software tools help to transform vast amounts of complex multi-channel imaging data into quantitative information.

Thanks to these advances high-throughput cellular imaging is not only used in all stages of target based drug development [15] but is also becoming a relevant tool for investigating more fundamental biological questions. Experiments *in vitro* allow to monitor cell fate, build artificial tumors for studying the link between cancer and inflammation [16,17], and chart the interactions in the microenvironment to new drugs [18–20]. While it is probably



**Fig. 1. Biological image examples.** These examples illustrate some of the challenges that need to be taken into account when analysing biological imaging data. (Left) shows an image of cytoskeletal and nuclear staining of mouse 3T3 cells. Here it is challenging to assess the precise shape of each cell given the occlusion that occurs. (Middle-Left) This image depicts the artifacts associated with phase-contrast microscopy in the form of clumping which in turn prevents the determination of colonies. While this label free method clearly provides advantages for live cell imaging studies these artifacts make the quantitative analysis more challenging. (Middle-right, Right) These images illustrate that it is not always possible to segment individual cells. The image on the very right shows a spheroid of MCF7 cells in a matrix. The images have been generated in Department of Oncology at the University of Oxford.



**Fig. 2. Imaging based phenotyping.** A typical cellular analysis workflow is depicted. An acquired image is processed by correcting for noise and enhancing the signal. Next, cellular objects including nuclei, vasculature, etc. are extracted from the image. If the application demands, cells are tracked and their states are monitored. Features including shape, appearance, and context or trajectory and lineages are gleaned and models are learnt. Finally, the various features are visualized in a phenotypical feature space. This figure also motivates the main section of this review. Image processing methods are being reviewed in Section 2. Algorithms for delineating objects and structures of interest are reviewed in Section 3. Section 4 provides a review of cell tracking methods. The visualisation methods are being discussed in Section 6. For this illustration we used an image out of the set BBBC015v1 provided by Ilya Ravkin, available from the Broad Bioimage Benchmark Collection published in [14].

too early to judge if live cell imaging will have any significant role in clinical practice, it is safe to assume that there is a growing need for processing and analysing images of more complex 2D/3D co-culture models at scale.

### 1.3. Imaging-based phenotyping

The process of extracting meaningful information from image data (see Fig. 1) is part of an emerging field Houle and collaborators [21] termed phenomics. It is an area of computational biology concerned with the comprehensive high-dimensional measurement of phenomes. Houle et al. [21] distinguish between two different paradigms. The first is to take a large set of measurements at a given point in time, which they refer to as extensive phenotyping. In our context one could extensively characterise a given cell line with a number of different assays. Intensive phenotyping on the other would require characterising a given phenotype in great detail. Imaging specific processes using video microscopy [22,23] would fall into this category. For the purpose of illustration a notional imaging based phenotyping workflow is shown in Fig. 2.

While the difference between intrinsic and extrinsic phenotyping is less relevant to this review, it is the notion of taking measurements that is of fundamental importance. A good measurement should be accurate, robust and precise. Additional key elements [24] are the limit of detection, the response function and specificity. Given that there is noise in the measurement, the limit of detection defines the level below which the response is not meaningful. The response function specifies the dependence of the signal on systematic changes in experimental conditions. Good measurements are necessary for generating reproducible data. The choice of a specific cell segmentation algorithm can, for example, affect the interpretation of an experiment. The systematic analysis of various segmentation methods [25,26] documents how

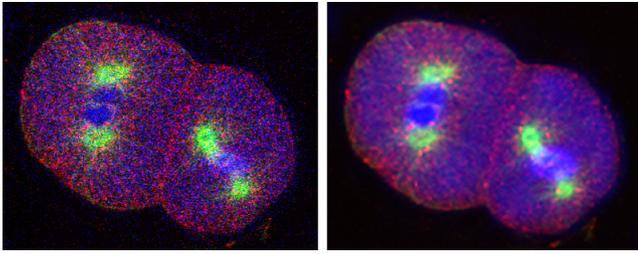
measurement statistics depend on algorithm choices and parameter settings.

The community of computer scientists, engineers and bioinformaticians that develops and advances mathematical methods and algorithms has grown substantially. Biological image analysis is now a broadly recognised area in leading international medical imaging and bioinformatics conferences. A number of challenge competitions, such as the “Particle Tracking Challenge” [27], “Cell Tracking Competition” [28] and Digital Reconstruction of Axonal and Dendritic Morphology Challenge (DIADEM) [29] have been initiated to advance the application specific algorithms. Related challenge competitions can be found on the *grand-challenge.org* web page.

A field which was started by a few enthusiasts (e.g. [30]) has now matured and produces very powerful algorithms that will continue to impact the life sciences. Thanks to fundamental methodological advances in image analysis, signal processing, medical imaging and computer vision the field will continue to evolve rapidly. Examples for one such development are advances in machine learning which will be discussed in this article. We expect that biomedical imaging will become a core part of the life science curriculum. Only with a certain understanding of the underlying methods it is possible to apply these in a thoughtful way to ensure studies produce reproducible results that ultimately help addressing key scientific questions. There is no doubt that microscopy has evolved from a technique of choice for producing stunning and impressive cover images for scientific publications to a technology that turns vast amounts of imaging data into quantitative information.

### 1.4. Purpose and outline

A number of review articles [2,31–34] have highlighted the opportunities biomedical imaging will provide. The term



**Fig. 3. Image denoising.** The example shows the non-redundant interscale wavelet thresholding method developed by Luisier and collaborators. It was specifically developed for Poisson data and presents a less computational intensive methods to other state of the art methods. Figure courtesy of: Florian Luisier, Biomedical Imaging Group, EPFL, Switzerland published in [38].

bio-image informatics evolved and is now used by some. Meijering [35] published a very comprehensive review on cell tracking methods. Dufour et al. [36] provide an overview of assessing 3D morphology. Here we strive to capture the impact more recent developments in computer vision and machine learning will have.

Enabled by machine learning computer vision has emerged as a key technology powering applications ranging from internet search to autonomous driving. The way in which we develop and design algorithms is changing dramatically. Rather than designing such methods from first principles, advanced machine learning techniques now allow us to learn computational models directly from the data. We will discuss how such deep architectures can be used to build reliable algorithms for biological imaging applications.

The workflow shown in Fig. 2 also provides a high level motivation for overall structure of this article. The analysis of shape and motion are the two central themes of this review. Only if we can delineate biological targets accurately, will it be possible to extract a set of meaningful and robust measurements. The segmentation of objects is discussed in Section 3. Methods for cell tracking are presented in Section 4. In the light of recent methodological developments, we believe that these are the areas which will advance most rapidly.

In many cases we cannot work on the raw images directly. Pre-processing methods play a central role in removing image noise and other artefacts. Given the broad range of microscopy methods, Section 2 does not permit a comprehensive review of pre-processing methods. Instead, more recent approaches for processing label-free microscopy images will be discussed.

In Section 5 we discuss some of the available software tools. Rather than limiting the discussion to software relevant to end users, we also provide an overview of toolkits to be used by algorithm developers. With larger data sets information visualisation starts to play a more prominent role. In Section 6 we discuss ideas of visualising the extracted data more effectively. Conclusions and directions for future work are being presented in Section 7.

## 2. Image preprocessing methods

Shading correction, removal of image noise and the suppression of out of focus light are perhaps the most important steps that would need to be addressed during image preprocessing. Given the focus of this review it is not possible to discuss the necessary calibration procedures that should be part of experimental protocols. Here we only highlight a few topics that should be taken into account during image processing. As the illumination across the field of view will not be uniform an explicit flat-field correction [37] is often necessary. Weak fluorescent staining or short exposure times can result in a low signal-to-noise ratio. In certain prac-

tical settings basic noise removal techniques such as median filtering or image smoothing will provide acceptable results.

### 2.1. Denoising and enhancing signals

A systematic characterisation of the various different noise sources [39] is necessary to achieve best results. More refined signal representations using wavelets [40] capable of providing multi-scale representation of images, or signals in general, gave rise to significant advances in de-noising methodologies. The general soft thresholding approach introduced by Donoho [41] needs to be mentioned in this context. Building on this approach, Luisier et al. [38] and Boulanger and colleagues [42] developed noise removal algorithms (see Fig. 3) that are specifically suited for microscopy applications. Additionally, there have been many fundamental contributions to methods that estimate both the background and foreground using non-local and multi-scale methods [43].

Deconvolution is an important and necessary preprocessing step that effects co-localisation analysis and image segmentation [39]. Wallace et al. [44] review the most common deconvolution methods. The problem can also be posed as a regularisation between the observed and an ideal that is modelled using the point spread function (PSF) of the microscope. Campisi and Egiazarian [45] present well-known taxonomy of methods that either systematically estimate the PSF or are blind to the exact form of the PSF.

### 2.2. Label-free imaging

Label free imaging methods such as phase-contrast, DIC or dark field microscopy provide some clear advantages as they do not require any sample preprocessing. However, processing such images can be challenging. For example, the varying appearance of cell boundaries under phase-contrast microscopy requires special consideration (see also Fig. 1). Su et al. [49] demonstrate that a physics based model allows to extract images features, so called phase retardation features, that enable a more robust processing of the data. An alternative is to apply machine learning techniques (e.g. [48]) for identifying structures such as cellular boundaries. Methods that are capable of extracting more quantitative information from brightfield images continue to improve. Joo et al. [46] developed a technique that can measure minute phase variations caused by changes in refractive index and thickness inside the specimen. Popescu and colleagues [47] propose the diffraction phase microscopy as a new technique for quantitative phase imaging of biological structures. These examples indicate that the collaboration between optical physicists and image analysis experts are necessary before the full promise of these techniques can be realised.

## 3. Delineating objects of interest

Digital images are represented as pixels associated with various intensity, or brightness values. Biologists however, are interested in objects such as cells, vesicles or tissue components such as blood vessels or glands. Image segmentation allows the identification of object boundaries which then can be used to quantify and analyse various attributes associated the objects of interest.

Despite the development of many segmentation methods over the last five decades image segmentation remains one of the most challenging image analysis tasks [31]. Even relatively minor changes in imaging conditions can require algorithm or parameter re-optimisation. Firstly, segmenting cellular images requires the identification of multiple objects in the image. Cells have heterogeneous shapes that are typically subject to dynamic changes. It is

therefore difficult to define shape in the form of mathematical models. Secondly, cell compartmentalization as well as intra- and inter-cell variability induces non-homogeneous marker distributions within and across cells, leading to undesirable image features such as intensity gradients. When analysing dense cell populations it is often very hard to assign features. While many splitting criteria for segmenting touching cells have been proposed they often only hold in specific settings. Unless a concrete objective or application context is given, the task of image segmentation is not a well defined problem.

In certain cases pixel based similarity measures are already sufficient for identifying objects of interest. One intuitive example is the identification of DAPI stained nuclei. Generally, purely data driven approaches are not sufficient for robust object delineation. Here, it is necessary to incorporate prior information through user interaction or mathematical models which are either designed manually or learnt from example data.

In many biological applications vast amounts of data are being generated and it is not feasible to rely on user input for the purpose of segmenting individual cells. Instead, it is necessary to design fully automated algorithms that are robust to changes in imaging conditions. Recent developments in computer vision and machine learning have helped in making significant progress towards this goal. In this review, more traditional methods are being referenced to point out what challenges need to be overcome. The remainder of this section focuses on more recent developments that will greatly impact our ability of developing robust algorithms to effectively process large data sets.

A discussion of traditional methods in Section 3.1 provides some historical context. These algorithms are now core components of most software packages (see Section 5) and are still used in many applications. Although there is no systematic theory of image segmentation, the concept of partitioning an image into segments by the means of clustering allows to present some of the more recent developments in a consistent fashion [50]. Modelling the image probabilistically is another important concept. Section 3.3 outlines how spatial relationships between pixels in the image can be modelled. This allows us to consider not only texture but also membrane boundaries which play an important role for delineating objects. Finally, we discuss some of the opportunities the emerging body of deep learning will offer. They provide an entirely different approach for learning computational models directly from imaging data but require a large set of annotated training images.

### 3.1. Traditional approaches

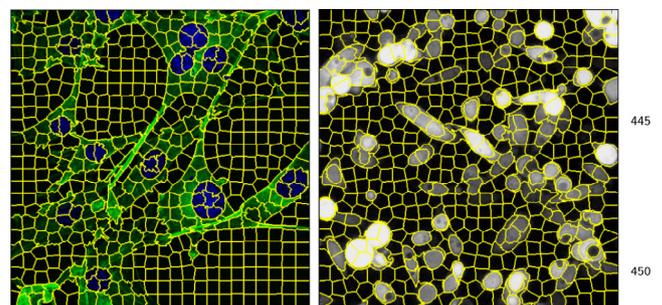
Many of the traditional image processing and computer vision methods [51] were developed for processing and analysing binary images. Meijering [52] provides a comprehensive review of the development of segmentation methods for biological imaging applications. Given that the basis for fluorescent microscopy is the use of differently coloured labels such images can be easily converted into binary images. At first sight, basic automated thresholding techniques (e.g [53]) combined with suitable pre-processing and post-processing appear to be a adequate solution. However, touching cells, image noise, inhomogeneous staining and uneven illumination are typical causes for segmentation errors. The field of mathematical morphology [54] was developed for analysing geometrical structures, based on set theory, lattice theory, topology and random functions. The well known watershed segmentation algorithm has been developed in this context. Watershed segmentations have been applied with considerable success to biomedical imaging [55–57]. The quality of watershed segmentations depends on selected seed points as well as suitable

image pre-processing steps. Generally, relatively minor changes in imaging conditions require further algorithm optimisation.

A recent review [58] highlights the impact active contour models continue to have on biological imaging. In general an active contour is a curve that evolves from some initial position towards the object of interest, a biological target. The contour evolution is governed by an energy function, called the snake energy. While we cannot provide a comprehensive review of the work on active contours, it should be mentioned that the development can be split into three broader topics. Point snakes denote the first category. Nominally, in active contour models [59] the curve is defined on the discrete pixel grid. The resulting representation unfortunately contains many parameters making it difficult to achieve robust performance. Based on the mathematical concept of implicit functions the idea of geodesic snakes [60–62] follows a more principled approach. The objective function captured by *snake energy* effectively controls the smoothness of the resulting contour and the model can deal with topological changes, which allows to segment highly complex objects. Unless constrained correctly geodesic contours can lead to overfitting. The optimisation process tends to be computationally expensive. Parametric snakes introduced through the work by Staib and Duncan [63] employ a parametric representation (e.g. B-splines) of the contour. As a result it is possible to design very fast algorithms which effectively incorporate shape priors. The idea of the *snakuscle* [64], a minuscule snake, takes this idea to an extreme. It enforces a circular shape and can be used for applications such as cell counting and vessel extraction. The use of active contour models is not limited to image segmentation. By initialising the contour on one image of a time lapse sequence and optimising it on the subsequent image, this method has been used effectively for cell tracking. Details will be discussed in Section 4.

### 3.2. Segmentation through clustering

The idea of grouping a set of pixels based on colour or local texture is a paradigm that is the basis of many image segmentation algorithms. The most basic example is the separation of background and foreground pixels. Various forms of clustering (see for example [66]) have been explored for generating image segments which are consistent in colour and texture. For characterising textures [67–69] and other structures [70,71] (e.g. dots of vessels) a vast number of low level image features have been developed. Methods that label each pixel site independently will not produce locally consistent image segments. The concept of superpixels addresses this problem by defining local clusters and limiting the search regions for assigning new pixels to a given region. One of the most successful methods proposed by Achanta et al. [65] reliably identifies regions which adhere well to object



**Fig. 4. Superpixels.** Superpixels provide a simple and effective method of generating locally consistent patches. The SLIC superpixel method [65] is used to illustrate the efficacy of this approach. Both images are results of the SLIC method after being deployed on images in Fig. 1. A set of 600 superpixels segment the images. This approach can also be used to make the processing of very large images more efficient.

boundaries and are efficient to compute. The examples shown in Fig. 4 illustrate that additional grouping is necessary before meaningful objects can be identified.

Similarities between pixel sites can be captured in form of an affinity matrix  $A$ . The set of pixel sites  $V$  together with the matrix can be used to define a graph  $G$ . The graph can now be cut into connected components that have relatively large interior weights by removing edges with low weights. Normalised cuts [72] segments an image by measuring the total dissimilarity between groups of pixels as well as the dissimilarity within groups. Cour et al. [73] provide a formulation that takes the edge information in an image into account. The iCut algorithm [74] illustrates how this approach can be applied successfully to biological imaging. As images grow in size the affinity matrices become prohibitively large. Defining the graph on a set of superpixels instead of the original pixel grid of the image is one approach to avoid this problem. Finally, a method inspired from the tracking literature has also demonstrated much practice. In [75], the graph cut algorithm is used in the context of tracking cells from image plane to image plane.

Yet another hybrid method was proposed in [76] where several methods were assembled to address the problem of segmentation of large microscopy images. Level sets, the geometry-inspired centroidal Voronoi tessellation (and  $K$ -Means), and local estimation of pixel intensity and texture were deployed in tandem to create 2D and 3D segmentations.

### 3.3. Probabilistic segmentation

By modelling the value at a given pixel as a random variable we can model images probabilistically. As images contain rich structure locally, the value of a given pixel depends on its neighbours. Markov Random Fields (MRFs) provide a formal framework for modelling complex probability distributions in form of a graph. Each node in the graph is associated with a random variable. Interactions between random variables are specified through edges in this graph. Such distributions are often expressed in terms of energy functions and clique potentials. Geman and Geman [77] originally introduced a MRF for reconstructing noisy images. Since then MRFs have been applied for a broad range of computer vision applications including image segmentation. Determining the maximum a posteriori estimate of such a model is an *NP-hard* problem. The originally proposed stochastic optimisation techniques are computationally expensive and slow.

By proposing a new class of energy minimisation algorithms based on graph cuts Boykov et al. [78,79] enabled the application of MRFs for a broad range of computer vision and image analysis task. This approach led not only to fast interactive segmentation methods for medical imaging [80], it also stimulated a number of successful segmentation methods for biological applications.

Building on the idea of graph cut based active contours [81] Chen and collaborators [82] developed a segmentation approach that is particularly suited by RNAi screens. Here, the information of different fluorescent channels is combined to obtain better object boundaries to aid image segmentation. The pattern-based cell segmentation approach proposed by Dimopoulos et al. [83] is fixated on the detection of membrane patterns; the approach effectively segments densely packed cells in an accurate manner. Probabilistic segmentation methods have been leveraged by Mosaliganti et al. [76] to enforce separation of overlapping cells.

### 3.4. Learning models from image data

Methods of machine learning play a vital role in a number of tasks including feature selection, classification and the discovery of latent structures. More recent developments in applying deep learning [84] to computer vision promise to have a very

fundamental impact on how we build and design algorithms for analysing biological image data. The fact that neural networks can be used to approximate almost any continuous function has been known for a while [85]. This theoretical result [86] even holds if the network consists of only a single intermediate layer between input and output neurons. While it illustrates the universal applicability of the approach, it does not provide any practical guidance for designing neural networks for specific problems.

Recently proposed convolutional networks for object recognition [87,88] and semantic segmentation [89] not only demonstrate that such approaches can outperform traditional methods, they also illustrate that it is possible to learn such models directly from raw image data. Data preprocessing, feature extraction and segmentation which were considered to be different steps in a traditional image analysis pipeline are now all integrated into one holistic computational model. Hence image analysis pipelines no longer depend on a set of user defined features.

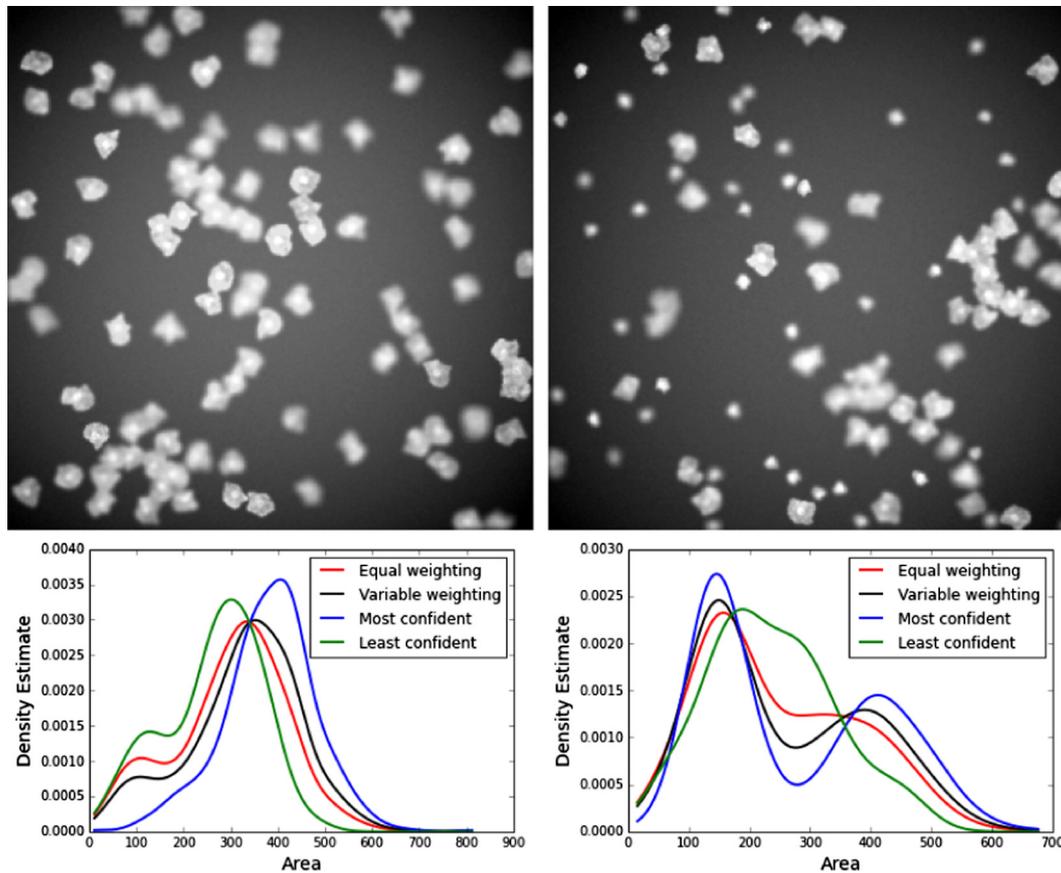
As layers can be fully connected these models explore a richer set of spatial interactions across the scales when compared to traditional MRFs. The price for this level of complexity is the fact that millions of free parameters need to be estimated from labelled training images. One solution is to train the model on a large corpus of natural images and then use transfer learning [90] to re-target the model onto a specific data set. Further, training models on simulated data [91] have shown extremely promising results. Here it would be possible to utilise earlier generative models for microscopy data [92,93].

Given the rapid development of the field, it is difficult to provide a comprehensive overview of deep learning to life science applications. Ning et al. [94] developed a automated phenotyping approach for *C. elegans* embryos obtained through DIC microscopy. A convolutional network is used to map raw pixels into output labels representing cell nuclei, nuclear membranes, cytoplasm and cell walls, and extracellular medium. Ciresian and collaborators [95] train a network for detecting membranes in electron microscopy data. Using data from the ISBI 2012 EM Segmentation Challenge they demonstrate that their approach outperforms competing techniques by a large margin with respect to different error metrics. One of the best scoring approaches in the Mitosis Detection Algorithms 2013 (AMIDA13) challenge included a convolutional network based architecture [96]. Ronneberger et al. [97] demonstrated how the same fully connected convolutional networks [98] can be applied towards the segmentation of electron microscopy and differential interference contrast imagery. Similarly, during the recently held CAMELYON challenge [99], which aims to identify histology pattern that are associated with metastatic breast cancer, many of the successful entries explored the use of CNNs and related recurring neural networks.

### 3.5. Open challenges

While these most recent advances open very exciting opportunities, we also need to advance methodologies that help to extract reliable data from imperfect segmentations. Overlapping and partially occluded cells will continue to cause uncertainty that even an apparently perfect segmentation approach will not be able to resolve. This problem will continue to confound the research community given the complications of occlusion that will arise when 3D objects are projected onto 2D imaging planes.

Mosaliganti and collaborators [76] present one approach where constraints are imposed based on ideal separation of nuclei. Subsequently, an objective function is optimised to determine the necessary separating plane. Another option is to discard data that does not satisfy certain assumptions. Meaburn et al. [101] present a systematic analysis of genome reorganization events during early tumorigenesis. Here a multi-stage classification [57] is being



**Fig. 5. Accounting for uncertainty in segmentation.** Nketia et al. estimate a confidence score for each of the identified candidate regions. This confidence score is then used to generate an overall estimate. Here the approach is illustrated with the help of two simulated images. The image on the left depicts a mono culture and the one on the right a mixed cell population. A comparison of the resulting distributions is shown below. Figure courtesy of: Thomas Nketia, Institute of Biomedical Engineering, University of Oxford published in [100].

applied to select a set of cells for analysis. Alternatively the inherent uncertainty of image segmentation can be taken into account. Nketia et al. [102] propose to estimate the confidence of each identified object and incorporate this confidence score in the measurement statistic with the help of a kernel density estimate [103]. An example on some simulated data is shown in Fig. 5. Further work is necessary to establish a more robust confidence metric for the resulting measurements.

In certain cases it is not necessary to segment the image. Arteta [104] developed a learning based framework that is designed to count cells in the presence of overlap. Here dot annotations are used to capture the required input from users. This approach has been extended to quantify *Drosophila* egg laying [105]. The accuracy of the original counting method is increased through correction of bias observed in the algorithm output.

Recently Ulmann et al. [106] also propose an interesting approach for a segmentation-free classification of images. Here over 900 features are being extracted from the image but only a subset of those accounting for 98 % of the variation are used for analysis. A linear discriminant analysis is used for classification. The method has been tested with success on the Broad Bioimage Benchmark Collection and data from the Human Protein Atlas [107].

While machine learning methods enhance our general ability for designing more robust algorithms, further work is needed to determine what type of preprocessing is required to ensure that these methods can be applied in a general setting. Here it might be possible to build on a set of tools that has been developed for processing histology images [108,109].

#### 4. Cell tracking

Cell tracking is essential in understanding the temporal dynamics of cell behaviour in time-lapse sequences. Cellular density in time lapse sequences obtained from typical biological experiments often tend to be high given the very closely-packed arrangement of cells in most frames. In addition, equipment and biological limitations including risk of photodamage reduce the frequency at which images can be captured over time. Consequently, this results in time lapse data with significantly low temporal resolution compared to that of conventional video. As a result, motile cells in the time series data seem to jump in a random fashion between consecutive frames showing minimal spatial overlap. These artefacts make cell tracking more complex and challenging than conventional object tracking in video [110]. The difficulties are further compounded by sudden changes in cell morphology over time [111].

Current methods for cell tracking build on the development of visual tracking algorithms for more general computer vision applications [112]. They can be broadly categorized into deformable models, state space models, and segmentation-based object association. In Section 3.1 deformable models were discussed in the context of image segmentation.

Deformable models involve a contour evolution approach, usually a level set or an active contour to obtain the boundary of an object in the current frame by evolving the contour from the previous frame [113–116]. In practice, it involves initializing the segmentation of objects in the first frame and updating in subsequent frames. Such contour evolution models fail in applications where

objects undergo sufficiently large displacements or show little overlap between consecutive frames [117]. These approaches also fail to detect cells that enter the field of view. Li et al. [115] address this by treating new cells using a local association scheme.

State space models for tracking often employ stochastic filtering techniques that rely on the underlying motion or appearance model of the tracked object. In this regard, the model expects the object to follow some assumed motion pattern but does not require an accurate segmentation of the tracked objects. They tend to be computationally demanding due to the large number of hypotheses pertaining to the motion-under-scrutiny and yet allow for complex observation models [118]. State space models are, however, better suited to handle larger displacements [119,120].

Segmentation based object association tracking models involve two major processes: object detection or segmentation in all frames and the association of objects in different frames to obtain a lineage. This allows cell tracking to be handled as two separate tasks; detection and association. Such association methods have been shown to be effective for cell tracking, to scale well and also achieve high accuracy in overall cell tracking [117,121–125]. The segmentation step varies widely depending on application or imaging modality.

Another method that could be loosely associated with the object association model is template matching with image registration between time points. The registration could be applied at the pixel level [126], at feature points [127], or for entire cellular objects [128]. The remainder of this section will focus on various approaches used in object association models as these have been most effective for cell tracking in time lapse phase contrast data [125,123,124]. Magnusson et al. [129] use the Viterbi algorithm to achieve a globally consistent cell tracks.

#### 4.1. Learning-based cell tracking

Tracking by assignment usually involves optimization features or objective functions which need to be adjusted for obtaining good assignments. Such parameters could be learned from training data by casting tracking as a learning problem [117]. Such methods have been applied to tracking pedestrians in conventional video with models derived from the Conditional Random Field (CRF) [130] and HybridBoost [131] techniques. The main challenge with learning based methods is obtaining accurate training data which usually involves costly annotations. In this regard, a cell labelling and active learning [132] have been applied to minimize the cost of annotation that is required in [117].

Lou and Hamprecht [117] present a more comprehensive learning-based formulation of the minimum cost flow theory method proposed by Padfield et al. [123] which is purposely built for cell tracking. The underlying model accounts for mitosis, appearance and disappearance of cells.

The approach *learns* the parameters associated with tracking events from training data to improve robustness and eliminate parameter tweaking. The learning of parameters hence provides a more robust approach to estimating model parameters. Such a learning-based approach however requires ground truth training data that contains frame-to-frame association pairs of all transition events for all cells. In the standard frame transition events, a cell may *move* or *divide*, and can *appear* or *disappear* from a frame. Accounting for appearing or disappearing cells is important as cells near the boundary of the field of view of the camera could move out and in depending on the direction of motion of the cells. Also, by introducing additional events *merge* and *split*, Lou et al. [117] account for the under-segmentation and over-segmentation errors typically observed in the detection step of closely-packed objects.

#### 4.2. Global object association tracking

As indicated earlier, the object association methods have shown to be effective in time-lapse. However, the challenge of accurately accounting for cells that leave or appear in a frame as well seeming merges and splits over many consecutive frames still remains. To resolve this, Bise et al. [124] proposes a global object association model rather than the local association methods involving association between consecutive frames only as in [123,117].

Global association methods associate objects over multiple frames by joining multi-frame trajectories. This could be achieved via *tracklet stitching* [133,134]. This involves first generating reliable fragments of object tracks (tracklets) and then merging these tracklets into more complete tracks. Connecting tracklets is achieved in [133] via the Hungarian algorithm [135] and by dynamic programming in [136].

The global association methods mentioned above are built for general object tracking and do not address specific cell behaviour such as cell division. The formulation proposed by Bise et al. [124] however accounts for division and possible cell segmentation errors. The tracklet association problem is formulated as a maximum-a posteriori (MAP) problem and solved by linear programming to obtain cell trajectories and lineage trees.

#### 4.3. Open challenges

Cell tracking methods have already enabled large scale biological studies that were previously not possible. The *Mitocheck consortium* [137] processed more than 190,000 time lapse movies of fluorescently labelled HeLa cells providing records of over 19 million cell divisions. This data was utilised to provide time resolved profiles of RNAi induced loss-of-function phenotypes resulting from siRNAs targeting the entire genome.

To study multi-generational differences between cerebral cortex neural progenitor cells Winter et al. [138] use phase contrast microscopy allowing image capture at a temporal resolution sufficient for accurate tracking through multiple rounds of cell division in a label-free manner. The cell tracking tool *Lever* [139] became the foundation of an entire research programme.

Extending these tracking approaches for analysing complex cellular behaviours is an open challenge. While the analysis of cell cycle phase [137,140,141] has been explored, other complex phenomena such as for example the various cellular events leading to cell death or collective cell migration have not been addressed. Another open an unexplored topic is also the multi-generational cellular response to drug compounds.

### 5. Software tools

Eliceiri et al. [33] provide a comprehensive overview of all the different software tools that are necessary for the implementation of an image informatics workflow. Here image acquisition, storage, analysis as well as image and data visualisation need to be taken into account. This section focuses on freely available open-source software tools. The following three sections provide an overview of software packages for end users, tool kits facilitating algorithm development and client server based environments. Open challenges and directions for future developments will be discussed in the final section.

#### 5.1. Supporting end users

Intuitive and well-supported software tools play a crucial role of providing end users access to this technology. Today, users can choose between very mature and sophisticated software tools.

Furthermore, users also benefit from support through tutorials and workshops which are provided by the developers themselves as well as through initiatives such as the Euro-BioImaging available on the website [www.eurobioimaging.eu](http://www.eurobioimaging.eu) and the Global-BioImaging project.

Undoubtedly ImageJ and the recently enhanced distribution of the same software, Fiji [142], are the oldest and most widely used tools for scientific image analysis. The software can be extended with help of a large number of plug-ins ranging from cell counting to co-localisation analysis. From the quick assessment of image quality to the scripted analysis of smaller image data sets it can be used for a wide range of tasks. As it is written in the language Java it can be easily deployed on all popular operating systems. While very versatile, some of the more advanced algorithms for segmentation and tracking are not available. It is probably less suited for the routine analysis of large data sets.

Cellprofiler [143] has been developed to facilitate the interactive data exploration, analysis, and classification of large biological image sets. Originally written in MATLAB, the software is now implemented in the computer language Python and includes a number of powerful features including a range of machine learning tools. Cellprofiler provides an open source alternative to commercially available high-content image analysis software packages such as Cellomics (ThermoFisher Scientific), Harmony (Perkin Elmer) or IN Cell Investigator (GE Healthcare Life Sciences). As an open source tool it provides the advantage that it can be extended very easily. However compared to other commercial tools it is not very well integrated with high-throughput microscopy hardware.

While Cellprofiler [144] is more geared towards 2D high-throughput screening data, Icy [144] provides a very comprehensive set of tools for the analysis of multi-channel 3D images. Developed and maintained by the biomedical image analysis group at the Institut Pasteur (Paris, France) it builds on a number of existing open source libraries and can also be extended through a broad range of plugins. The ilastic [145] platform is a simple and user-friendly tool for interactive image classification, segmentation and analysis. Notably, it provides access to some very advanced machine learning based algorithms for segmentation and tracking.

## 5.2. Supporting algorithm developers

Engineers and computer scientists rely on sophisticated software libraries and toolkits. These software tools provide important core capabilities which include the reading of application specific file formats and access to previously developed technology. In some sense they form the bases of extending the capabilities of the software packages discussed previously.

The National Library of Medicine Insight Segmentation and Registration Toolkit (ITK) [146,147] is the most widely used open source medical image analysis toolkit. It is designed to support *N*-dimensional images. Still, working and extending the C++ based libraries does require training. However, as the algorithms that have been included in the toolkit have been carefully validated, ITK made an invaluable contribution to the scientific community. SimpleITK [148] now provides a scripting interface to the underlying libraries, hence making them accessible in languages such as Python and R. The newest release now includes access to the registration algorithms which will be of interest to a number of groups. Farsight [149,150], which is a more specialised set of tools for the analysis of multi-channel fluorescent microscopy images is built directly on ITK.

Scripting languages not only provide the advantage of fast prototyping, they also allow the systematic integration of well established libraries for numerical computing, linear optimisation and now machine learning. With the increase in data volume it will be necessary to process data sets on dedicated compute clusters.

Python based environments allow one to make this transition. Possible performance bottlenecks can be eliminated through targeted optimisation. Python is increasingly viewed as the *lingua franca* of data analytics. It provides a great environment for developing analysis scripts that can be deployed on end user as well as specialised compute clusters. Rather than entering into a debate whether or not it should be used instead of statistics package based on the language R, we argue that these two languages, which are both distributed under open source licences, complement each other. With the Python interface of OpenCV [151] and the scikit-image libraries [152] Python provides a very rich set of image analysis tools. To date there is a lack of libraries that address microscopy specific analysis solutions. The Python Microscopy Environment (PyME) [153] is being developed to address this gap. It is targeted to high-resolution microscopy. In addition, all the underlying software modules provided by Cellprofiler can be accessed as Python libraries. SimpleITK can now also be installed directly as part of leading Python distributions. Python also provides access to popular deep learning libraries such as TensorFlow [154,155] and Theano [155].

## 5.3. Creating integrated environments

As data sets grow larger in size it is often no longer feasible to analyse the experimental data on a personal workstation or laptop. Environments such as OMERO [156] and Bisque [157], which were originally developed for storing and retrieving microscopy image data sets have now been extended to provide client-server software the visualisation, analysis and management of microscope images. Specific expertise and hardware is necessary for the setup and maintenance of these environments. Currently such capabilities will be limited to industrial laboratories and larger centres as they have the capacity to maintain and customise these solutions.

## 5.4. Open challenges

Today, it can still be difficult to reproduce image analysis experiments. A change in implementation details or parameter settings might actually result in different results. Provenance based methods are certainly par for the course. Further, standardization will be required to establish norms for analysis and discovery. It will be necessary to develop publications standards that are similar to those established for microarray analysis [158,159]. Furthermore the tools for processing large data sets need to be improved. What is already true for the processing of sequencing data today, will very likely become a reality for many future imaging studies. The necessary data will have to be hosted on large databases and will have to be processed on dedicated compute clusters. More intuitive and user friendly interfaces will have to be developed. In addition various analysis and visualisation capabilities will have to be integrated. Substantial further work will be needed to achieve this goal. But one day users who analyse their images using their favorite tools will routinely process large image collections on server hosted environments.

## 6. Information visualisation

Image analysis is only the first step in imaging studies that results in large and complex high dimensional datasets describing the phenomena-of-interest or the differences between biological samples. Visualisation plays an important role in understanding and analysing complex microscopy data. Dedicated software is needed to effectively view and inspect 3D time-lapse data that has been acquired in multiple channels. Segmentation and tracking results need to be displayed in context of the original data to

support algorithm development and analysis. Furthermore visualisation enables the identification of interesting phenotypes or artefacts, and the development statistical models of cellular shape and appearance. The effective visualisation of these datasets is crucial for identifying possible relationships and hypothesis in the data and choosing appropriate statistical or mathematical modelling methods.

Imaging information contributes to building quantitative models of the cell and cellular function. Including a detailed discussion of these methods is outside the scope of this review article. The work of Murphy and collaborators [160] and the Virtual Cell Project [161] serve as excellent examples for such efforts. In this context the recently NIH funded 4D Nucleome project should also be mentioned. It aims to advance our understanding of the principles underlying nuclear organization in space and time.

### 6.1. Viewing complex imaging data sets

Many tools have been developed to allow efficient exploration of complex imaging datasets from different angles. Walter et al. [162] provide a very comprehensive and broad overview of different visualisation methods for images. For example, *Icy* provides synchronised viewers to allow inspecting different time-points or locations of the image simultaneously [144]. Furthermore, the results of image analysis can be overlaid on raw images which enable validation of the analysis methods. Another useful tool for exploring large and complex 3D imaging data is *Volume3D* which utilises surface rendering to allow real-time visualisation of gigabytes-sized 3D imaging datasets on a typical laptop or a personal computer [163]. *Clear Volume* [164] is a dedicated open-source tool for visualising light sheet microscopy. The tools enables viewing of the data during image acquisition. 3D image data can also be streamed over the internet for remote viewing. *CellProfiler Tracer* allows visualisation of time-lapse data and exploration of the resulting cell trajectories, lineage tree as well as the progression of selected cells along time simultaneously [165]. These different tools aim to facilitate quantification of phenotypes in separate imaging datasets.

### 6.2. Shape and appearance variation

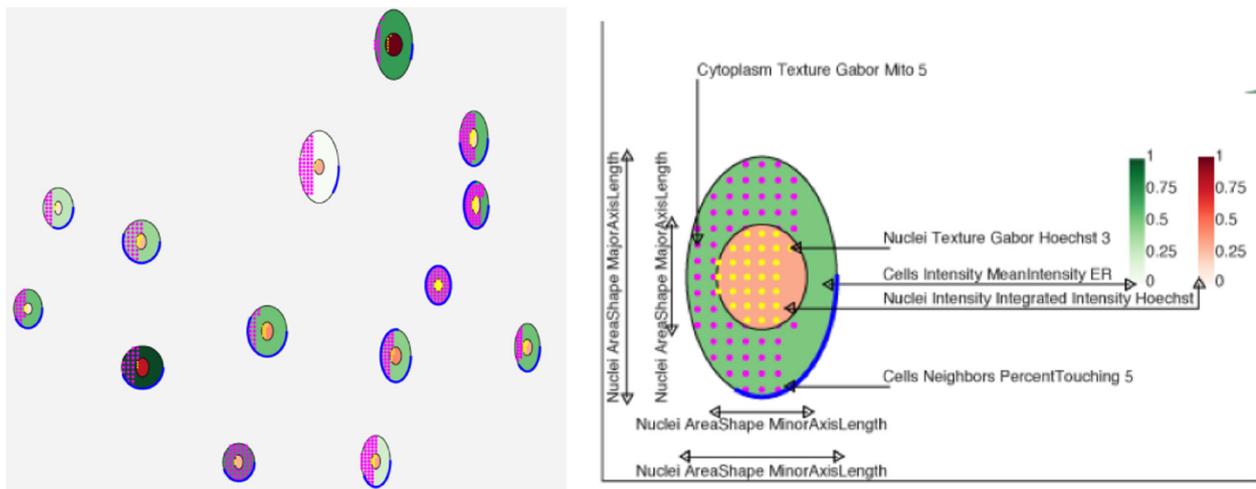
One goal of cell biology is to understand how cells adopt different shapes and expression profiles in response to varying environ-

mental and cellular conditions. A statistical analysis of the quantitative data derived from imaging studies is one first step towards advancing our understanding on how molecular processes govern cellular morphology. Here we need to differentiate between discriminative and generative models. The PhenoRipper software [167] was developed in the Altschuler and Wu laboratory with the aim to enable the rapid exploration of high-content microscopy images. PhenoRipper permits comparison of images obtained under different experimental conditions based on image phenotype similarity. Because the software analyses the images on a set of blocks rather than individual cells the analysis does not rely on an explicit cell segmentation method. The set of features is selected automatically. As a result of this analysis the software will map the data into a high-dimensional feature space and identify a number of classes. The benefit of this approach is that it hides the details of the machine learning procedure from the user.

Generative approaches can help to relate the features to interpretable visual models. Sailem and collaborators [168] present an approach for modelling cellular shape. Based on image derived features they employ Gaussian mixture modelling and hierarchical clustering for developing a graphical model that explains how the cross talk between Rac and Rho contribute to shape changes in wild-type *Drosophila* BG-2 cells. By formulating a transport based framework Wang et al. [169] develop a generative approach for modelling cellular shape and appearance directly in the image space. The transport based framework allows to address several tasks such as discriminating nuclear chromatin patterns in cancer cells, decoding differences in facial expressions, galaxy morphologies, as well as sub cellular protein distributions. It is for example possible to learn a visual dictionary exemplars and then compare the instances of this with new data points. In order to improve the representation an appearance manifold Sing et al. [170] learn a distance metric from labelled data. This metric is locally adaptive to account for heterogeneity in the data. This approach allows to analyse the heterogeneous expression patterns in cell nuclei.

### 6.3. Visualising phenotypic structures

The development of visualisation tools to address the nature of image-based datasets lags behind and is still largely restricted to the traditional plotting toolbox that includes scatter plots, parallel coordinates, and heat maps [171]. Visual analytics tools combine different traditional visualisation methods to enable the user to



**Fig. 6. PhenoPlots phenotypic measurements.** Phenoplots of 9 phenotypic measurements of the Human Bone Osteosarcoma Epithelial Cells from the *U2OS* cell line are shown. The position of cells reflects their X and Y coordinates. Image derived phenotypically relevant measurements are being mapped into pictorial structures to provide a more intuitive visualisation. This approach allows to visualise a set of phenotypic traits in form of a compact illustration. Mito: mitochondria, ER: Endoplasmic Reticulum. Figure courtesy of: Heba Sailem while at Insitute of Cancer Research, London published in [166].

interact with the data in order to drill down or zoom up different views of the data simultaneously [172,173]. Recently, few methods have been developed to address the specific needs in large biological datasets. For example, *glyph SPLOMs* allow efficient exploration of the associations and dependencies between tens of variables by representing dependency relationships as a certain number symbolic categories [174].

Another interesting method is *PhenoPlot* that provides pictorial elements to intuitively and quantitatively depict underlying phenotypic structures [166]. Such intuitive representations greatly facilitate interpreting the information in high dimension and easily relating it to the measured structures. The *PhenoPlot* representation of 9 variables describing the shape, texture, and intensity of different organelles in a given image of Human Bone Osteosarcoma Epithelial Cells from the U2OS cell lines [175] is shown in Fig. 6. Taylor and Noble [176] developed an approach for the interactive exploration of large image data sets. Here user generated measurements are utilised to arrange images virtual light table. This way the user can effectively associate images with meta data.

## 7. Conclusion & future directions

In recent years the field of biological imaging has grown significantly. The community has developed new approaches and tools that have become an integrative part of biological studies. Algorithms and methodologies continue to evolve. The recent advances in machine learning and computer vision, which allow learning capable and robust algorithms with the aid of deep learning directly from the data, have been highlighted in this review. Rather than setting up pipelines that utilise a number of algorithmic components, users can now learn new or adapt pre-trained models for certain tasks. Importantly, such an approach could reduce the number of parameters that would need to be determined by a user.

Increasingly, it will become necessary to process very large image data sets. Perhaps these should not be confused with the overused term *big data*. However, these data sets are sufficiently large to require dedicated server based solutions. Rather than struggling to process increasing large data sets on their own computers, users will store and process their data on dedicated compute clusters. Principles from bioinformatics will have to be adopted for allowing users to deal with such data sets in a transparent fashion.

The systematic integration of quantitative image analysis with mathematical models to obtain more biophysically relevant measurements is another important new research area. For example, Olivo-Marín and collaborators [177,178] propose a novel approach for measuring pressure and forces in the cell. Here the intracellular material is being represented as a 2D incompressible fluid and this model then acts as a control parameter for solving the classical optical flow equations. In general image derived data could be used for providing more realistic initial conditions for a mathematical model. In turn the underlying model could act as a prior for the extraction of image derived measurements. It is our contention that bioimage analysis and informatics will continue to play a very important role in gaining insights into the functioning and maintenance of living cells way beyond the relatively simple observations and methods of the early pioneers like van Leeuwenhoek.

## Acknowledgements

TAN received support from the RCUK Centre for Doctoral Training in Healthcare Innovation, RCUK Digital Economy Theme (EP/G036861/1), and the Wellcome Institutional Strategic Support Fund of the University of Oxford. EPSRC and the Wellcome

Institutional Strategic Support Fund of the University Of Oxford. JR and HS are supported by the EPSRC SeeBiByte Programme Grant (EP/M013774/1). In addition, JR is supported by the Ludwig Institute for Cancer Research.

## References

- [1] R.Y. Tsien, The green fluorescent protein, *Ann. Rev. Biochem.* 67 (1) (1998) 509–544.
- [2] C. Vonesch, F. Aguet, J.L. Vonesch, M. Unser, The colored revolution of bioimaging, *IEEE Signal Process. Mag.* 23 (3) (2006) 20–31.
- [3] E. Betzig, G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacio, M.W. Davidson, J. Lippincott-Schwartz, H.F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution, *Science* 313 (5793) (2006) 1642–1645, <http://dx.doi.org/10.1126/science.1127344>. URL <http://www.ncbi.nlm.nih.gov/pubmed/16902090>.
- [4] S.W. Hell, S.J. Sahl, M. Bates, X. Zhuang, R. Heintzmann, M.J. Booth, J. Bewersdorf, G. Shtengel, H. Hess, P. Tinnefeld, A. Honigsmann, S. Jakobs, I. Testa, L. Cognet, B. Lounis, H. Ewers, S.J. Davis, C. Eggeling, D. Klenerman, K.I. Willig, G. Vicidomini, M. Castello, A. Diaspro, T. Cordes, The 2015 super-resolution microscopy roadmap, *J. Phys. D: Appl. Phys.* 48 (44) (2015) 443001, <http://dx.doi.org/10.1088/0022-3727/48/44/443001>. URL <http://iopscience.iop.org/article/10.1088/0022-3727/48/44/443001>.
- [5] S.G. Megason, S.E. Fraser, Imaging in systems biology, *Cell* 130 (5) (2007) 784–795, <http://dx.doi.org/10.1016/j.cell.2007.08.031>.
- [6] S.G. Megason, In toto imaging of embryogenesis with confocal time-lapse microscopy, *Zebrafish: Methods Protoc.* (2009) 317–332.
- [7] J. Zhang, R.E. Campbell, A.Y. Ting, R.Y. Tsien, Creating new fluorescent probes for cell biology, *Nat. Rev. Mol. Cell Biol.* 3 (12) (2002) 906–918, <http://dx.doi.org/10.1038/nrm976>. URL <http://www.nature.com/doi/10.1038/nrm976>.
- [8] X. Li, D.V. Gutierrez, M.G. Hanson, J. Han, M.D. Mark, H. Chiel, P. Hegemann, L. T. Landmesser, S. Herlitz, Fast noninvasive activation and inhibition of neural and network activity by vertebrate rhodopsin and green algae channelrhodopsin, *Proc. Natl. Acad. Sci. U.S.A.* 102 (49) (2005) 17816–17821, <http://dx.doi.org/10.1073/pnas.0509030102>. URL <http://www.ncbi.nlm.nih.gov/pubmed/16306259>.
- [9] E.S. Boyden, F. Zhang, E. Bamberg, G. Nagel, K. Deisseroth, Millisecond-timescale, genetically targeted optical control of neural activity, *Nat. Neurosci.* 8 (9) (2005) 1263–1268, <http://dx.doi.org/10.1038/nn1525>. URL <http://www.ncbi.nlm.nih.gov/pubmed/16116447>.
- [10] G. Nagel, M. Brauner, J.F. Liewald, N. Adeishvili, E. Bamberg, A. Gottschalk, Light activation of Channelrhodopsin-2 in excitable cells of caenorhabditis elegans triggers rapid behavioral responses, *Curr. Biol.* 15 (24) (2005) 2279–2284, <http://dx.doi.org/10.1016/j.cub.2005.11.032>.
- [11] L. Fenno, O. Yizhar, K. Deisseroth, The development and application of optogenetics, *Ann. Rev. Neurosci.* 34 (1) (2011) 389–412, <http://dx.doi.org/10.1146/annurev-neuro-061010-113817>. URL <http://www.annualreviews.org/doi/10.1146/annurev-neuro-061010-113817>.
- [12] D.C. Swinney, J. Anthony, How were new medicines discovered?, *Nat. Rev. Drug Discovery* 10 (7) (2011) 507–519, <http://dx.doi.org/10.1038/nrd3480>. URL <http://www.ncbi.nlm.nih.gov/pubmed/21701501>.
- [13] P. Horvath, N. Aulner, M. Bickle, A.M. Davies, E.D. Nery, D. Ebner, M.C. Montoya, P. Östling, V. Pietiäinen, L.S. Price, S.L. Shorte, G. Turcatti, C. von Schantz, N.O. Carragher, Screening out irrelevant cell-based models of disease., *Nat. Rev. Drug Discovery*. doi:<http://dx.doi.org/10.1038/nrd.2016.175>. URL <http://www.ncbi.nlm.nih.gov/pubmed/27616293>
- [14] V. Ljosa, K.L. Sokolnicki, A.E. Carpenter, Annotated high-throughput microscopy image sets for validation, *Nat. Methods* 9 (7) (2012) 637.
- [15] P. Lang, K. Yeow, A. Nichols, A. Scheer, Cellular imaging in drug discovery, *Nat. Rev. Drug Discovery* 5 (April) (2006) 343–356, <http://dx.doi.org/10.1038/nrd2008>. URL <http://www.nature.com/doi/10.1038/nrd2008>.
- [16] A. Mantovani, P. Allavena, A. Sica, F. Balkwill, Cancer-related inflammation, *Nature* 454 (7203) (2008) 436–444, <http://dx.doi.org/10.1038/nature07205>. URL <http://www.nature.com/doi/10.1038/nature07205>.
- [17] S.M. Cruz, F.R. Balkwill, Inflammation and cancer: advances and new agents, *Nat. Rev. Clin. Oncol.* 12 (10) (2015) 1–13, <http://dx.doi.org/10.1038/nrclinonc.2015.105>. URL <http://www.nature.com/doi/10.1038/nrclinonc.2015.105>.
- [18] N. Sachs, H. Clevers, Organoid cultures for the analysis of cancer phenotypes, *Curr. Opin. Genet. Dev.* 24 (2014) 68–73, <http://dx.doi.org/10.1016/j.gde.2013.11.012>. URL <http://www.ncbi.nlm.nih.gov/pubmed/24657539>.
- [19] H.E. Francies, M.J. Garnett, What role could organoids play in the personalization of cancer treatment?, *Pharmacogenomics* 16 (14) (2015) 1523–1526, <http://dx.doi.org/10.2217/pgs.15.114>. URL <http://www.ncbi.nlm.nih.gov/pubmed/26485224>.
- [20] C. Pauli, L. Puca, J.M. Mosquera, B.D. Robinson, H. Beltran, M.A. Rubin, R.A. Rao, An emerging role for cytopathology in precision oncology, *Cancer Cytopathol.* 124 (3) (2016) 167–173, <http://dx.doi.org/10.1002/cncy.21647>. URL <http://www.ncbi.nlm.nih.gov/pubmed/26641771>.
- [21] D. Houle, D.R. Govindaraju, S. Ohmlot, Phenomics: the next challenge, *Nat. Rev. Genet.* 11 (12) (2010) 855–866, <http://dx.doi.org/10.1038/nrg2897>. URL <http://www.ncbi.nlm.nih.gov/pubmed/21085204>.

- [22] L. Ji, J. Lim, G. Danuser, Fluctuations of intracellular forces during cell protrusion, *Nat. Cell Biol.* 10 (12) (2008) 1393–1400, <http://dx.doi.org/10.1038/ncb1797>.
- [23] J.F. Dorn, G. Danuser, G. Yang, Computational processing and analysis of dynamic fluorescence image data, *Methods Cell Biol.* 85 (2008) 497–538, [http://dx.doi.org/10.1016/S0091-679X\(08\)85022-4](http://dx.doi.org/10.1016/S0091-679X(08)85022-4).
- [24] A.L. Plant, L.E. Locascio, W.E. May, P.D. Gallagher, Improved reproducibility by assuring confidence in measurements in biomedical research, *Nat. Methods* 11 (9) (2014) 895–898, <http://dx.doi.org/10.1038/nmeth.3076>.
- [25] A.A. Dima, J.T. Elliott, J.J. Filliben, M. Halter, A. Peskin, J. Bernal, M. Kociolek, M. C. Brady, H.C. Tang, A.L. Plant, Comparison of segmentation algorithms for fluorescence microscopy images of cells, *Cytometry A* 79 (7) (2011) 545–559, <http://dx.doi.org/10.1002/cyto.a.21079>. URL <http://www.ncbi.nlm.nih.gov/pubmed/21674772>.
- [26] P. Bajcsy, A. Cardone, J. Chalfoun, M. Halter, D. Juba, M. Kociolek, M. Majurski, A. Peskin, C. Simon, M. Simon, A. Vandecreme, M. Brady, Survey statistics of automated segmentations applied to optical imaging of mammalian cells, *BMC Bioinf.* 16 (1) (2015) 330, <http://dx.doi.org/10.1186/s12859-015-0762-2>. URL <http://www.biomedcentral.com/1471-2105/16/330>.
- [27] N. Chenouard, I. Smal, F. de Chaumont, M. Maška, I.F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, A.R. Cohen, W.J. Godínez, K. Rohr, Y. Kalaidzidis, L. Liang, J. Duncan, H. Shen, Y. Xu, K.E.G. Magnusson, J. Jaldén, H.M. Blau, P. Paul-Gilloteaux, P. Roudot, C. Kervrann, F. Waharte, J.-Y. Tinevez, S.L. Shorte, J. Willems, K. Celler, G.P. van Wezel, H.-W. Dan, Y.-S. Tsai, C. Ortiz de Solórzano, J.-C. Olivo-Marín, E. Meijering, Objective comparison of particle tracking methods, *Nat. Methods* 11 (3) (2014) 281–289, <http://dx.doi.org/10.1038/nmeth.2808>. URL <http://www.ncbi.nlm.nih.gov/pubmed/24441936>.
- [28] M. Maška, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbíola, T. España, S. Venkatesan, D.M.W. Balak, P. Karas, T. Bolcková, M. Streitová, C. Carthel, S. Coraluppi, N. Harder, K. Rohr, K.E.G. Magnusson, J. Jaldén, H.M. Blau, O. Dzyubachyk, P. Křížek, G.M. Hagen, D. Pastor-Escuredo, D. Jimenez-Carretero, M.J. Ledesma-Carbayo, A. Muñoz-Barrutia, E. Meijering, M. Kozubek, C. Ortiz-de Solórzano, A benchmark for comparison of cell tracking algorithms, *Bioinformatics* 30 (11) (2014) 1609–1617, <http://dx.doi.org/10.1093/bioinformatics/btu080>. URL <http://www.ncbi.nlm.nih.gov/pubmed/24526711>.
- [29] Y. Liu, *The diadem and beyond*, *Neuroinformatics* 9 (2) (2011) 99–102.
- [30] A.K. Jain, S.P. Smith, E. Backer, Segmentation of muscle cell pictures: a preliminary study, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI 2 (3) (1980) 232–242, <http://dx.doi.org/10.1109/TPAMI.1980.4767010>. URL <http://www.ncbi.nlm.nih.gov/pubmed/21868896>.
- [31] H. Peng, *Bioimage informatics: a new area of engineering biology*, *Bioinformatics* 24 (17) (2008) 1827–1836, <http://dx.doi.org/10.1093/bioinformatics/btn346>. URL <http://www.ncbi.nlm.nih.gov/pubmed/18603566>.
- [32] J. Rittscher, Characterization of biological processes through automated image analysis, *Ann. Rev. Biomed. Eng.* 12 (2010) 315–344, <http://dx.doi.org/10.1146/annurev-bioeng-070909-105235>. URL <http://www.ncbi.nlm.nih.gov/pubmed/20482277>.
- [33] K.W. Eliceiri, M.R. Berthold, I.G. Goldberg, L. Ibáñez, B.S. Manjunath, M.E. Martone, R.F. Murphy, H. Peng, A.L. Plant, B. Roysam, et al., Biological imaging software tools, *Nat. Methods* 9 (7) (2012) 697–710.
- [34] C. Kervrann, C.Ó.S. Sorzano, S.T. Acton, J.-C. Olivo-Marín, M. Unser, A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy, *IEEE J. Sel. Top. Signal Process.* 10 (1) (2016) 6–30.
- [35] E. Meijering, O. Dzyubachyk, I. Smal, Methods for cell and particle tracking, *Methods Enzymol.* 504 (2012) 183–200, <http://dx.doi.org/10.1016/B978-0-12-391857-4.00009-4>. URL <http://www.ncbi.nlm.nih.gov/pubmed/22264535>.
- [36] A.C. Dufour, T.-Y. Liu, C. Ducrocq, R. Tournemene, B. Cummings, R. Thibeaux, N. Guillen, A.O. Hero, J.-C. Olivo-Marín, Signal processing challenges in quantitative 3-D cell morphology: more than meets the eye, *IEEE Signal Process. Mag.* 32 (1) (2015) 30–40, <http://dx.doi.org/10.1109/MSP.2014.2359131>. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6975298>.
- [37] M.A. Model, Intensity calibration and shading correction for fluorescence microscopes, *Current Protoc. Cytometry Chapter 10* (2006), <http://dx.doi.org/10.1002/0471142956.cy1014s37>. Unit10.14., URL <http://www.ncbi.nlm.nih.gov/pubmed/18770832>.
- [38] F. Luisier, C. Vonesch, T. Blu, M. Unser, Fast interscale wavelet denoising of Poisson-corrupted images, *Signal Process.* 90 (2) (2010) 415–427, <http://dx.doi.org/10.1016/j.sigpro.2009.07.009>. URL <http://linkinghub.elsevier.com/retrieve/pii/S0165168409003016>.
- [39] L. Landmann, Deconvolution improves colocalization analysis of multiple fluorochromes in 3D confocal data sets more than filtering techniques, *J. Microsc.* 208 (Pt 2) (2002) 134–147. URL <http://www.ncbi.nlm.nih.gov/pubmed/12423263>.
- [40] S. Mallat, *A theory for multiresolution signal decomposition: The wavelet representation*, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 674–693.
- [41] D.L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inf. Theory* 41 (3) (1995) 613–627, <http://dx.doi.org/10.1109/18.382009>. arXiv:0611061v2, URL <http://ieeexplore.ieee.org/document/382009>.
- [42] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J.-B. Sibarita, J. Salamero, Patch-based nonlocal functional for denoising fluorescence microscopy image sequences, *IEEE Trans. Med. Imaging* 29 (2) (2010) 442–454, <http://dx.doi.org/10.1109/TMI.2009.2033991>. URL <http://www.ncbi.nlm.nih.gov/pubmed/19900849>.
- [43] C. Kervrann, C.O. Sanchez Sorzano, S.T. Acton, J.-C. Olivo-Marín, M. Unser, A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy, *IEEE J. Sel. Top. Signal Process.* 10 (1) (2016) 6–30, <http://dx.doi.org/10.1109/JSTSP.2015.2505402>.
- [44] W. Wallace, L.H. Schaefer, J.R. Swedlow, A workingperson's guide to deconvolution in light microscopy, *Biotechniques* 31 (5) (2001) 1076–1078. 1080, 1082 passim. URL <http://www.ncbi.nlm.nih.gov/pubmed/11730015>.
- [45] P. Campisi, K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*, CRC Press, 2016.
- [46] C. Joo, T. Akkin, B. Cense, B.H. Park, J.F. de Boer, Spectral-domain optical coherence phase microscopy for quantitative phase-contrast imaging, *Opt. Lett.* 30 (16) (2005) 2131–2133.
- [47] G. Popescu, T. Ikeda, R.R. Dasari, M.S. Feld, Diffraction phase microscopy for quantifying cell structure and dynamics, *Opt. Lett.* 31 (6) (2006) 775–777.
- [48] D.H. Theriault, M.L. Walker, J.Y. Wong, M. Betke, Cell morphology classification and clutter mitigation in phase-contrast microscopy images using machine learning, *Mach. Vis. Appl.* 23 (4) (2012) 659–673.
- [49] H. Su, Z. Yin, S. Huh, T. Kanade, Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features, *Med. Image Anal.* 17 (7) (2013) 746–765.
- [50] D.A. Forsyth, J. Ponce, *Computer vision: a modern approach*, Prentice Hall Professional Technical Reference, 2002.
- [51] R.C. Gonzales, P. Wintz, *Digital Image Processing*, Addison-Wesley, 1987.
- [52] E. Meijering, Cell segmentation: 50 years down the road [life sciences], *IEEE Signal Process. Mag.* 29 (5) (2012) 140–145.
- [53] N. Otsu, A threshold selection method from gray-scale histogram, *IEEE Trans. Systems Man Cybernetics* 8 (1978) 62–66.
- [54] J. Serra, P. Soille, *Mathematical Morphology and Its Applications to Image Processing*, vol. 2, Springer Science & Business Media, 2012.
- [55] G. Lin, U. Adiga, K. Olson, J.F. Guzowski, C.A. Barnes, B. Roysam, A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks, *Cytometry A* 56 (1) (2003) 23–36.
- [56] C. Wahlbl, I.M. Sintorn, F. Erlandsson, G. Borgfors, E. Bengtsson, Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections, *J. Microsc.* 215 (Pt 1) (2004) 67–76, <http://dx.doi.org/10.1111/j.0022-2720.2004.01338.x>.
- [57] K. Nandy, P.R. Gudla, R. Amundsen, K.J. Meaburn, T. Misteli, S.J. Lockett, Automatic segmentation and supervised learning-based selection of nuclei in cancer tissue images, *Cytometry A* 81 (9) (2012) 743–754.
- [58] R. Delgado-Gonzalo, V. Uhlmann, D. Schmitter, M. Unser, Snakes on a plane: a perfect snap for bioimage analysis, *IEEE Signal Process. Mag.* 32 (1) (2015) 41–48.
- [59] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, *Int. J. Comput. Vision* 1 (4) (1988) 321–331, <http://dx.doi.org/10.1007/BF00133570>. URL <http://link.springer.com/10.1007/BF00133570>.
- [60] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, *Int. J. Comput. Vision* 22 (1) (1997) 61–79.
- [61] R. Malladi, J.A. Sethian, B.C. Vemuri, Shape modeling with front propagation: a level set approach, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (2) (1995) 158–175.
- [62] K. Zhang, L. Zhang, H. Song, W. Zhou, Active contours with selective local or global segmentation: a new formulation and level set method, *Image Vis. Comput.* 28 (4) (2010) 668–676.
- [63] L.H. Staib, J.S. Duncan, Boundary finding with parametrically deformable models, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (11) (1992) 1061–1075, <http://dx.doi.org/10.1109/34.166621>. URL <http://ieeexplore.ieee.org/document/166621/>.
- [64] P. Thévenaz, M. Unser, Snakuscules, *IEEE Trans. Image Process.* 17 (4) (2008) 585–593, <http://dx.doi.org/10.1109/TIP.2007.914742>. URL <http://ieeexplore.ieee.org/document/4429291/>.
- [65] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Susstrunk, SLIC superpixels compared to state-of-the-art superpixel methods, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (11) (2012) 2274–2282, <http://dx.doi.org/10.1109/TPAMI.2012.120>. URL <http://ieeexplore.ieee.org/document/6205760/>.
- [66] C.M. Bishop et al., *Pattern Recognition and Machine Learning*, Springer New York, 2006.
- [67] R.M. Haralick, Statistical and structural approaches to texture, *Proc. IEEE* 67 (5) (1979) 786–804, <http://dx.doi.org/10.1109/PROC.1979.11328>. URL <http://ieeexplore.ieee.org/document/1455597/>.
- [68] A.K. Jain, F. Farrokhnia, Unsupervised texture segmentation using Gabor filters, *Pattern Recogn.* 24 (12) (1991) 1167–1186, [http://dx.doi.org/10.1016/0031-3203\(91\)90143-S](http://dx.doi.org/10.1016/0031-3203(91)90143-S).
- [69] W. Freeman, E. Adelson, The design and use of steerable filters, *IEEE Trans. Pattern Anal. Mach. Intell.* 13 (9) (1991) 891–906, <http://dx.doi.org/10.1109/34.93808>. URL <http://ieeexplore.ieee.org/document/93808/>.
- [70] D. Ballard, Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recogn.* 13 (2) (1981) 111–122, [http://dx.doi.org/10.1016/0031-3203\(81\)90009-1](http://dx.doi.org/10.1016/0031-3203(81)90009-1).
- [71] A.F. Frangi, W.J. Niessen, K.L. Vincken, M.A. Viergever, Multiscale vessel enhancement filtering 1 Introduction 2 Method, in: MICCAI, 1496, 1998, pp. 130–137.
- [72] J. Shi, J. Malik, Normalized cuts and image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (8) (2000) 888–905.

- [73] T. Cour, F. Benezit, J. Shi, Spectral segmentation with multiscale graph decomposition, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2, IEEE, 2005, pp. 1124–1131.
- [74] Y. He, H. Gong, B. Xiong, X. Xu, A. Li, T. Jiang, Q. Sun, S. Wang, Q. Luo, S. Chen, iCut: an integrative cut algorithm enables accurate segmentation of touching cells, *Sci. Rep.* 5.
- [75] T. Pécot, S. Singh, E. Caserta, K. Huang, R. Machiraju, G. Leone, Non parametric cell nuclei segmentation based on a tracking over depth from 3d fluorescence confocal images, in: 9th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2012, May 2–5, 2012, Barcelona, Spain, Proceedings, 2012, pp. 170–173. doi:<http://dx.doi.org/10.1109/ISBI.2012.6235511>.
- [76] K. Mosaliganti, L. Cooper, R. Sharp, R. Machiraju, G. Leone, K.J. Saltz, Reconstruction of cellular biological structures from optical microscopy data, *IEEE Trans. Visual Comput. Graphics* 14 (4) (2008) 863–876.
- [77] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian relation of images, *IEEE Trans. Pattern Anal. Mach. Intell.* 6 (1984) 721–741.
- [78] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (11) (2001) 1222–1239.
- [79] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *IEEE Trans. Pattern Anal. Mach. Intell.* (2004) 1124–1137.
- [80] Y. Boykov, M.-P. Jolly, Interactive organ segmentation using graph cuts, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2000, pp. 276–286.
- [81] N. Xu, N. Ahuja, R. Bansal, Object segmentation using graph cuts based active contours, *Comput. Vis. Image Understanding* 107 (3) (2007) 210–224. <http://dx.doi.org/10.1016/j.cviu.2006.11.004>. URL <http://www.sciencedirect.com/science/article/pii/S1077314206001792>.
- [82] C. Chen, H. Li, X. Zhou, S.T.C. Wong, Constraint factor graph cut-based active contour method for automated cellular image segmentation in RNAi screening, *J. Microsc.* 230 (2008) 177–191. <http://dx.doi.org/10.1111/j.1365-2818.2008.01974.x>. arXiv:NIHMS150003, URL <http://www.ncbi.nlm.nih.gov/pubmed/18445146>.
- [83] S. Dimopoulos, C.E. Mayer, F. Rudolf, J. Stelling, Accurate cell segmentation in microscopy images using membrane patterns, *Bioinformatics* 30 (18) (2014) 2644–2651.
- [84] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444.
- [85] M.A. Nielsen, Neural Networks and Deep Learning, 2015. URL <http://neuralnetworksanddeeplearning.com/>.
- [86] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* 2 (5) (1989) 359–366. [http://dx.doi.org/10.1016/0893-6080\(89\)90020-8](http://dx.doi.org/10.1016/0893-6080(89)90020-8).
- [87] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv Preprint arXiv:1409.1556.
- [88] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
- [89] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
- [90] C.H. Lampert, H. Nickisch, S. Harmeling, Learning to detect unseen object classes by between-class attribute transfer, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, 2009, pp. 951–958.
- [91] A. Gupta, A. Vedaldi, A. Zisserman, Synthetic data for text localisation in natural images, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [92] D. Svoboda, M. Kašík, M. Maška, J. Hubený, S. Stejskal, M. Zimmermann, On simulating 3d fluorescent microscope images, in: D. Svoboda (Ed.), Proceedings of the 12th International Conference on Computer Analysis of Images and Patterns, CAIP 07, Springer-Verlag, 2007, pp. 309–316.
- [93] D. Svoboda, M. Kozubek, S. Stejskal, Generation of digital phantoms of cell nuclei and simulation of image formation in 3d image cytometry, *Cytometry A* 75 (6) (2009) 494–509.
- [94] F. Ning, D. Delhomme, Y. LeCun, F. Piano, L. Bottou, P.E. Barbano, Toward automatic phenotyping of developing embryos from videos, *IEEE Trans. Image Process.* 14 (9) (2005) 1360–1371.
- [95] D. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Deep neural networks segment neuronal membranes in electron microscopy images, in: Proceedings of the Conference on Advances in Neural Information Processing Systems, 2012, pp. 2843–2851.
- [96] M. Veta, P.J. van Diest, S.M. Willems, H. Wang, A. Madabhushi, A. Cruz-Roa, F. Gonzalez, A.B.L. Larsen, J.S. Vestergaard, A.B. Dahl, D.C. Ciresan, J. Schmidhuber, A. Giusti, L.M. Gambardella, F.B. Tek, T. Walter, C.W. Wang, S. Kondo, B.J. Matuszewski, F. Precioso, V. Snell, J. Kittler, T.E. de Campos, A.M. Khan, N.M. Rajpoot, E. Arkoumani, M.M. Lacle, M.A. Viergever, J.P.W. Pluim, Assessment of algorithms for mitosis detection in breast cancer histopathology images, *Med. Image Anal.* 20 (1) (2015) 237–248. <http://dx.doi.org/10.1016/j.media.2014.11.010>. arXiv:1411.5825, URL <http://www.ncbi.nlm.nih.gov/pubmed/25547073>.
- [97] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 9351, Springer, 2015, pp. 234–241.
- [98] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
- [99] D. Wang, A. Khosla, R. Gargeya, H. Irshad, A.H. Beck, Deep learning for identifying metastatic breast cancer, arXiv Preprint arXiv:1606.05718.
- [100] T. Nketia, J. Rittscher, A. Noble, Utilizing phase retardation features to segment cells in phase contrast microscopy images, in: Conference in Medical Image Understanding and Analysis, MIUA 2014, 2014.
- [101] K.J. Meaburn, T. Misteli, Locus-specific and activity-independent gene repositioning during early tumorigenesis, *J. Cell Biol.* 180 (1) (2008) 39–50. <http://dx.doi.org/10.1083/jcb.200708204>.
- [102] T.A. Nketia, J.A. Noble, J. Rittscher, Towards quantifying the impact of cell boundary estimation on morphometric analysis for phenotypic screening, in: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), IEEE, 2015, pp. 781–784.
- [103] B.W. Silverman, Density estimation for statistics and data analysis, vol. 26, CRC Press, 1986.
- [104] C. Arteta, V. Lempitsky, J.A. Noble, A. Zisserman, Detecting overlapping instances in microscopy images using extremal region trees, *Med. Image Anal.* 27 (2016) 3–16.
- [105] D. Waithe, P. Rennert, G. Brostow, M.D. Piper, Quantify: robust trainable software for automated drosophila egg counting, *PLoS ONE* 10 (5) (2015) e0127659.
- [106] V. Uhlmann, S. Singh, A.E. Carpenter, CP-CHARM: segmentation-free image classification made accessible, *BMC Bioinf.* 17. doi:<http://dx.doi.org/10.1186/s12859-016-0895-y>.
- [107] M. Uhlen, P. Oksvold, L. Fagerberg, E. Lundberg, K. Jonasson, M. Forsberg, M. Zwahlen, C. Kampf, K. Wester, S. Hober, et al., Towards a knowledge-based human protein atlas, *Nat. Biotechnol.* 28 (12) (2010) 1248–1250.
- [108] P.U. Adiga, B. Chaudhuri, Region based techniques for segmentation of volumetric histo-pathological images, *Comput. Methods Programs Biomed.* 61 (1) (2000) 23–47.
- [109] A.M. Khan, N. Rajpoot, D. Treanor, D. Magee, A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution, *IEEE Trans. Biomed. Eng.* 61 (6) (2014) 1729–1738.
- [110] M. Dewan, M. Ahmad, M. Swamy, Tracking biological cells in time-lapse microscopy: an adaptive technique combining motion and topological features, *IEEE Trans. Biomed. Eng.* 58 (6) (2011) 1637–1647.
- [111] K. Thirusittampalam, M. Hossain, O. Ghita, P. Whelan, A novel framework for cellular tracking and mitosis detection in dense phase contrast microscopy images, *IEEE J. Biomed. Health Inf.* 17 (3) (2013) 642–653.
- [112] A.W. Smeulders, D.M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, M. Shah, Visual tracking: an experimental survey, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (7) (2014) 1442–1468.
- [113] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillén-Aghion, J.-C. Olivo-Marin, C. Zimmer, Segmenting and tracking fluorescent cells in dynamic 3-d microscopy with coupled active surfaces, *IEEE Trans. Image Process.* 14 (9) (2005) 1396–1410.
- [114] F. Yang, M. Mackey, F. Ianzini, G. Gallardo, M. Sonka, Cell segmentation, tracking, and mitosis detection using temporal context, *Medical Image Computing and Computer Assisted Intervention MICCAI 3749* (2005) 302–309.
- [115] K. Li, E.D. Miller, M. Chen, T. Kanade, L.E. Weiss, P.G. Campbell, Cell population tracking and lineage construction with spatiotemporal context, *Med. Image Anal.* 12 (5) (2008) 546–566.
- [116] O. Dzyubachyk, W. Van Cappellen, J. Essers, W. Niessen, E. Meijering, Advanced level-set-based cell tracking in time-lapse fluorescence microscopy, *IEEE Trans. Med. Imaging* 29 (3) (2010) 852–867.
- [117] X. Lou, F. Hamprecht, Structured learning for cell tracking, in: Proceedings of the Conference on Advances in Neural Information Processing Systems.
- [118] E. Meijering, O. Dzyubachyk, I. Smal, W.A. van Cappellen, Tracking in cell and developmental biology, *Semin. Cell Dev. Biol.* 20 (8) (2009) 894–902. imaging in Cell and Developmental Biology Planar Cell Polarity.
- [119] W.J. Godinez, M. Lampe, R. Eils, B. Müller, K. Rohr, Tracking multiple particles in fluorescence microscopy images via probabilistic data association, in: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, 2011, pp. 1925–1928.
- [120] T.-C. Ku, Y.-N. Huang, C.-C. Huang, D.-M. Yang, L.-S. Kao, T.-Y. Chiu, C.-F. Hsieh, P.-Y. Wu, Y.-S. Tsai, C.-C. Lin, An automated tracking system to measure the dynamic properties of vesicles in living cells, *Microsc. Res. Tech.* 70 (2) (2007) 119–134.
- [121] X. Lou, F.O. Kaster, M.S. Lindner, B.X. Kausler, U. Kothe, B. Hockendorf, J. Wittbrodt, H. Janicke, F. Hamprecht, Deltr: Digital embryo lineage tree reconstructor, *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* 2011. pp. 1557–1560.
- [122] F. Li, X. Zhou, J. Ma, S. Wong, Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis, *IEEE Trans. Med. Imaging* 29 (1) (2010) 96–105.
- [123] D. Padfield, J. Rittscher, B. Roysam, Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis, *Med. Image Anal.* 15 (4) (2011) 650–668. <http://dx.doi.org/10.1016/j.media.2010.07.006>. URL <http://www.ncbi.nlm.nih.gov/pubmed/20864383>.
- [124] R. Bise, Z. Yin, T. Kanade, Reliable cell tracking by global data association, *Biomedical Imaging: From Nano to Macro, IEEE International Symposium on* 2011. pp. 1004–1010.

- [125] T. Kanade, Z. Yin, R. Bise, S. Huh, S. Eom, M. Sandbothe, M. Chen, Cell image analysis: algorithms, system and applications, in: Applications of Computer Vision (WACV), 2011 IEEE Workshop on, 2011, pp. 374–381.
- [126] A. Hand, T. Sun, D. Barber, D. Hose, S. Macneil, Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration, *J. Microsc.* 234 (1) (2009) 62–79.
- [127] P. Matula, P. Matula, M. Kozubek, V. Dvorak, Fast point-based 3-d alignment of live cells, *IEEE Trans. Image Process.* 15 (8) (2006) 2388–2396.
- [128] C. Wilson, J. Theriot, A correlation-based approach to calculate rotation and translation of moving cells, *IEEE Trans. Image Process.* 15 (7) (2006) 1939–1951.
- [129] K.E. Magnusson, J. Jaldén, P.M. Gilbert, H.M. Blau, Global linking of cell tracks using the viterbi algorithm, *IEEE Trans. Med. Imaging* 34 (4) (2015) 911–929.
- [130] B. Yang, R. Nevatia, An online learned crf model for multi-target tracking, in: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 2034–2041.
- [131] Y. Li, C. Huang, R. Nevatia, Learning to associate: Hybridboosted multi-target tracker for crowded scene, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 2953–2960.
- [132] X. Lou, M. Schiegg, F. Hamprecht, Active structured learning for cell tracking: algorithm, framework, and usability, *IEEE Trans. Med. Imaging* 33 (4) (2014) 849–860.
- [133] C. Huang, B. Wu, R. Nevatia, Robust object tracking by hierarchical association of detection responses, in: Computer Vision—ECCV 2008, Springer, 2008, pp. 788–801.
- [134] C. Huang, Y. Li, R. Nevatia, Multiple target tracking by learning-based hierarchical association of detection responses, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (4) (2013) 898–910.
- [135] H.W. Kuhn, The Hungarian method for the assignment problem, *Nav. Res. Logist. Q.* 2 (1–2) (1955) 83–97.
- [136] S. Bonneau, M. Dahan, L. Cohen, Single quantum dot tracking based on perceptual grouping using minimal paths in a spatiotemporal volume, *IEEE Trans. Image Process.* 14 (9) (2005) 1384–1395.
- [137] B. Neumann, T. Walter, J.-K. Hériché, J. Bulkescher, H. Erfle, C. Conrad, P. Rogers, I. Poser, M. Held, U. Liebel, C. Cetin, F. Sieckmann, G. Pau, R. Kabbe, A. Wünsche, V. Satagopam, M.H.A. Schmitz, C. Chapuis, D.W. Gerlich, R. Schneider, R. Eils, W. Huber, J.-M. Peters, A.A. Hyman, R. Durbin, R. Pepperkok, J. Ellenberg, Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes, *Nature* 464 (7289) (2010) 721–727, <http://dx.doi.org/10.1038/nature08869>. URL <http://www.nature.com/doi/10.1038/nature08869>.
- [138] M. Winter, M. Liu, D. Montealeone, J. Melunis, U. Hershberg, S. Goderie, S. Temple, A. Cohen, Computational image analysis reveals intrinsic multigenerational differences between anterior and posterior cerebral cortex neural progenitor cells, *Stem Cell Rep.* 5 (4) (2015) 609, <http://dx.doi.org/10.1016/j.stemcr.2015.08.002>.
- [139] M. Winter, W. Mankowski, E. Wait, S. Temple, A.R. Cohen, Lever: software tools for segmentation, tracking and lineaging of proliferating cells, *Bioinformatics* (2016) btw406.
- [140] D. Padfield, J. Rittscher, N. Thomas, B. Roysam, Spatio-temporal cell cycle phase analysis using level sets and fast marching methods, *Med. Image Anal.* 13 (1) (2009) 143–155, <http://dx.doi.org/10.1016/j.media.2008.06.018>. URL <http://www.ncbi.nlm.nih.gov/pubmed/18752984>.
- [141] A. El-Labban, A. Zisserman, Y. Toyoda, A.W. Bird, A. Hyman, Temporal models for mitotic phase labelling, *Med. Image Anal.* 18 (7) (2014) 977–988.
- [142] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (7) (2012) 676–682, <http://dx.doi.org/10.1038/nmeth.2019>. URL <http://www.ncbi.nlm.nih.gov/pubmed/22743772>.
- [143] A.E. Carpenter, T.R. Jones, M.R. Lamprecht, C. Clarke, I.H. Kang, O. Friman, D.A. Guertin, J.H. Chang, R.A. Lindquist, J. Moffat, P. Golland, D.M. Sabatini, Cell Profiler: image analysis software for identifying and quantifying cell phenotypes, *Genome Biol.* 7 (10) (2006) R100, <http://dx.doi.org/10.1186/gb-2006-7-10-r100>. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1794559&tool=pmcentrez&rendertype=abstract>.
- [144] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, J.-C. Olivo-Marin, Icy: an open bioimage informatics platform for extended reproducible research, *Nat. Methods* 9 (7) (2012) 690–696, <http://dx.doi.org/10.1038/nmeth.2075>. URL <http://www.ncbi.nlm.nih.gov/pubmed/22743774>.
- [145] C. Haubold, M. Schiegg, A. Kreshuk, S. Berg, U. Koethe, F.A. Hamprecht, Segmenting and tracking multiple dividing targets using ilastik, *Adv. Anat. Embryol. Cell Biol.* 219 (2016) 199–229, <http://dx.doi.org/10.1007/978-3-319-28549-88>. URL <http://www.ncbi.nlm.nih.gov/pubmed/27207368>.
- [146] T.S. Yoo, M.J. Ackerman, W.E. Lorensen, W. Schroeder, V. Chalana, S. Aylward, D. Metaxas, R. Whitaker, Engineering and algorithm design for an image processing API: a technical report on ITK-The Insight Toolkit, *Stud. Health Technol. Inf.* (2002) 586–592.
- [147] K. Martin, L. Ibáñez, L. Avila, S. Barré, J.H. Kaspersen, Integrating segmentation methods from the Insight Toolkit into a visualization application, *Med. Image Anal.* 9 (6) (2005) 579–593, <http://dx.doi.org/10.1016/j.media.2005.04.009>. URL <http://www.ncbi.nlm.nih.gov/pubmed/16185910>.
- [148] B.C. Lowekamp, D.T. Chen, L. Ibáñez, D. Blezek, The Design of SimpleITK, *Front. Neuroinf.* 7 (2013) 45, <http://dx.doi.org/10.3389/fninf.2013.00045>. URL <http://www.ncbi.nlm.nih.gov/pubmed/24416015>.
- [149] B. Roysam, W. Shain, C. Barnes, W. Mohler, G. Lin, C. Björnsson, Y. Chen, Y. Al-Kofahi, A. Narayanaswamy, The FARSIGHT project: associative multi-dimensional image analysis methods for optical microscopy, in: J. Rittscher, R. Machiraju, S.T. Wong (Eds.), *Microscopic Image Analysis for Life Science Applications*, Artech House, Oxford, 2008, pp. 115–152. Ch. 5.
- [150] J. Luisi, A. Narayanaswamy, Z. Galbreath, B. Roysam, The FARSIGHT trace editor: an open source tool for 3-D inspection and efficient pattern analysis aided editing of automated neuronal reconstructions, *Neuroinformatics* 9 (2–3) (2011) 305–315, <http://dx.doi.org/10.1007/s12021-011-9115-0>. URL <http://www.ncbi.nlm.nih.gov/pubmed/21487683>.
- [151] G. Bradski, A. Kaehler, *Learning OpenCV: computer vision with the OpenCV library*, O'Reilly Media Inc, 2008.
- [152] S. van der Walt, J.L. Schönberger, F. Nunez-Iglesias, F. Boulogne, J.D. Warner, N. Yager, E. Goullart, T. Yu, scikit-image contributors, Scikit-image: image processing in Python, *PeerJ* 2 (2014) e453, <http://dx.doi.org/10.7717/peerj.453>. URL <http://www.ncbi.nlm.nih.gov/pubmed/25024921>.
- [153] D.J. Crossman, Y. Hou, I. Jayasinghe, D. Baddeley, C. Soeller, Combining confocal and single molecule localisation microscopy: a correlative approach to multi-scale tissue imaging, *Methods* 88 (2015) 98–108, <http://dx.doi.org/10.1016/j.jmeth.2015.03.011>.
- [154] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, et al., Tensorflow: Large-scale machine learning on heterogeneous distributed systems, *arXiv Preprint arXiv:1603.04467*.
- [155] L. Rampasek, A. Goldenberg, Tensorflow: biology's gateway to deep learning?, *Cell systems* 2 (1) (2016) 12–14.
- [156] C. Allan, J.-M. Burel, J. Moore, C. Blackburn, M. Linkert, S. Loynton, D. Macdonald, W.J. Moore, C. Neves, A. Patterson, M. Porter, A. Tarkowska, B. Lorange, J. Avondo, I. Lagerstedt, L. Lianas, S. Leo, K. Hands, R.T. Hay, A. Patwardhan, C. Best, G.J. Kleywegt, G. Zanetti, J.R. Swedlow, OMERO: flexible, model-driven data management for experimental biology, *Nat. Methods* 9 (3) (2012) 245–253, <http://dx.doi.org/10.1038/nmeth.1896>. URL <http://www.ncbi.nlm.nih.gov/pubmed/22373911>.
- [157] K. Kvilekval, D. Fedorov, B. Obara, A. Singh, B.S. Manjunath, Bisque: a platform for bioimage analysis and management, *Bioinformatics* 26 (4) (2010) 544–552, <http://dx.doi.org/10.1093/bioinformatics/btp699>. URL <http://www.ncbi.nlm.nih.gov/pubmed/20031971>.
- [158] A. Brazma, P. Hingamp, J. Quackenbush, G. Sherlock, P. Spellman, C. Stoeckert, J. Aach, W. Ansorge, C.A. Ball, H.C. Causton, et al., Minimum information about a microarray experiment (miame)—toward standards for microarray data, *Nat. Genet.* 29 (4) (2001) 365–371.
- [159] A. Brazma, Minimum information about a microarray experiment (miame)—successes, failures, challenges, *Sci. World J.* 9 (2009) 420–423.
- [160] T.E. Buck, J. Li, G.K. Rohde, R.F. Murphy, Toward the virtual cell: automated approaches to building models of subcellular organization learned from microscopy images, *Bioessays* 34 (9) (2012) 791–799.
- [161] B.M. Slepchenko, J.C. Schaff, I. Macara, L.M. Loew, Quantitative cell biology with the virtual cell, *Trends Cell Biol.* 13 (11) (2003) 570–576.
- [162] T. Walter, D.W. Shattuck, R. Baldock, M.E. Bastin, A.E. Carpenter, S. Duce, J. Ellenberg, A. Fraser, N. Hamilton, S. Pieper, M.A. Ragan, J.E. Schneider, P. Tomancak, J.K. Hériché, Visualization of image data from cells to organisms, *Nat. Methods* 7 (3 Suppl) (2010) S26–S41, <http://dx.doi.org/10.1038/nmeth.1431>. URL <http://www.nature.com/doi/10.1038/nmeth.1431>.
- [163] H. Peng, A. Bria, Z. Zhou, G. Iannello, F. Long, Extensible visualization and analysis for multidimensional images using Vaa3D., *Nat. Protoc.* 9(1) (2014) 193–208. doi:<http://dx.doi.org/10.1038/nprot.2014.011>. URL <http://www.nature.com/nprot/journal/v9/n1/full/nprot.2014.011.html>.
- [164] L.A. Royer, M. Weigert, U. Günther, N. Maghelli, F. Jug, I.F. Sbalzarini, E.W. Myers, ClearVolume: open-source live 3D visualization for light-sheet microscopy, *Nat. Methods* 12 (6) (2015) 480–481, <http://dx.doi.org/10.1038/nmeth.3372>. URL <http://www.nature.com/doi/10.1038/nmeth.3372>.
- [165] M.-A. Bray, A.E. Carpenter, Cell Profiler Tracer: exploring and validating high-throughput, time-lapse microscopy image data, *BMC Bioinf.* 16 (2015) 368, <http://dx.doi.org/10.1186/s12859-015-0759-x>. URL <http://www.ncbi.nlm.nih.gov/pubmed/26537300>.
- [166] H.Z. Sailem, J.E. Sero, C. Bakal, Visualizing cellular imaging data using Phenoplots, *Nat. Commun.* 6 (2015) 1–6, <http://dx.doi.org/10.1038/ncomms6825>.
- [167] S. Rajaram, B. Pavie, L.F. Wu, S.J. Altschuler, Phenoripper: software for rapidly profiling microscopy images, *Nat. Methods* 9 (7) (2012) 635–637.
- [168] H. Sailem, V. Bousgouni, S. Cooper, C. Bakal, Cross-talk between rho and rac gtpases drives deterministic exploration of cellular shape space and morphological heterogeneity, *Open Biol.* 4(1). doi:<http://dx.doi.org/10.1098/rsob.130132>. URL <http://rsob.royalsocietypublishing.org/content/4/1/130132>.
- [169] W. Wang, D. Slepčev, S. Basu, J.A. Ozolek, G.K. Rohde, A linear optimal transportation framework for analyzing and visualizing variations in sets of images, *Int. J. Comput. Vision* 101 (2) (2013) 254–269.
- [170] S. Singh, F. Janoos, T. Pécot, E. Caserta, G. Leone, J. Rittscher, R. Machiraju, Identifying nuclear phenotypes using semi-supervised metric learning, in: G. Székely, H.K. Hahn (Eds.), *Information Processing in Medical Imaging: 22nd International Conference, IPMI 2011, Kloster Irsee, Germany, July 3–8, 2011*.

- Proceedings, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 398–410. doi:<http://dx.doi.org/10.1007/978-3-642-22092-033>.
- [171] T.L. Weissgerber, N.M. Milic, S.J. Winham, V.D. Garovic, Beyond bar and line graphs: time for a new data presentation paradigm, *PLoS Biol.* 13 (4) (2015) e1002128, <http://dx.doi.org/10.1371/journal.pbio.1002128>. URL <http://www.ncbi.nlm.nih.gov/pubmed/25901488>.
- [172] B. Antal, A. Chessel, R.E. Carazo Salas, Mineotaur: a tool for high-content microscopy screen sharing and visual analytics, *Genome Biol.* 16 (1) (2015) 283, <http://dx.doi.org/10.1186/s13059-015-0836-5>.
- [173] H. Ding, C. Wang, K. Huang, R. Machiraju, iGPSe: a visual analytic system for integrative genomic based cancer patient stratification, *BMC Bioinf.* 15 (1) (2014) 203, <http://dx.doi.org/10.1186/1471-2105-15-203>. URL <http://www.ncbi.nlm.nih.gov/pubmed/25000928>.
- [174] A. Yates, A. Webb, M. Sharpnack, H. Chamberlin, K. Huang, R. Machiraju, Visualizing multidimensional data with glyph SPLOMs, *Comput. Graphics Forum* 33 (3) (2014) 301–310, <http://dx.doi.org/10.1111/cgf.12386>. URL <http://doi.wiley.com/10.1111/cgf.12386>.
- [175] S.M. Gustafsdottir, V. Ljosa, K.L. Sokolnicki, J.A. Wilson, D. Walpita, M.M. Kemp, K.P. Seiler, H.A. Carrel, T.R. Golu, S.L. Schreiber, P.A. Clemons, A.E. Carpenter, A.F. Shamji, Multiplex cytological profiling assay to measure diverse, *PLoS ONE* 8 (12) (2013) e80999, <http://dx.doi.org/10.1371/journal.pone.0080999>. URL <http://dx.plos.org/10.1371/journal.pone.0080999>.
- [176] S. Taylor, R. Noble, HTML5 pivotviewer: high-throughput visualization and querying of image data on the web, *Bioinformatics* 30 (18) (2014) 2691–2692.
- [177] T. Lecomte, R. Thibeaux, N. Guillen, A. Dufour, J.C. Olivo-Marin, Fluid optical flow for forces and pressure field estimation in cellular biology, in: *Proceedings – International Conference on Image Processing, ICIP, IEEE, 2012*, pp. 69–72. URL <http://ieeexplore.ieee.org/document/6466797/>.
- [178] A.B. Pujadas, M. Manich, N. Guillen, J.-C. Olivo-Marin, A.C. Dufour, Biophysical measurements in 2d and 3d live cell imaging using fluid dynamics and optical flow, in: *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on, IEEE, 2016*, pp. 980–983.



# An Overview of data science uses in bioimage informatics



Anatole Chessel

LOB, Ecole Polytechnique, CNRS, INSERM, Université Paris-Saclay, 91128 Palaiseau cedex, France

## ARTICLE INFO

### Article history:

Received 31 August 2016

Received in revised form 9 December 2016

Accepted 30 December 2016

Available online 3 January 2017

### Keywords:

Data science

Bioimage informatics

High content screening

Image analysis

## ABSTRACT

This review aims at providing a practical overview of the use of statistical features and associated data science methods in bioimage informatics. To achieve a quantitative link between images and biological concepts, one typically replaces an object coming from an image (a segmented cell or intracellular object, a pattern of expression or localisation, even a whole image) by a vector of numbers. They range from carefully crafted biologically relevant measurements to features learnt through deep neural networks. This replacement allows for the use of practical algorithms for visualisation, comparison and inference, such as the ones from machine learning or multivariate statistics. While originating mainly, for biology, in high content screening, those methods are integral to the use of data science for the quantitative analysis of microscopy images to gain biological insight, and they are sure to gather more interest as the need to make sense of the increasing amount of acquired imaging data grows more pressing.

© 2017 Elsevier Inc. All rights reserved.

## Contents

1. Introduction	110
2. Numerical features computation	111
2.1. Features computation strategy	112
2.2. Feature quality control and post-processing	112
3. Storage, management and sharing of image-derived data	112
4. Features comparison	113
5. Features analysis and interpretation	114
6. Implementations	115
7. Conclusion	115
References	115

## 1. Introduction

For the last couple of decades the development of increasingly efficient fluorescence probes along with technological advances in microscopy has led to terabytes of increasingly resolved images being acquired across models and conditions. The discipline of bioimage informatics (BII) is rising to develop the means to provide a quantitative analysis of those data, and integrate them into larger biological questions and studies, inline with a more general trend of the life sciences toward more quantitative and integrative approaches.

Numerous previous recent reviews, from the more specific to the more generic, have been documenting that rise and the methods developed and used. [1] is a fairly wide review of applications of classical computer vision techniques to biology, while [2] provide a more recent and in depth overview of bioimage informatics. [3,4] are older reviews focusing on High Content Screening (HCS); more recently, [5,6] are comprehensive reviews on HCS while [7] focuses on some example of phenotype analysis. [8] is an earlier review on the uses of machine learning on biological imaging which focuses more specifically on supervised and unsupervised learning. [9] provide a more methodological review and comparison of segmentation methods.

Here we will more specifically focus on the next steps after image segmentation, i.e. the computation and use of statistical

E-mail address: [anatole.chessel@polytechnique.edu](mailto:anatole.chessel@polytechnique.edu)

features, at the intersection with wider data science techniques. Answering biological questions involve performing comparison and inference on complex objects, be it cell shape for morphology, vesicles distributions, actin or microtubule structures, etc... whose handling is not trivial: how to compute the average of a set of cell shape? Or compare the intracellular organisation of segmented microtubules within cells? A classical path to an answer to those questions is to associate  $n$  real numbers to each of them, the so-called features (i.e. equivalently send those objects to  $\mathbb{R}^n$ ), and use the wealth of algorithms and methods available to handle such numerical data.

Analysing images and image derived data in biology has a long history [10], and has grown in scope and visibility in the last thirty years, concomitant with the increased importance of digital microscopy in the life sciences and the rise of quantitative biology. High throughput and high content microscopy, given the scale at which they do experiments, spanning 100s to 1000s of conditions or more, has always been key in moving those developments forward. They include full genome screens [11–13], systematic investigation of genetic interactions [14,15], or more focused investigation of the influence of cell context [16] or of cell motility [17], the variability of cell shape [18], of iPSC cell lines [19] or the investigation of small molecules effects for drug developments [20–22].

But similar techniques have been used outside of HCS, to analyse time lapse data [23,24], In-Situ Hybridisation experiments [25,26], cell lineages [27], perform content based image retrieval [28] or build models of protein localisation [29]. Recent advance in microscopy, leading to the acquisition of very large images [30,27,31], where a single image can weight 100 Gb to several Tb, will also benefit from such large scale data analysis.

All those apparently very diverse applications have at their core the computation of numerical features on their objects of interests, to be used as representation. The main purpose of feature representations is to quantitatively describe complex objects and concepts, essential for their further quantitative analysis. Conversely, one will only be able to access the aspects of an objects that are described by the representation used, hence the choice of the right features, and the right way of visualising and comparing them corresponds to hypothesis on the data at hand and their variabilities.

The precise nature of the study will of course determine the objects for which features are computed, and how. Computing features for whole images or systematic tiles of images has been done. More often, a segmentation of the objects of interest is performed: cells, nuclei or other intracellular objects, etc... That subject is talked about elsewhere and from here on we will assume that those objects of interests have been obtained, and focus on what can be done with them. For more information, one could look at the aforementioned reviews. More information on software implementations will be given Section 6.

This review aims at being accessible to non-expert in computational methods and provide an entry point to data science for bioimage informatics to interested biologists. To separate descriptions of classical data science methods for the readers less familiar with them from their application to biology, a few ‘Method boxes’ have been separated from the main text, describing succinctly classical data science frameworks (box 1), supervised and unsupervised learning (box 2) or dimensionality reduction (box 3). Box 4 provide some generic advice, in particular regarding evaluation, that may be useful to keep in mind when starting out in data science while Box 5 give some more details and advices on available tools. Out of necessity, some descriptions will be succinct

and technical but interested readers will be able to find more in depth tutorial and courses online; see also Section 6 for pointers.

The plan will follow the chronology of projects using those methods, beginning by how the features themselves are computed (Section 2), and handled for storage and visualisation (Section 3). We will then focus specifically on comparison in Section 4 as generic multivariate two-sample tests are an important but tricky topic and look into wider statistical and machine learning techniques for inference and interpretation in Section 5. We will finish by a few word about the practical implementations of all those algorithms in Section 6. A lot of the examples will come from high content screening (HCS) studies for historical reasons, but are now increasingly of use across bioimage informatics. From Section 3 onward, we leave the realm of images per se and will be looking at rather generic biological data science questions of wider relevance.

## 2. Numerical features computation

As said, the aim of the computation of numerical features is to send the objects of interests to  $\mathbb{R}^n$ , i.e. associate  $n$  real number to each of them. Let us note first it’s only one solution (the most commonly used and usually the most practical), among others. For example one could try and work directly in object space, by computing an innate distance between objects, and avoid numerical features entirely. Examples include [33], where a representation of drosophila embryo is built with which computation are done, or [34] which compute kernels on graphs. Those approaches were particularly pursued for shape analysis, using for examples shape diffeomorphisms to compute geodesics in shape space [35–37] but will not be presented further here as they tend to be more complex and more closely tied to particular type of data or analysis frameworks.

Some data science framework ‘Data science’ is a recent label that aggregates various fields and practices that aim at analysing actual data. Classical framework includes:

*Statistics.* From a statistical point of view, features are considered random variables, i.e. variable whose values follow a specific distribution, with the feature set being a *multivariate* random variable, and each object under study to which a feature set is associated (a individual cell, a track, a gene) can be seen as a particular realisation of that random variable. Statistical methods can be parametric, if they assume random variables follow a specific distribution or non-parametric otherwise. Descriptive statistics tries to accurately and usefully describe and characterise samples while statistical inference tries to infer properties of the distribution it comes from.

*Machine learning.* Outlined in box 2, machine learning refers to techniques originally from artificial intelligence concerned with learning rules from data, typically allowing predictions from new data points. Statistics is a major ingredient of most machine learning algorithms. While there is a clear overlap with statistical inference in that both aim at guessing properties of a dataset, machine learning (and the more recent neural nets in particular) tend to focus on results and efficiency, sometime at the expense of a rigorous understanding of the methods.

*Geometry.* A last way to view a feature set is geometric, as a  $n$  dimensional point cloud. Techniques using that viewpoint include nearest neighbour or other distance and projection based techniques. More recently, computational geometry/computational topology is starting to be a useful data science paradigm [32].

## 2.1. Features computation strategy

One key question about statistical image features is their interpretability, i.e. whether they can be directly related to biologically relevant quantities.

A first possibility usually leading to interpretable features is to build hand-crafted features specially adapted to the problem at hand, drawn from the knowledge one may have on the biology at play. One can compute the number, size, intensity, positions of vesicles [12], the length, number and curvature of individual microtubule [13] or simple geometrical properties of shapes ([19] for example). They will have the advantage of being meaningful and to directly relate to biological questions and conclusions, but may be limited in their expressivity and prevent the discovery of truly novel and unexpected results. Other examples include [38] which uses the repartition of intensity around detected vesicles and Ripley's K-function from spatial statistics, [16] where features specific to cell local context are defined, [39] where careful analysis of intensity is performed to allow for semi-quantitative interpretation or [17] which define features on tracks of moving cells.

Another strategy (often used in conjunction of the first) is to compute more complex mathematical quantities which, while harder to relate to will express a wider range of behaviours of the objects under study. Texture features, Fourier moments, spherical harmonics, quantities derived from mathematical morphology or fractal analysis and many more has been used. The extreme version of this strategy are the CHARM features [40,41], which compute up to 3000 quantities from the image and transforms of the image. See also Section 6 for software implementations used to compute those features.

Lastly among classical techniques, a 'multi-tiered' approach has often been used with great success, where features computed from the raw objects of interests are used to feed a supervised learning (see box 2) step where carefully chosen, biologically relevant, classes are used to represent the data. [42,12] classify nuclei into a few classes relevant for mitosis analysis, [13] used a few cell shape, well defined from literature to analyse yeast shape and [18] does the same for mammalian cells. In those cases, features derived from those classes (like the ratio of single cell in each class) are used for further analysis. Other uses of learning are presented Section 5.

Finally a relatively new set of techniques, evolving from older Artificial Neural Networks, and often summarised by the term 'deep learning' have been rising in the last few years, with the first applications to bioimage informatics starting to appear. A good way to think of them and their applications so far in this context is as supervised 'feature learning', although potential uses are much bigger than that. Typically, features are learnt as the ones which best differentiate given training labels for example. Interestingly they are inherently 'multi-tiered' (multi layered in that case, hence why they are called 'deep' network), and thus are not entirely dissimilar to the last set of techniques. A good high level description and review can be found in [43]. Uses in BII include [44], which uses a Convolutional Neural Network (CNN) as a features learning step which compare favourably to classical features and supervised learning pipeline, [45] which uses stacked autoencoders and transfer learning to predict mechanism of action of small molecules or [46] which use weakly supervised CNN to perform joint segmentation and classification. Related works include [47] which used CNN for cell tracking or [48] for histopathological slide classification.

**Supervised and unsupervised learning** Classical machine learning algorithms are usually classified into supervised and unsupervised, depending on the data at hand.

*Supervised.* Supervised learning starts with a set of labelled objects and their associated features and try, from them, to infer a rule (that's the training phase). This rule will allow the prediction of the (unknown) label of a new set of features (testing phase). This is done trying to minimise the generalisation error, i.e. limit over-fitting, which is when one stops learning about the properties of the generic classes the data is but an example of, and starts learning about the peculiarities of that particular, finite, dataset.

*Unsupervised.* Unsupervised learning, or clustering, aims at guessing the structure of the data knowing only the features, by grouping together similar objects. The most classically used algorithm is hierarchical clustering, which uses distances between data points to group close point together starting with the closest ones and build tree-like dendograms. The efficiency of those methods depends a lot on the data themselves and how much of a 'clustered' structure they have and can be hard to evaluate, but can be invaluable in data exploration, especially if little is known a priori.

Other classical learning tasks include regression (learning a continuous value instead of a label), semi-supervised methods, generative methods etc. . .

## 2.2. Feature quality control and post-processing

Importantly before any analysis and interpretation, features quality control and preprocessing needs to be performed. This is in part very much study and experimental design dependant and we will limit ourselves to an overview of problems encountered and solutions proposed.

Depending on the experimental protocol and the study, quality controls done at the level of features include checking the heterogeneity and specificity of RNAi silencing, or controlling for various batch effect (per well or plate for example) [12,11,49], detecting outliers etc. . .Increasingly, an efficient and versatile way to handle those tasks is to use supervised learning with positive and negative controls and bootstrap resampling (see [38] for example). Inference with hierarchical Bayesian models is also possible [50].

Feature selection or reduction (see method box 3) is an important preprocessing step to ensure the features used for analysis are informative and adapted to the task. Most classical methods and some less classical have been used. Classical LDA [51,52,41,25] and PCA [53–55,18] are still often used. Other approaches include [38] which used a SVM recursive feature elimination procedure, while [56] used genetic algorithms. [49] proposes a method based on lasso logistic regression.

## 3. Storage, management and sharing of image-derived data

The computation of features is just one step in a larger pipeline, and soon enough, for a given project/set of experiments, a lot of different related things have to co-exist: the images them-

selves, but also the associated experimental metadata (reagents, microscope settings), segmented regions of interest and corresponding features, possibly training or learnt labels or other analysis results. When adding links with others data sources, or ontologies [57], the need for specific tools and framework for storage and organisation, or for structured data sharing, becomes pressing.

Several such projects have been developed in the last few years. OMERO [58] is an open source client-server software using the bioformat library for microscopy image reading and it's associated data model [59]. It is centered around a database for the storage of images and associated metadata, and has recently been extended for HCS more specific metadata uses [60]. Similar in mindset but coming from cytological slide management, Cytomine [61] focuses more on very large images. Other similar tools include OpenBIS [62], which also include management of proteomics or deep sequencing data and Bisque [63]. Outside of microscopy per se, InterMine [64] is a similar framework aiming at bringing together data from several sources and databases.

The sharing of structured data that include images along with the derived metadata involve the definition of a suitable file format standard. OME-TIFF [59] is an open format which stores images along with the experimental metadata, but excluding analysis metadata. Attempts focusing on those include [65] which proposes to manage experimental data using semantically typed data hypercubes (SDCubes), combining hierarchical data format 5 (HDF5) and extensible markup language (XML) file types, and [66] which propose the CellH5 format, also based on HDF5, to specifically allow the storage of graph-like objects relations (like the lineages segmented cells are part of).

As an alternative to a specific file format for data sharing, several projects built websites to publish results in an accessible and searchable way and visualise imaging data. [www.mitochek.org](http://www.mitochek.org) show the results of [11], [sysgro.org](http://sysgro.org) of [13] and [67] of [39], to name but a few. To interactively visualise and/or publish a large number of features, [68] provides a visual analytics framework using the NoSQL database Neo4j. All those efforts are parts of the 'open data' idea, which promote easy and unrestricted data sharing [69]. Particularly for bioimage informatics, it allow for a wider community of non computational experts to be able to browse the data and for data reuse, which is an established practice in data science in general and increasingly biology in particular (used in [70,71,44–46,72] for example).

Despite all those works, the actual sharing of large image data sets poses logistic issues due to their size, complexity and diversity. Many data sharing solutions are now available to scientists, but few are adapted to the scale and complexity of imaging data. One such example is the Image Data Repository (IDR [107], <http://idr-demo.openmicroscopy.org>), an online, public resource that seeks to store, integrate and serve image datasets from published scientific studies.

**Feature selection and dimensionality reduction** One way to view data in  $R^n$  is as an  $n$  dimensional point cloud. But that point cloud is not evenly distributed in those dimensions and often occupy a much smaller subspace. Often, some features are uninformative, or too noisy, or redundant. It is thus very often possible to find  $p$  dimension, with  $p \ll n$ , such that the same dataset can be expressed in  $R^p$  more usefully and with little (or controlled) loss of information.

*Feature selection.* Feature selection aims at selecting a subset of the computed features according to some criteria, either non-supervised (uncorrelated and with a minimal signal/noise ratio for example), or supervised (providing the best separation of known and relevant training labels). Example are provided in the main text; [73] provides a generic review and [74] one more recent but focused on bioinformatics.

*Manifold learning.* Another set of technique tries to build new features out of the old ones, potentially losing interpretability along the way. Here again supervised and non-supervised approaches exist, with respectively Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) being the linear algorithms most often use in practice (see main text). Finding generic, non-linear and non-parametric algorithms have been an active area of research with [75] providing a comparative review. [76] provides some perspective in comparing the various differing algorithms.

#### 4. Features comparison

One of the key questions asked by biology to data is to test the difference between two conditions: is the phenotype of interest different in condition A compared to condition B? While the statistical tools to answer that question are well defined for a single measurement, with t-test under a normality assumption and the Kolmogorov–Smirnov (KS) test as a non-parametric test for example, the situation is much more complex for multivariate measurements. There can't be said to exist one single accepted solution, and many different approaches have been developed, both generic and specifically for bioimage informatics.

Generic test proposed in statistics include extension of the KS test [77], of rank tests [78] or based on distances and bootstrap permutation [79]. More recently, kernel tests have been developed [80] which have been shown to be a more general class of earlier tests [81]. Those works are usually hard to use in practice however as implementation can be hard to find, slow for realistic data, or limited in the number of dimension (i.e. features) they can handle. A rather large survey of method used in a variety of application can be found in [82].

More specifically in bioimage informatics, a lot of different methods have been used and proposed. For example [12] proposes to break the features vector down into it's norm, which

can be interpreted as the amplitude of the phenotype, and the angle, which represent the phenotype itself. [13] used two methods, one based on KS test of individual features and another based on euclidian distance of the full vector, to differentiate conditions which differs greatly in only one features to those with a widespread but milder difference. [53] proposes a comparison of some classical methods on small molecule perturbations showing that the simplest method, based on difference of the means, perform at least as well as more complex ones. [55] (also used in [38,72] for example) propose a supervised learning method to compute a dissimilarity score between two populations, by checking how well is a SVM classifier able to distinguish between them. [54] focus on the non parametric redistribution of phenotypes using Earth Movers Distance. [50] propose an integrated model using hierarchical bayesian modeling and copulas to account for the within population variability.

In HCS more specifically, the aim is often to provide a clear decision, where each given condition is either different or not different from the negative controls. A mathematically well defined path is to use a statistical test, which provide a p-value, and use multiple test correction [83] to set the significance threshold, to account for the fact that for a given significance threshold, the more tests are performed, the more false positive one will get. Such a threshold, while useful or even necessary is usually recognised as being somewhat arbitrary as phenotypes happen in a continuum rather than as clear presence/absence. Methods providing a finer look into the extracted data will be presented in the next section.

**Building and evaluating a data analysis pipeline** *Methods evaluation.* Given the breadth of techniques presented in this review, even by limiting its scope to more traditional features based methods, the question of methods' assessment and evaluation of performance, and choice of the right tool for the right purpose is an important one. Obviously, given the focus on quantitative analysis, each part of a data analysis pipeline could and should be evaluated. Two main ways of doing so is against a supervised 'ground truth', if one is available or time can be spent annotating data, or comparing with simulations, if enough simplifying assumptions exists that some can be made. Both methods are used within the papers cited here and generally accepted, and have as a caveat that they only evaluate known behaviour of the system.

*Visualisations and controls.* Indeed stacking algorithmic methods one after the other until the link with the raw data and the initial aims becomes unclear is fairly easy. Having regular controls and visualisations enables one to check for relevance and correct mistakes before they propagate down an analysis pipeline.

*Exploratory analysis and interpretation.* In particular, during exploratory analysis, it is customary to try a lot of methods and parameters, potentially leading to over-fitting of data or cherry-picking results. On top of checks and controls, biology is an experimental science and the final answers will always be given by an experimental validation or corroboration with other results. . .

*Reproducible research.* In data science, one can consider the code like an electronic lab book, recording everything tried and the results obtained. Modern version control (using github -github.com-, bitbucket -bitbucket.org or gitlab -gitlab.com- for example) allows one to store the whole history of changes of a project in a way that can be easily shared privately or publicly.

## 5. Features analysis and interpretation

The amount of data at this point can be staggering, with up to millions of objects having tens to hundreds or more of features, and the methods to do data mining, i.e. used to dig up informations from the data, are very varied indeed. We will only attempt here to provide pointers to some classical and less classical methods and the way they have been used in bioimage informatics.

Data visualisation in a useful and informative way is of crucial importance in such studies. Classical plots [84] include bar plot and box plot [85], which tend to hide the true distribution of the data, histograms and heatmaps that show it but are less synthetic, scatterplots to look at the dependence between two variables, and scatter plot matrices if there is more than two. Examples of more exotic visualisation can be found in [86]. Increasingly the ability to generate interactive plots is becoming crucial (see Section 6 for possible implementation). Particularly in BII the ability to go back to the image, i.e. to show the actual raw acquired image (or the part of it) which is at the basis of the performed analysis, is important for both quality control and illustration purposes. The data handling frameworks presented Section 3 can be very handy for that task.

An issue most of the classical visualisation techniques have is their inability to usefully plot data of more than a few dimensions. Dimensionality reduction or manifold learning are often used to reduce data dimension to a plotable size while retaining useful information (see Method Box 3 for more information on the techniques). While PCA and LDA remain the most classical methods for pre-processing, prior to machine learning for example (see Section 2 and the remainder of this section), the t-SNE method [87] and it's variant (like [88] in biology) is specifically optimized for boiling down a high dimensional dataset to two or three dimensions, or even one [89], for plotting. More recently, methods based on local clustering to provide a topological map of the data are being developed, like SPADE [90] or PhenoGraph [91].

Machine learning methods are a staple of data science pipelines and a very powerful way to try and interpret data. So called 'phenotypic clustering' for example refer to the use of unsupervised learning to try and assign conditions or objects to groups with similar phenotypes [13,92,14,17]. Hierarchical clustering is the most often used method for it's versatility and ease of interpretation but by no mean the only one. A more indepth look is provided by phenotypic profiling techniques [93], in particular when additional information is used to compare to or annotate, like DAVID for gene ontology enrichment [94] or STRING for protein interaction [95,26]. While used more often earlier in a pipeline, to provide a more interpretable view of the features when an annotated training set is available (see Section 2), supervised clustering is also used as an analysis tool [13,16]. [52] provide a comparison of learning algorithm in the context of BII.

Networks, i.e. nodes connected by edges, is a very useful representation in biology, used for interaction networks, transcription networks, metabolic or signalling networks, etc. . It's also a very common data analysis and visualisation tool. A common assumption used to build networks out of multivariate features is to build edges between similar objects. For example in [26] the gene pairs with the 5% smallest distance in feature space are linked together, with [92] using a similar techniques. [71] proposes a more complex statistical method accounting for the sign of the association between genes and [96] provides an overview of those sort of techniques. Another network inference technique uses conditional independence between features; the presence of an edge imply a causal relationship between the two features. Starting from [97], it has been used in [13,12,16] to provide a systemic view of feature

dependence. [98] develops a Gaussian Graphical Model method that includes multiple instance learning to perform a similar task.

**Tools for feature extraction and analysis** Here are some more details and comments on a few selected software solutions for feature computation and analysis. While very far from exhaustive, it shows the range of solution used, from specific to bioimage informatics like ImageJ to generic programming languages.

*ImageJ.* ImageJ is the most well known and used tool in bioimage informatics. It can be used to compute features from regions of interest (ROIs) using the built-in measurement tool. Various plug-ins, too numerous to list here, exist to compute more complex features. An interesting effort toward an integrated and generic tool to analyse large datasets is TANGO [99], with an integrated link to R.

*Cell Profiler.* The most user friendly and versatile tool for high content analysis is arguably Cell Profiler, along with Cell Profiler Analyst. It includes a user-friendly GUI, access to a number of modules for processing, segmentation, feature computation and analysis (including CHARM [41]), and tools for annotation and evaluation. A downside, as for most integrated packages, is that it could be hard to expand if one's problem is not handled out of the box.

*Knime.* Knime is an open-source data analytics platform. Using a graphical programming framework, it can interface with various data sources, preprocessing and analysis algorithms or visualisations, including image analysis. It's an interesting solution, both versatile but without resorting to full fledged programming.

*Lower level languages.* Matlab, R and Python are certainly the most used in practice. While the learning curve for a non-specialist is certainly steeper, they give access to a very wide range of state-of-the-art methods. R is historically strong in statistics, Matlab is versatile but has been much used in the image analysis community while python is (relatively) newer but with a dedicated and fast growing community.

Another kind of analysis, not necessarily based on statistical features but very fruitful are generative models (see [35,100,72,101] for a recent review). So far we were focused on descriptive models, aiming at quantitatively describe complex datasets for further interpretation and comparison. Generative models aim instead (or in addition) at being able to generate (simulate) new typical examples of the data under study. Aims include the study of statistical variability, a more accurate and intuitive way to describe a particular localisation or further uses in modelling. In particular in [72] conditional generative models of punctuate pattern knowing microtubule localisation are learnt, showing the potential of those methods in studying the relative sub-cellular positions of several organelles.

One of the uses of BII is to provide quantitative measurements to nourish and calibrate quantitative models and simulations from wider computational work. Building those using the very high number of feature of potentially individual object or events could be key in understanding complex biological processes at the system level. Building and comparing networks as described earlier goes in that direction. A more direct attempt can be found in [102] where a pde model is build to study cell population dynamics from siRNA treatments, from time resolved SVN classes. Similarly, [103] aims to infer endosomal network dynamics from multiparametric quantitative imaging.

## 6. Implementations

All the methods and algorithms presented cannot, in practice, be separated from their implementations. Thanks to open source code and the trend toward open science, the availability one way or another of implementations of published computational works is getting ever closer to being natural and expected. A short description is given in the following section, with method box 5 providing more hands on advice.

The software classically used in BII like ImageJ [104] can be used to compute a segmentation or pre-process images but are not necessarily the most adapted ones for feature computation and analysis, unless specific plugins are used, like [99]. More specifically, CellProfiler [105] and CellProfilerAnalyst [106], modular software for high content image analysis with a user friendly interface, or Knime ([www.knime.org](http://www.knime.org)), a versatile data science platform with a graphical programming interface and rich image analysis modules, could be used instead.

But to perform those analysis, an interactive interpreted language, such as R ([r-project.org](http://r-project.org)), Python ([python.org](http://python.org)) or Matlab ([mathworks.com](http://mathworks.com), closed source commercial software) is typically used, allowing for rapid prototyping and visualisation. They all have a rich library of algorithms for image analysis and feature computation, data mining both classical and state of the art and interactive visualisation, each with their specificity. Python in particular, with the scipy stack ([scipy.org](http://scipy.org)), is rapidly growing in both the machine learning and more applied data analysis community. Interactive notebook environment like Jupyter ([jupyter.org](http://jupyter.org)) is also proving a popular data science environment allowing for code, results and explanations to coexist in one convenient place. Thanks to the availability of those software and the ease to share code online, the interested reader will easily be able to find online tutorials covering the basics (and more!) of the use of those environments for data analysis.

## 7. Conclusion

BII is about using quantitative imaging to further biological understanding, and the current explosion of data science is key to making that link between quantitative measurements and biological questions. Here we proposed an overview of methods used to tackle those issues, focusing on statistical feature computation and analysis.

We saw that most data science techniques, both standard (features selection, multivariate statistics or supervised and non-supervised learning) as well as state of the art (feature learning via deep learning in particular) have been applied to image based biological data analysis. Increasingly, methods are being developed specifically for use in biological settings. But most importantly, taking into account the specifics of the experimental settings, both technical and biological, as well as of the questions under study is of course paramount.

While a lot of the works cited in the present article come from the high content screening world as this sub-field of BII is the one where the biggest amount of data were acquired early on, quantitative data analysis techniques is sure to spread to the whole of bioimaging, as microscopy throughput and resolution gets higher and biological questions get more refined.

## References

- [1] G. Danuser, Computer vision in cell biology, *Cell* 147 (5) (2011) 973–978, <http://dx.doi.org/10.1016/j.cell.2011.11.001>. <http://www.ncbi.nlm.nih.gov/pubmed/22118455>.
- [2] I.F. Sbalzarini, Seeing is believing: quantifying is convincing: computational image analysis in biology, *Adv. Anat. Embryol. Cell Biol.* 219 (2016). <http://>

- [www.ncbi.nlm.nih.gov/pubmed/27207361](http://www.ncbi.nlm.nih.gov/pubmed/27207361), <http://link.springer.com/10.1007/978-3-319-28549-8>.
- [3] C. Conrad, D.W. Gerlich, Automated microscopy for high-content RNAi screening, *J. Cell Biol.* 188 (4) (2010) 453–461, <http://dx.doi.org/10.1083/jcb.200910105>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2828931&tool=pmcentrez&rendertype=abstract>.
  - [4] S. Wong, X. Zhou, Informatics challenges of high-throughput microscopy, *IEEE Signal Process. Mag.* 23 (3) (2006) 63–72, <http://dx.doi.org/10.1109/MSP.2006.1628879>. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1628879>.
  - [5] M. Boutros, F. Heigwer, C. Laufer, Microscopy-based high-content screening, *Cell* 163 (6) (2015) 1314–1325, <http://dx.doi.org/10.1016/j.cell.2015.11.007>. <http://linkinghub.elsevier.com/retrieve/pii/S0092867415014877>.
  - [6] M. Mattiazzi Usaj, E.B. Styles, A.J. Verster, H. Friesen, C. Boone, B.J. Andrews, High-content screening for quantitative cell biology, *Trends Cell Biol.* (2016). <http://www.ncbi.nlm.nih.gov/pubmed/27118708>.
  - [7] W. Chen, X. Xia, Y. Huang, X. Chen, J.-D. J. Han, Bioimaging for quantitative phenotype analysis, *Methods* 102 (2016) 20–25, <http://dx.doi.org/10.1016/j.ymeth.2016.01.017>.
  - [8] C. Sommer, D.W. Gerlich, Machine learning in cell biology – teaching computers to recognize phenotypes, *J. Cell Sci.* 126 (Pt 24) (2013) 5529–5539, <http://dx.doi.org/10.1242/jcs.123604>. <http://www.ncbi.nlm.nih.gov/pubmed/24259662>.
  - [9] P. Bajcsy, A. Cardone, J. Chalfoun, M. Halter, D. Juba, M. Kociolek, M. Majurski, A. Peskin, C. Simon, M. Simon, A. Vandecreme, M. Brady, Survey statistics of automated segmentations applied to optical imaging of mammalian cells, *BMC Bioinf.* 16 (2015) 330, <http://dx.doi.org/10.1186/s12859-015-0762-2>. <http://europepmc.org/articles/PMC4608288?report=abstract>.
  - [10] E.B. King, L.K. Kromhout, K.L. Chew, B.H. Mayall, N.L. Petrakis, R.H. Jensen, I.T. Young, Analytic studies of foam cells from breast cancer precursors, *Cytometry* 5 (1984) 124–130.
  - [11] B. Neumann, T. Walter, J.-K. Hériché, J. Bulkescher, H. Erfle, C. Conrad, P. Rogers, I. Poser, M. Held, U. Liebel, C. Cetin, F. Sieckmann, G. Pau, R. Kabbe, A. Wünsche, V. Satagopam, M.H.A. Schmitz, C. Chapuis, D.W. Gerlich, R. Schneider, R. Eils, W. Huber, J.-M. Peters, A.a. Hyman, R. Durbin, R. Pepperkok, J. Ellenberg, Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes, *Nature* 464 (7289) (2010) 721–727, <http://dx.doi.org/10.1038/nature08869>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3108885&tool=pmcentrez&rendertype=abstract>.
  - [12] C. Collinet, M. Stöter, C.R. Bradshaw, N. Samusik, J.C. Rink, D. Kenski, B. Habermann, F. Buchholz, R. Henschel, M.S. Mueller, W.E. Nagel, E. Fava, Y. Kalaidzidis, M. Zerial, Systems survey of endocytosis by multiparametric image analysis, *Nature* 464 (7286) (2010) 243–249, <http://dx.doi.org/10.1038/nature08779>. <http://www.ncbi.nlm.nih.gov/pubmed/20190736>.
  - [13] V. Graml, X. Studera, J.L. Lawson, A. Chessel, M. Geymonat, M. Bortfeld-Miller, T. Walter, L. Wagstaff, E. Piddini, R.E. Carazo-Salas, A genomic multiprocess survey of machineries that control and link cell shape, microtubule organization, and cell-cycle progression, *Dev. Cell* 31 (2) (2014) 227–239, <http://dx.doi.org/10.1016/j.devcel.2014.09.005>. <http://linkinghub.elsevier.com/retrieve/pii/S1534580714005917>.
  - [14] C. Laufer, B. Fischer, M. Billmann, W. Huber, M. Boutros, Mapping genetic interactions in human cancer cells with RNAi multiparametric phenotyping, *Nat. Methods* 10 (5) (2013) 427–431, <http://dx.doi.org/10.1038/nmeth.2436>. <http://www.ncbi.nlm.nih.gov/pubmed/23563794>.
  - [15] A. Roguev, D. Talbot, G.L. Negri, M. Shales, G. Cagney, S. Bandyopadhyay, B. Panning, N.J. Krogan, Quantitative genetic-interaction mapping in mammalian cells, *Nat. Methods* 10 (5) (2013) 432–437, <http://dx.doi.org/10.1038/nmeth.2398>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3641890&tool=pmcentrez&rendertype=abstract>.
  - [16] B. Snijder, R. Sacher, P. Rämö, P. Liberali, K. Mench, N. Wolfrum, L. Burleigh, C. C. Scott, M.H. Verheije, J. Mercer, S. Moese, T. Heger, K. Theusner, A. Jurgeit, D. Lamparter, G. Balistreri, M. Schelhaas, C.A.M. De Haan, V. Marjomäki, T. Hyypiä, P.J.M. Rottier, B. Sodeik, M. Marsh, J. Gruenberg, A. Amara, U. Greber, A. Helenius, L. Pelkmans, Single-cell analysis of population context advances RNAi screening at multiple levels, *Mol. Syst. Biol.* 8 (579) (2012) 579, <http://dx.doi.org/10.1038/msb.2012.9>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3361004&tool=pmcentrez&rendertype=abstract>.
  - [17] A. Schoenauer Sebag, S. PlanCADE, C. Raulet-Tomkiewicz, R. Barouki, J.-P. Vert, T. Walter, A generic methodological framework for studying single cell motility in high-throughput time-lapse data, *Bioinformatics* 31 (12) (2015) i320–i328, <http://dx.doi.org/10.1093/bioinformatics/btv225>. <http://bioinformatics.oxfordjournals.org/lookup/doi/10.1093/bioinformatics/btv225>.
  - [18] Z. Yin, A. Sadok, H. Sailem, A. McCarthy, X. Xia, F. Li, M.A. Garcia, L. Evans, A.R. Barr, N. Perrimon, C.J. Marshall, S.T.C. Wong, C. Bakal, A screen for morphological complexity identifies regulators of switch-like transitions between discrete cell shapes, *Nat. Cell Biol.* 15 (7) (2013) 860–871, <http://dx.doi.org/10.1038/ncb2764>. <http://www.ncbi.nlm.nih.gov/pubmed/23748611>.
  - [19] A. Leha, N. Moens, R. Melecky, O.J. Culley, M.K. Gervasio, M. Kerz, A. Reimer, S.A. Cain, I. Streeter, A. Folarin, O. Stegle, C.M. Kielty, R. Durbin, F.M. Watt, D. Danovi, A high-content platform to characterise human induced pluripotent stem cell lines, *Methods* 96 (2016) 85–96, <http://dx.doi.org/10.1016/j.ymeth.2015.11.012>.
  - [20] M. Breinig, F.A. Klein, W. Huber, M. Boutros, A chemical-genetic interaction map of small molecules using high-throughput imaging in cancer cells, *Mol. Syst. Biol.* 11 (12) (2015), <http://dx.doi.org/10.15252/msb.20156400>. 846–846, <http://msb.embopress.org/cgi/doi/10.15252/msb.20156400>.
  - [21] M.J. Wawer, K. Li, S.M. Gustafsdottir, V. Ljosa, N.E. Bodycombe, M.A. Marton, K.L. Sokolnicki, M.-A. Bray, M.M. Kemp, E. Winchester, B. Taylor, G.B. Grant, C. S.-Y. Hon, J.R. Duvall, J.A. Wilson, J.A. Bittker, V. DanÅ ik, R. Narayan, A. Subramanian, W. Winckler, T.R. Golub, A.E. Carpenter, A.F. Shamji, S.L. Schreiber, P.A. Clemons, Toward performance-diverse small-molecule libraries for cell-based phenotypic screening using multiplexed high-dimensional profiling, *Proc. Natl. Acad. Sci.* 111 (30) (2014) 10911–10916, <http://dx.doi.org/10.1073/pnas.1410933111>. <http://www.pnas.org/cgi/doi/10.1073/pnas.1410933111>.
  - [22] S.M. Gustafsdottir, V. Ljosa, K.L.K. Sokolnicki, J. Anthony Wilson, D. Walpita, M.M. Kemp, K. Petri Seiler, H.A. Carrel, T.R. Golub, S.L. Schreiber, P.A. Clemons, A.A.E. Carpenter, A.F. Shamji, Multiplex cytological profiling assay to measure diverse cellular states, *PLoS one* 8 (12) (2013) e80999, <http://dx.doi.org/10.1371/journal.pone.0080999>. <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0080999#>, <http://dx.plos.org/10.1371/journal.pone.0080999>.
  - [23] Q. Zhong, A.G. Busetto, J.P. Fededa, J.M. Buhmann, D.W. Gerlich, Unsupervised modeling of cell morphology dynamics for time-lapse microscopy, *Nat. Methods* 9 (7) (2012) 711–713, <http://dx.doi.org/10.1038/nmeth.2046>. <http://www.nature.com/doi/10.1038/nmeth.2046>.
  - [24] K.T. Roybal, T.E. Buck, X. Ruan, B.H. Cho, D.J. Clark, R. Ambler, H.M. Tunbridge, J. Zhang, P. Verkade, C. Wülfing, R.F. Murphy, Computational spatiotemporal analysis identifies WAVE2 and cofilin as joint regulators of costimulation-mediated T cell actin dynamics, *Sci. Signaling* 9 (424) (2016) 1–13, <http://dx.doi.org/10.1126/scisignal.aad4149>.
  - [25] K. Puniyani, C. Faloutsos, E.P. Xing, SPEX2: automated concise extraction of spatial gene expression patterns from Fly embryo ISH images, *Bioinformatics* 26 (12) (2010) i47–i56, <http://dx.doi.org/10.1093/bioinformatics/btq172>. <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btq172>.
  - [26] N. Battich, T. Stoeger, L. Pelkmans, Image-based transcriptomics in thousands of single human cells at single-molecule resolution, *Nat. Methods* 10 (11) (2013) 1127–1133, <http://dx.doi.org/10.1038/nmeth.2657>. <http://www.ncbi.nlm.nih.gov/pubmed/24097269>.
  - [27] F. Amat, W. Lemon, D.P. Mousing, K. McDole, Y. Wan, K. Branson, E.W. Myers, P.J. Keller, Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data, *Nat. Methods* 11 (9) (2014) 951–958, <http://dx.doi.org/10.1038/nmeth.3036>. <http://www.nature.com/doi/10.1038/nmeth.3036>.
  - [28] B.H. Cho, I. Cao-Berg, J.A. Bakal, R.F. Murphy, Omero searcher: content-based image search for microscope images, *Nat. Methods* 9 (7) (2012) 633–634, <http://dx.doi.org/10.1038/nmeth.2086>. <http://www.nature.com/doi/10.1038/nmeth.2086>.
  - [29] L.P. Coelho, J.D. Kangas, A.W. Naik, E. Osuna-Highley, E. Glory-Afshar, M. Fuhrman, R. Simha, P.B. Berget, J.W. Jarvik, R.F. Murphy, Determining the subcellular location of new proteins from microscope images using local features, *Bioinformatics* 29 (18) (2013) 2343–2349, <http://dx.doi.org/10.1093/bioinformatics/btt392>.
  - [30] E.H.K. Stelzer, Light-sheet fluorescence microscopy for quantitative biology, *Nat. Methods* 12 (1) (2014) 23–26, <http://dx.doi.org/10.1038/nmeth.3219>. <http://www.nature.com/doi/10.1038/nmeth.3219>.
  - [31] P. Mahou, M. Zimmerley, K. Loulier, K.S. Matho, G. Labroille, X. Morin, W. Supatto, J. Livet, D. Débarre, E. Beaupaire, Multicolor two-photon tissue imaging by wavelength mixing, *Nat. Methods* 9 (8) (2012) 815–818, <http://dx.doi.org/10.1038/nmeth.2098>. <http://dx.doi.org/10.1038/nmeth.2098>.
  - [32] G. Carlsson, Topology and data, *Bull. Am. Math. Soc.* 46 (2) (2009) 255–308, <http://dx.doi.org/10.1090/S0273-0979-09-01249-X>. <http://www.ams.org/journal-getitem?pii=S0273-0979-09-01249-X>.
  - [33] E. Frise, A.S. Hammonds, S.E. Celniker, Systematic image-driven analysis of the spatial Drosophila embryonic expression landscape, *Mol. Syst. Biol.* 6 (345) (2010) 345, <http://dx.doi.org/10.1038/msb.2009.102>. <http://msb.embopress.org/cgi/doi/10.1038/msb.2009.102>. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2824522&tool=pmcentrez&rendertype=abstract>.
  - [34] J.-K. Hériché, J.G. Lees, I. Morilla, T. Walter, B. Petrova, M.J. Roberti, M.J. Hossain, P. Adler, J.M. Fernández, M. Krallinger, C.H. Haering, J. Vilo, A. Valencia, J.A. Ranea, C. Orenzo, J. Ellenberg, Integration of biological data by kernels on graph nodes allows prediction of new genes involved in mitotic chromosome condensation, *Mol. Biol. Cell* 25 (16) (2014) 2522–2536, <http://dx.doi.org/10.1091/mbc.E13-04-0221>. <http://europepmc.org/articles/PMC4142622?report=abstract>.
  - [35] T. Peng, W. Wang, G.K. Rohde, R.F. Murphy, Instance-based generative biological shape modeling, *ISBI 2009*, vol. 1, 2009, pp. 690–693.
  - [36] C. Hagwood, J. Bernal, M. Halter, J. Elliott, T. Brennan, Testing equality of cell populations based on shape and geodesic distance, *IEEE Trans. Med. Imaging* 32 (12) (2013) 2230–2237, <http://dx.doi.org/10.1109/TMI.2013.2278467>. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6587104>.
  - [37] M. Zhang, P. Golland, Statistical shape analysis: From landmarks to diffeomorphisms, *Med. Image Anal.* (2016), <http://dx.doi.org/10.1016/j.media.2016.06.025>.
  - [38] P. Liberali, B. Snijder, L. Pelkmans, A hierarchical map of regulatory genetic interactions in membrane trafficking, *Cell* 157 (6) (2014) 1473–1487, <http://dx.doi.org/10.1016/j.cell.2014.04.029>.

- [39] M. Breker, M. Gymrek, M. Schuldiner, A novel single-cell screening platform reveals proteome plasticity during yeast stress responses, *J. Cell Biol.* 200 (6) (2013) 839–850, <http://dx.doi.org/10.1083/jcb.201301120>.
- [40] N. Orlov, L. Shamir, T. Macura, J. Johnston, D.M. Eckley, I.G. Goldberg, WND-CHARM: Multi-purpose image classification using compound image transforms, *Pattern Recognit. Lett.* 29 (11) (2008) 1684–1693, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2573471&tool=pmcentrez&rendertype=abstract>.
- [41] V. Uhlmann, S. Singh, A.E. Carpenter, CP-CHARM: segmentation-free image classification made accessible, *BMC Bioinf.* 17 (1) (2016) 51, <http://dx.doi.org/10.1186/s12859-016-0895-y>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4729047&tool=pmcentrez&rendertype=abstract>.
- [42] T. Walter, M. Held, B. Neumann, J.-K. Hériché, C. Conrad, R. Pepperkok, J. Ellenberg, Automatic identification, clustering of chromosome phenotypes in a genome wide RNAi screen by time-lapse imaging, *J. Struct. Biol.* 170 (1) (2010) 1–9, <http://dx.doi.org/10.1016/j.jsb.2009.10.004>, <http://www.ncbi.nlm.nih.gov/pubmed/19854275>.
- [43] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444, <http://dx.doi.org/10.1038/nature14539>, <http://dx.doi.org/10.1038/nature14539>.
- [44] O. Dürr, B. Sick, Single-cell phenotype classification using deep convolutional neural networks, *J. Biomol. Screening* (2016), <http://dx.doi.org/10.1177/1087057116631284>, <http://jbx.sagepub.com/cgi/doi/10.1177/1087057116631284>.
- [45] C. Kandhaswamy, L.M. Silva, L.A. Alexandre, J.M. Santos, High-content analysis of breast cancer using single-cell deep transfer learning, *J. Biomol. Screening* 21 (3) (2016) 252–259, <http://dx.doi.org/10.1177/1087057115623451>, <http://jbx.sagepub.com/content/21/3/252.abstract>.
- [46] O.Z. Kraus, J.L. Ba, B.J. Frey, Classifying and segmenting microscopy images with deep multiple instance learning, *Bioinformatics* 32 (12) (2016) i52–i59, <http://dx.doi.org/10.1093/bioinformatics/btw252>, <http://bioinformatics.oxfordjournals.org/lookup/doi/10.1093/bioinformatics/btw252>.
- [47] A. Jackson-smith, Cell Tracking using Convolutional Neural Networks.
- [48] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using convolutional neural networks, *International Joint Conference on Neural Networks (IJCNN 2016)*, 2016.
- [49] J. Rameseder, K. Krismer, Y. Dayma, T. Ehrenberger, M.K. Hwang, E.M. Airoldi, S.R. Floyd, M.B. Yaffe, A multivariate computational method to analyze high-content RNAi screening data, *J. Biomol. Screening* (2015), <http://dx.doi.org/10.1177/1087057115583037>, <http://jbx.sagepub.com/content/early/2015/04/24/1087057115583037.full>.
- [50] E. Pauwels, D. Surdez, G. Stoll, A. Lescuré, E. Del Nery, O. Delattre, V. Stoven, A probabilistic model for cell population phenotyping using HCS data, *PLoS one* 7 (8) (2012) e42715, <http://dx.doi.org/10.1371/journal.pone.0042715>, <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0042715>.
- [51] S. Huh, D. Lee, R.F. Murphy, Efficient framework for automated classification of subcellular patterns in budding yeast, *Cytometry Part A* 75 (11) (2009) 934–940, <http://dx.doi.org/10.1002/cyto.a.20793>, <http://doi.wiley.com/10.1002/cyto.a.20793>.
- [52] S.S. Abbas, T.M.H. Dijkstra, T. Heskes, A comparative study of cell classifiers for image-based high-throughput screening, *BMC Bioinf.* 15 (1) (2014) 342, <http://dx.doi.org/10.1186/1471-2105-15-342>, <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-15-342>.
- [53] V. Ljosa, P.D. Caie, R. Ter Horst, K.L. Sokolnicki, E.L. Jenkins, S. Daya, M.E. Roberts, T.R. Jones, S. Singh, A. Genovesio, P.A. Clemons, N.O. Carragher, A.E. Carpenter, Comparison of methods for image-based profiling of cellular morphological responses to small-molecule treatment, *J. Biomol. Screening* (2013), <http://dx.doi.org/10.1177/1087057113503553>.
- [54] S. Singh, F. Janoos, T. Pécot, E. Caserta, K. Huang, J. Rittscher, G. Leone, R. Machiraju, Non-parametric population analysis of cellular phenotypes, *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 14 (Pt 2)* (2011) 343–351, <http://www.ncbi.nlm.nih.gov/pubmed/2199504>.
- [55] X. Zhang, M. Boutros, A novel phenotypic dissimilarity method for image-based high-throughput screens, *BMC Bioinf.* 14 (1) (2013) 336, <http://dx.doi.org/10.1186/1471-2105-14-336>, <http://www.biomedcentral.com/1471-2105/14/336>.
- [56] P. Charoenkwan, E. Hwang, R.W. Cutler, H.-C. Lee, L.-W. Ko, H.-L. Huang, S.-Y. Ho, HCS-Neurons: identifying phenotypic changes in multi-neuron images upon drug treatments of high-content screening, *BMC Bioinf.* 14 Suppl 1 (Suppl 16) (2013) S12, <http://dx.doi.org/10.1186/1471-2105-14-S16-S12>, <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-S16-S12>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3853092&tool=pmcentrez&rendertype=abstract>.
- [57] R. Hoehndorf, M.A. Harris, H. Herre, G. Rustici, G.V. Gkoutos, Semantic integration of physiology phenotypes with an application to the cellular phenotype ontology, *Bioinf. (Oxford England)* 28 (13) (2012) 1783–1789, <http://dx.doi.org/10.1093/bioinformatics/bts250>, <http://www.ncbi.nlm.nih.gov/pubmed/22539675>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3381966>.
- [58] C. Allan, J.-M. Burel, J. Moore, C. Blackburn, M. Linkert, S. Loynton, D. Macdonald, W.J. Moore, C. Neves, A. Patterson, M. Porter, A. Tarkowska, B. Loranger, J. Avondo, I. Lagerstedt, L. Lianas, S. Leo, K. Hands, R.T. Hay, A. Patwardhan, C. Best, G.J. Kleywegt, G. Zanetti, J.R. Swedlow, OMERO: flexiblemodel-driven data management for experimental biology, *Nat. Methods* 9 (3) (2012) 245–253, <http://dx.doi.org/10.1038/nmeth.1896>.
- [59] M. Linkert, C.T. Rueden, C. Allan, J.-M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, C. Neves, D. Macdonald, A. Tarkowska, C. Sticco, E. Hill, M. Rossner, K.W. Eliceiri, J.R. Swedlow, Metadata matters: access to image data in the real world, *J. Cell Biol.* 189 (5) (2010) 777–782, <http://dx.doi.org/10.1083/jcb.201004104>, <http://www.ncbi.nlm.nih.gov/pubmed/20513764>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2878938>.
- [60] S. Li, S. Besson, C. Blackburn, M. Carroll, R.K. Ferguson, H. Flynn, K. Gillen, R. Leigh, D. Lindner, M. Linkert, W.J. Moore, B. Ramalingam, E. Rozbicki, G. Rustici, A. Tarkowska, P. Walczysko, E. Williams, C. Allan, J.-M. Burel, J. Moore, J.R. Swedlow, Metadata management for high content screening in OMERO, *Methods (San Diego Calif.)* 96 (2016) 27–32, <http://dx.doi.org/10.1016/j.jmeth.2015.10.006>, <http://www.ncbi.nlm.nih.gov/pubmed/26476368>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4773399>.
- [61] R. Marée, L. Rollus, B. Stévens, R. Hoyoux, G. Louppe, R. Vandaele, J.-M. Begon, P. Kainz, P. Geurts, L. Wehenkel, Collaborative analysis of multi-gigapixel imaging data using Cytomine, *Bioinformatics* 32 (9) (2016) 1395–1401, <http://dx.doi.org/10.1093/bioinformatics/btw013>, <http://bioinformatics.oxfordjournals.org/lookup/doi/10.1093/bioinformatics/btw013>.
- [62] A. Bauch, I. Adamczyk, P. Buczek, F.-J. Elmer, K. Enimanev, P. Gylzowski, M. Kohler, T. Pylak, A. Quandt, C. Ramakrishnan, C. Beisel, L. Malmström, R. Aebersold, B. Rinn, openBIS: a flexible framework for managing and analyzing complex data in biology research, *BMC bioinformatics* 12 (1) (2011) 468, <http://dx.doi.org/10.1186/1471-2105-12-468>, <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-468>, <http://www.biomedcentral.com/1471-2105/12/468>.
- [63] K. Kvilekval, D. Fedorov, B. Obara, A. Singh, B.S. Manjunath, Bisque: a platform for bioimage analysis and management, *Bioinf. (Oxford, England)* 26 (4) (2010) 544–552, <http://dx.doi.org/10.1093/bioinformatics/btp699>, <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btp699>, <http://www.ncbi.nlm.nih.gov/pubmed/20031971>.
- [64] R.N. Smith, J. Aleksic, D. Butano, A. Carr, S. Contrino, F. Hu, M. Lyne, R. Lyne, A. Kalderimis, K. Rutherford, R. Stepan, J. Sullivan, M. Wakeling, X. Watkins, G. Micklem, InterMine: a flexible data warehouse system for the integration and analysis of heterogeneous biological data, *Bioinf. (Oxford England)* 28 (23) (2012) 3163–3165, <http://dx.doi.org/10.1093/bioinformatics/bts577>, <http://bioinformatics.oxfordjournals.org/content/28/23/3163.short>.
- [65] B.L. Millard, M. Niepel, M.P. Menden, J.L. Muhlich, P.K. Sorger, Adaptive informatics for multifactorial and high-content biological data, *Nat. Methods* 8 (6) (2011), <http://dx.doi.org/10.1038/NMETH.1600>.
- [66] C. Sommer, M. Held, B. Fischer, W. Huber, D.W. Gerlich, CellH5: a format for data exchange in high-content screening, *Bioinf. (Oxford England)* 29 (12) (2013) 1580–1582, <http://dx.doi.org/10.1093/bioinformatics/btt175>, <http://bioinformatics.oxfordjournals.org/content/29/12/1580.short>.
- [67] M. Breker, M. Gymrek, O. Moldavski, M. Schuldiner, LoQAtE-Localization and Quantitation Atlas of the yeast proteome. A new tool for multiparametric dissection of single-protein behavior in response to biological perturbations in yeast, *Nucleic Acids Res.* 42 (1) (2014) D726–D730, <http://dx.doi.org/10.1093/nar/gkt933>, <http://www.ncbi.nlm.nih.gov/pubmed/24150937>.
- [68] B. Antal, A. Chessel, R.E. Carazo Salas, Mineotaur: a tool for high-content microscopy screen sharing and visual analytics, *Genome Biol.* 16 (1) (2015) 283, <http://dx.doi.org/10.1186/s13059-015-0836-5>, <http://genomebiology.com/2015/16/1/283>.
- [69] J.C. Mollo, The open knowledge foundation: open data means better science, *PLoS Biol.* 9 (12) (2011) e1001195, <http://dx.doi.org/10.1371/journal.pbio.1001195>, <http://dx.plos.org/10.1371/journal.pbio.1001195>.
- [70] J. Li, J.Y. Newberg, M. Uhlén, E. Lundberg, R.F. Murphy, Automated analysis and reannotation of subcellular locations in confocal images from the human protein atlas, *PLoS ONE* 7 (11) (2012), <http://dx.doi.org/10.1371/journal.pone.0050514>.
- [71] X. Wang, M.A. Castro, K.W. Mulder, F. Markowitz, Posterior association networks and functional modules inferred from rich phenotypes of gene perturbations, *PLoS Comput. Biol.* 8 (6) (2012) e1002566, <http://dx.doi.org/10.1371/journal.pcbi.1002566>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3386165&tool=pmcentrez&rendertype=abstract>.
- [72] G.R. Johnson, J. Li, A. Shariff, G.K. Rohde, R.F. Murphy, Automated learning of subcellular variation among punctate protein patterns and a generative model of their relation to microtubules, *PLoS Comput. Biol.* 11 (12) (2015) e1004614, <http://dx.doi.org/10.1371/journal.pcbi.1004614>, <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004614#references>.
- [73] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [74] Y. Saeyns, I. Inza, P. Larrañaga, A review of feature selection techniques in bioinformatics, *Bioinformatics* (2007), <http://dx.doi.org/10.1093/bioinformatics/btm344>, <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btm344>.
- [75] L. van der Maaten, E. Postma, J. van den Herik, Dimensionality Reduction: A Comparative Review, *Tech. rep.*, TICC, Tilburg University (2009).
- [76] J.M. Lewis, L. Van Der Maaten, V.R. De Sa, A behavioral investigation of dimensionality reduction, *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, 2012, pp. 671–676.
- [77] J.H. Friedman, L.C. Rafsky, Multivariate generalizations of the Wald-Wolfowitz and Smirnov two-sample tests, *Ann. Stat.* 7 (14) (1979) 697–717.
- [78] B.Y.P. Hall, N. Tajvidi, Permutation tests for equality of distributions in high-dimensional settings, *Biometrika* 89 (2) (2002) 359–374.

- [79] L. Baringhaus, C. Franz, On a new multivariate two-sample test, *J. Multivariate Anal.* 88 (1) (2004) 190–206, [http://dx.doi.org/10.1016/S0047-259X\(03\)00079-4](http://dx.doi.org/10.1016/S0047-259X(03)00079-4), <http://dl.acm.org/citation.cfm?id=969852.969866>.
- [80] W. Zaremba, A. Gretton, M. Blaschko, B-test : A Non-parametric, Low variance kernel two-sample test, *Adv. Neural Inf. Process. Syst.* 26 (2013) 1–9, <http://papers.nips.cc/paper/5081-b-test-a-non-parametric-low-variance-kernel-two-sample-test.pdf>.
- [81] D. Sejdinovic, A. Gretton, B. Sriperumbudur, K. Fukumizu, Hypothesis testing using pairwise distances and associated kernels, 29th International Conference on Machine Learning (2012). 1205.0411v2.
- [82] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, *ACM Comput. Surveys (CSUR)* 41 (2009) 1–58, <http://dx.doi.org/10.1145/1541880.1541882>, <http://portal.acm.org/citation.cfm?id=1541882>, <http://dl.acm.org/citation.cfm?id=1541882>, <http://portal.acm.org/citation.cfm?doi=1541880.1541882b>.
- [83] J.D. Storey, R. Tibshirani, Statistical significance for genomewide studies, *Proc. Natl. Acad. Sci. U.S.A.* 100 (16) (2003) 9440–9445, <http://dx.doi.org/10.1073/pnas.1530509100>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=170937&tool=pmcentrez&rendertype=abstract>.
- [84] H.Z. Sailem, S. Cooper, C. Bakal, Visualizing quantitative microscopy data: history and challenges, *Crit. Rev. Biochem. Mol. Biol.* 51 (2) (2016) 96–101, <http://dx.doi.org/10.3109/10409238.2016.1146222>, <http://www.tandfonline.com/doi/full/10.3109/10409238.2016.1146222>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4819578&tool=pmcentrez&rendertype=abstract>.
- [85] D.J. Matuszewski, C. Wählby, J.C. Puigvert, I.-M. Sintorn, M. Bickle, A. Carpenter, T. Jones, M. Lamprecht, C. Clarke, I. Kang, O. Friman, A. Gough, N. Chen, T. Shun, T. Lezon, R. Boltz, C. Reese, V. Ljosa, P. Caie, R. Horst, K. Sokolnicki, E. Jenkins, S. Daya, T. Jones, I. Kang, D. Wheeler, R. Lindquist, A. Papallo, D. Sabatini, C. Sommer, D. Gerlich, G. Chan, T. Kleinheinz, D. Peterson, J. Moffat, PopulationProfiler: a tool for population analysis and visualization of image-based cell screening data, *PLOS ONE* 11 (3) (2016) e0151554, <http://dx.doi.org/10.1371/journal.pone.0151554>, <http://dx.plos.org/10.1371/journal.pone.0151554>.
- [86] J. Heer, M. Bostock, V. Ogievetsky, A Tour through the Visualization Zoo, *Commun. ACM* 53 (5) (2010) 59–67, <http://dx.doi.org/10.1145/1743546>, <http://cat.inist.fr/?aModele=afficheN&cpsidt=22906879>.
- [87] L.J.P. Van Der Maaten, G.E. Hinton, Visualizing high-dimensional data using t-sne, *J. Mach. Learn. Res.* 9 (2008) 2579–2605, <http://dx.doi.org/10.1007/s10479-011-0841-3>, [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list\\_uids=7911431479148734548related:VOiAgwMNy20J](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=7911431479148734548related:VOiAgwMNy20J).
- [88] E.-a. D. Amir, K.L. Davis, M.D. Tadmor, E.F. Simonds, J.H. Levine, S.C. Bendall, D. K. Shenfeld, S. Krishnaswamy, G.P. Nolan, D. Pe'er, viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia, *Nat. Biotechnol.* 31 (6) (2013) 545–552, <http://dx.doi.org/10.1038/nbt.2594>, <http://www.ncbi.nlm.nih.gov/pubmed/23685480>, <http://www.nature.com/doi/10.1038/nbt.2594>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4076922&tool=pmcentrez&rendertype=abstract>.
- [89] Y. Cheng, M.T. Wong, L. van der Maaten, E.W. Newell, Categorical analysis of human T cell heterogeneity with one-dimensional soli-expression by nonlinear stochastic embedding, *J. Immunol.* 196 (2016) 924–932, <http://dx.doi.org/10.4049/jimmunol.1501928>.
- [90] P. Qiu, E.F. Simonds, S.C. Bendall, K.D. Gibbs, R.V. Bruggner, M.D. Linderman, K. Sachs, G.P. Nolan, S.K. Plevritis, Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE, *Nat. Biotechnol.* 29 (10) (2011) 886–891, <http://dx.doi.org/10.1038/nbt.1991>, <http://www.nature.com/doi/10.1038/nbt.1991>, <http://dx.doi.org/10.1038/nbt.1991>.
- [91] J.H. Levine, E.F. Simonds, S.C. Bendall, K.L. Davis, E.-a. D. Amir, M.D. Tadmor, O. Litvin, H.G. Fienberg, A. Jager, E.R. Zunder, P. Finck, A.L. Gedman, I. Radtke, J.R. Downing, D. Pe'er, G.P. Nolan, Data-driven phenotypic dissection of AML, *Cell* 162 (1) (2015) 184–197, <http://dx.doi.org/10.1016/j.cell.2015.05.047>.
- [92] F. Fuchs, G. Pau, D. Kranz, O. Sklyar, C. Budjan, S. Steinbrink, T. Horn, A. Pedal, W. Huber, M. Boutros, Clustering phenotype populations by genome-wide RNAi and multiparametric imaging, *Mol. Syst. Biol.* 6 (370) (2010) 370, <http://dx.doi.org/10.1038/msb.2010.25>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2913390&tool=pmcentrez&rendertype=abstract>.
- [93] J.C. Caicedo, S. Singh, A.E. Carpenter, Applications in image-based profiling of perturbations, *Curr. Opin. Biotechnol.* 39 (2016) 134–142, <http://dx.doi.org/10.1016/j.copbio.2016.04.003>.
- [94] B.T. Sherman, D.W. Huang, Q. Tan, Y. Guo, S. Bour, D. Liu, R. Stephens, M.W. Baseler, H.C. Lane, R.A. Lempicki, DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis, *BMC Bioinf.* 8 (2007) 426, <http://dx.doi.org/10.1186/1471-2105-8-426>, <http://www.ncbi.nlm.nih.gov/pubmed/17980028>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2186358>.
- [95] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K.P. Tsafou, M. Kuhn, P. Bork, L.J. Jensen, C. von Mering, STRING v10: protein-protein interaction networks, integrated over the tree of life, *Nucleic Acids Res.* 43 (Database issue) (2015) D447–D452, <http://dx.doi.org/10.1093/nar/gku1003>, <http://www.ncbi.nlm.nih.gov/pubmed/25352553>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4383874>.
- [96] F. Markowetz, How to understand the cell by breaking it: network analysis of gene perturbation screens, *PLoS Comput. Biol.* 6 (2) (2010) 8, <http://arxiv.org/abs/0910.2938>.
- [97] K. Sachs, O. Perez, D. Pe'er, D.A. Lauffenburger, G.P. Nolan, Causal protein-signaling networks derived from multiparameter single-cell data, *Science (New York N.Y.)* 308 (5721) (2005) 523–529, <http://dx.doi.org/10.1126/science.1105809>, <http://www.ncbi.nlm.nih.gov/pubmed/15845847>.
- [98] K. Punjani, E.P. Xing, GINI: from ISH images to gene interaction networks, *PLoS Comput. Biol.* 9 (10) (2013) e1003227, <http://dx.doi.org/10.1371/journal.pcbi.1003227>, <http://dx.plos.org/10.1371/journal.pcbi.1003227>.
- [99] J. Ollion, J. Cochenne, F. Loll, C. Escudé, T. Boudier, TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization, *Bioinf. (Oxford England)* 29 (14) (2013) 1840–1841, <http://dx.doi.org/10.1093/bioinformatics/btt276>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3702251&tool=pmcentrez&rendertype=abstract>.
- [100] T. Zhao, R.F. Murphy, Automated learning of generative models for subcellular location: Building blocks for systems biology, *Cytometry Part A* 71 (12) (2007) 978–990, <http://dx.doi.org/10.1002/cyto.a.20487>.
- [101] R.F. Murphy, Building cell models and simulations from microscope images, *Methods* 96 (2016) 33–39, <http://dx.doi.org/10.1016/j.jmeth.2015.10.011>, <http://www.sciencedirect.com/science/article/pii/S1046202315301298>, <http://linkinghub.elsevier.com/retrieve/pii/S1046202315301298>.
- [102] G. Pau, T. Walter, B. Neumann, J.-K. Heriche, J. Ellenberg, W. Huber, Dynamical modelling of phenotypes in a genome-wide RNAi live-cell imaging assay, *BMC Bioinf.* 14 (1) (2013) 308, <http://dx.doi.org/10.1186/1471-2105-14-308>, <http://www.biomedcentral.com/1471-2105/14/308>.
- [103] L. Foret, J.E. Dawson, C. Collinet, A. Deutsch, L. Bruschi, M. Zerial, Y. Kalaidzidis, A general theoretical framework to infer endosomal network dynamics from quantitative image analysis, *Curr. Biol.* 22 (2012) 1381–1390.
- [104] K.W. Eliceiri, M.R. Berthold, I.G. Goldberg, L. Ibáñez, B.S. Manjunath, M.E. Martone, R.F. Murphy, H. Peng, A.L. Plant, B. Roysam, N. Stuurman, J.R. Swedlow, P. Tomancak, A.E. Carpenter, Biological imaging software tools, *Nat. Methods* 9 (7) (2012) 697–710, <http://dx.doi.org/10.1038/nmeth.2084>, <http://www.nature.com/doi/10.1038/nmeth.2084>.
- [105] L. Kametsky, T.R. Jones, A. Fraser, M.-A. A. Bray, D.J. Logan, K.L. Madden, V. Ljosa, C. Rueden, K.W. Eliceiri, A.E. Carpenter, Improved structure, function and compatibility for cellprofiler: modular high-throughput image analysis software, *Bioinformatics* 27 (8) (2011) 1179–1180, <http://dx.doi.org/10.1093/bioinformatics/btr095>, <http://bioinformatics.oxfordjournals.org/content/27/8/1179.short>.
- [106] T.R. Jones, I.H. Kang, D.B. Wheeler, R.A. Lindquist, A. Papallo, D.M. Sabatini, P. Golland, A.E. Carpenter, CellProfiler analyst: data exploration and analysis software for complex image-based screens, *BMC Bioinf.* 9 (1) (2008) 482, <http://dx.doi.org/10.1186/1471-2105-9-482>, <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-482>, <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2614436&tool=pmcentrez&rendertype=abstract>.
- [107] E. Williams, J. Moore, S.W. Li, G. Rustici, A. Tarkowska, A. Chessel, S. Leo, B. Antal, R.K. Ferguson, U. Sarkans, A. Brazma, R.E. Carazo-Salas, J. Swedlow, The image data resource: A scalable platform for biological image data access, integration, and dissemination, *bioRxiv* (2016), <http://dx.doi.org/10.1101/089359>.



# DeconvolutionLab2: An open-source software for deconvolution microscopy



Daniel Sage<sup>a,\*</sup>, Laurène Donati<sup>a</sup>, Ferréol Soulez<sup>a</sup>, Denis Fortun<sup>b</sup>, Guillaume Schmit<sup>a</sup>, Arne Seitz<sup>c</sup>, Romain Guiet<sup>c</sup>, Cédric Vonesch<sup>b</sup>, Michael Unser<sup>a</sup>

<sup>a</sup>Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

<sup>b</sup>Center for Biomedical Imaging-Signal Processing Core (CIBM-SP), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

<sup>c</sup>Biolmaging and Optics Platform, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

## ARTICLE INFO

### Article history:

Received 7 September 2016  
Received in revised form 21 December 2016  
Accepted 30 December 2016  
Available online 3 January 2017

### Keywords:

Deconvolution microscopy  
Open-source software  
Standard algorithms  
Textbook approach  
Reference datasets

## ABSTRACT

Images in fluorescence microscopy are inherently blurred due to the limit of diffraction of light. The purpose of deconvolution microscopy is to compensate numerically for this degradation. Deconvolution is widely used to restore fine details of 3D biological samples. Unfortunately, dealing with deconvolution tools is not straightforward. Among others, end users have to select the appropriate algorithm, calibration and parametrization, while potentially facing demanding computational tasks. To make deconvolution more accessible, we have developed a practical platform for deconvolution microscopy called *DeconvolutionLab*. Freely distributed, *DeconvolutionLab* hosts standard algorithms for 3D microscopy deconvolution and drives them through a user-oriented interface. In this paper, we take advantage of the release of *DeconvolutionLab2* to provide a complete description of the software package and its built-in deconvolution algorithms. We examine several standard algorithms used in deconvolution microscopy, notably: Regularized inverse filter, Tikhonov regularization, Landweber, Tikhonov–Miller, Richardson–Lucy, and fast iterative shrinkage-thresholding. We evaluate these methods over large 3D microscopy images using simulated datasets and real experimental images. We distinguish the algorithms in terms of image quality, performance, usability and computational requirements. Our presentation is completed with a discussion of recent trends in deconvolution, inspired by the results of the Grand Challenge on deconvolution microscopy that was recently organized.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Contents

1. Introduction .....	29
2. <i>DeconvolutionLab2</i> : A Java open-source software package .....	30
2.1. <i>DeconvolutionLab</i> : The original ImageJ deconvolution tool. ....	30
2.2. <i>DeconvolutionLab2</i> : The remastered Java deconvolution tool .....	30
2.2.1. Practical details. ....	31
3. Deconvolution algorithms .....	31
3.1. Image-formation model .....	31
3.2. Naive inverse filtering. ....	32
3.3. Tikhonov regularization .....	32
3.4. Regularized inverse filtering. ....	32
3.5. Landweber .....	32
3.6. Tikhonov–Miller .....	32
3.7. Fast iterative soft-thresholding. ....	32
3.8. Richardson–Lucy .....	33
3.9. Richardson–Lucy with total-variation regularization .....	33

\* Corresponding author.

E-mail addresses: [daniel.sage@epfl.ch](mailto:daniel.sage@epfl.ch) (D. Sage), [laurene.donati@epfl.ch](mailto:laurene.donati@epfl.ch) (L. Donati), [ferreol.soulez@epfl.ch](mailto:ferreol.soulez@epfl.ch) (F. Soulez), [denis.fortun@epfl.ch](mailto:denis.fortun@epfl.ch) (D. Fortun).

4.	Deconvolution in practice	33
4.1.	Image acquisition	33
4.2.	Point-spread function	33
4.3.	Setting of parameters	33
4.3.1.	Ghosts and ringing	33
5.	Experimental illustrations	34
5.1.	Synthetic data	34
5.2.	Isolated bead	35
5.3.	Widefield data	35
6.	Discussion: trends in deconvolution	36
6.1.	Blind deconvolution	39
6.2.	Space-varying deconvolution	39
	Acknowledgement	39
Appendix A.	Implementation aspects	39
A.1.	FFT Libraries	39
A.2.	Dissection of an algorithm	39
A.2.1.	Implementation of the Landweber algorithm	39
A.2.2.	Java snippet of Landweber	40
	References	40

## 1. Introduction

The widespread development of fluorescent-labeling techniques has rendered fluorescent microscopy one of the most popular imaging modalities in biology. An epifluorescence (a.k.a. widefield) microscope is indeed the simplest modality for observing cellular structures: After labelling with a fluorescent dye, the biological specimen is illuminated at the excitation wavelength. The fluorescence emission is used to form the image. A 3D acquisition of the cell is built as a z-stack of 2D images, by moving the focal plane through the sample.

Unfortunately, the resolution of 3D micrographs is intrinsically limited by the diffraction of light; structures closer than the Rayleigh criterion cannot be distinguished. For a popular fluorophore (DAPI, emission wavelength  $\lambda = 470$  nm) and for the standard numerical aperture  $NA = 1.4$  and diffraction index  $n_i = 1.51$  nm, the Rayleigh criterion predicts that it is impossible to observe details smaller than about  $0.61 \frac{\lambda}{NA} \approx 200$  nm in the lateral sections and  $2 \frac{n_i \lambda^2}{NA} \approx 700$  nm along the optical axis [1]. Thus, the resolution is anisotropic, i.e., the resolution along the depth axis is lower than the resolution in the lateral dimensions. Moreover, this resolution is usually insufficient to satisfy the current demands of biological research for the visualisation of intracellular organelles. The impact of diffraction is perceived as a blur, where fine details are obscured by the haze produced by out-of-focus light. The acquired blurred image can be mathematically modeled as the result of convolving the observed objects with a 3D point-spread function (PSF). This PSF is the diffraction pattern of the light that would be emitted from an infinitesimal point-like object and collected by the microscope. In other words, the PSF sums up the effects of the imaging setup on the observations.

There are two approaches to improve the resolution: (i) changing the microscope design to improve the shape of the PSF (e.g. confocal, multiphoton and most super-resolution microscopy modalities), (ii) numerically inverting the blurring process to enhance micrographs: the deconvolution. The ultimate goal of deconvolution is to restore the original signal that was degraded by the acquisition system (see Fig. 1). It relies on methods that have to be carefully optimized to preserve biological information. We present these methods in Section 3.

Deconvolution is a versatile restoration technique that has been found useful in various contexts such as biomedical signal processing, electro-encephalography, seismic signal (1D), astronomy (2D), or biology (3D). It performs well in 1D or 2D, but its results are the most impressive for 3D volumetric data, especially when the PSF is

large axially. In this case, 3D deconvolution has the capability to combine lateral and axial information when restoring the original signal.

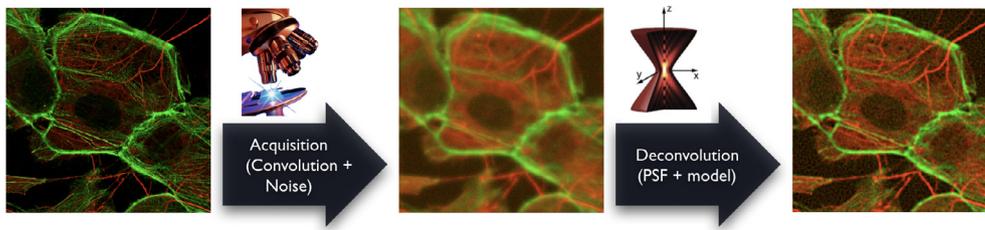
Deconvolution has multiple advantages. It is applicable to even the simplest optical setup, reducing financial costs and streamlining the acquisition pipeline. In addition to the resolution improvement, indirect benefits of deconvolution are contrast enhancement and noise reduction. As it mitigates the effect of noise, it can be used in low-light condition. The dim excitation light lowers bleaching probability of fluorophores and is therefore beneficial in terms of photo-toxicity in living cells. Not surprisingly, deconvolution is used routinely by microscopists and has become a popular pre-processing tool to further image-analysis steps such as segmentation and tracking. Unfortunately, without a proper tuning of the algorithms parameters, the deconvolved volume can be corrupted by artifacts that might prevent sound biological interpretation. Among such possible degradations, the most notable ones are noise amplification, ringing (known as Gibbs or Runge phenomenon) and aliasing (both spatial and spectral).

The deconvolution of micrographs was first investigated by Agard and Sedat [2]. Many variations and improvements have been proposed since then [3–7]. Some of these “deconvolution microscopy” methods led to various commercial and open-source software implementations [8,9]. The typical cost of a commercial package varies between USD 5000 and USD 10,000. At the time of writing this paper, the most popular ones are: Huygens (Scientific Volume Imaging); DeltaVision Deconvolution (Applied Precision, GE Healthcare Life Science); and AutoQuant (MediaCybernetics). Some of these commercial solutions (e.g., Huygens) specialize in the processing of large data and are capable of running unattended deconvolution in batch mode [10].

Meanwhile, several open-source deconvolution solutions exist too, often taking the form of an ImageJ<sup>1</sup> plugin. One of the first such platform that was made available is the popular DeconvolutionLab software developed at the Biomedical Imaging Group (EPFL) and detailed in the present paper. Freely distributed, DeconvolutionLab hosts various algorithms for 3D microscopy deconvolution and drives them through a user-oriented interface. Other open-source softwares also exist, including Nick Linnenbrügger’s DeconvolutionJ, Bob Dougherty’s Iterative Deconvolve 3D<sup>2</sup> which implements a deconvolution approach for the mapping of

<sup>1</sup> <http://imagej.nih.gov/ij/>.

<sup>2</sup> <http://www.optinav.info/Iterative-Deconvolve-3D.htm>.



**Fig. 1.** Principle of the deconvolution of a z-stack of images, presented here as the maximum-intensity projection of the volumetric data.

acoustic sources, Piotr Wendykier's *Parallel Iterative Deconvolution*<sup>3</sup> which proposes four iterative algorithms, and the *MiTiv*<sup>4</sup> project that proposes blind deconvolution software.

The deconvolution of three-dimensional data is a computationally heavy process. Fortunately, the last decades have seen a strong increase in the general accessibility to computing power. Without special equipment, it has now become possible to deconvolve data of practical size ( $512 \times 512 \times 64$ ) on a 8 GB consumer-grade computer in less than a couple minutes. Thus, the number of users having gained access to deconvolution has grown markedly through the years, which stresses the need for *accessible* and *user-friendly* software packages for deconvolution microscopy. This need is heightened by the fact that many potential users are biologists or life-science students, who may lack in computer and algorithmic literacy, so that they would have to be educated about the different available algorithms. Among others, the required skills address the selection of parameters, the control of computational and memory costs, and the recognition of restoration artifacts.

In this paper, we take advantage of the release of *DeconvolutionLab2*, the revamped sequel of *DeconvolutionLab*, to provide a complete description of the software and its built-in deconvolution algorithms. In regards to the aforementioned pedagogical aspects, the present paper equally intends as a step toward the education of inexperienced users.

## 2. *DeconvolutionLab2*: A Java open-source software package

Although microscope manufacturers may sometimes propose well-integrated software packages, their solutions are often mere black boxes. This situation prevents users to make an informed choice on which commercial deconvolution software is the most appropriate for their task at hand. Conversely, many deconvolution methods have been described in the scientific literature over the past twenty years, sometimes accompanied by open-source implementations. But even then, end users who do not master the underlying principles of deconvolution might face difficulties in selecting the method best suited to their needs. Moreover, academic packages meant to investigate some aspects of an algorithm are usually poorly designed in terms of user interface and applicable only to a specific class of signals.

At the Biomedical Imaging Group (EPFL), we have taken upon ourselves to develop the freely available software package *DeconvolutionLab*<sup>5</sup> to experiment with 3D deconvolution microscopy. *DeconvolutionLab* is a software platform that hosts various algorithms and drives them through a unified and user-friendly interface. After ten years of experience with this package, we have revamped it and renamed it *DeconvolutionLab2*. This second version keeps the same key ingredients that made the success of the

first version: Java source code, efficient FFT (fast Fourier transform), pluggable algorithms and an accommodating user interface.

### 2.1. *DeconvolutionLab*: The original *ImageJ* deconvolution tool

*DeconvolutionLab* was initially developed for educational purposes at EPFL. For over a decade it has been allowing students to conduct deconvolution experiments with the most representative classical algorithms, as well as with some more recent ones such as fast iterative soft-thresholding [11], Richardson–Lucy total variation [12], and thresholded Landweber [13]. Nowadays, we still train students with the help of *DeconvolutionLab*.

We have made *DeconvolutionLab* freely available since its release as an *ImageJ* plugin. As *ImageJ* is the *de facto* standard software tool of biological imaging, most biologists know how to install *DeconvolutionLab* on their own and can rapidly experiment with it. The package permits the deconvolution of large biological images at least as efficiently as commercial software packages [9]. With the passing years, our contribution has also gained popularity in several microscopic core facilities, where one of its favorite uses is for internal training. Moreover, from an academic perspective, *DeconvolutionLab* was deployed in more than seventy-five publications for various modalities (widefield, confocal [14], 2-photons [15], STED [16], light-sheet [14]). These works cover a wide range of applications, including neuroscience [15,17], osteology [16], microbiology [18], plant science [14] and material science [19].

### 2.2. *DeconvolutionLab2*: The remastered *Java* deconvolution tool

The present paper focuses on the complete description of *DeconvolutionLab2*, the sequel to *DeconvolutionLab*. It is a freely accessible and open-source software package running on Windows, Linux, and Mac OS operating systems. The package can be linked to well-known imaging software platforms. The backbone of the software architecture is a library that contains the number-crunching elements of the deconvolution task. The current list of built-in algorithms includes:

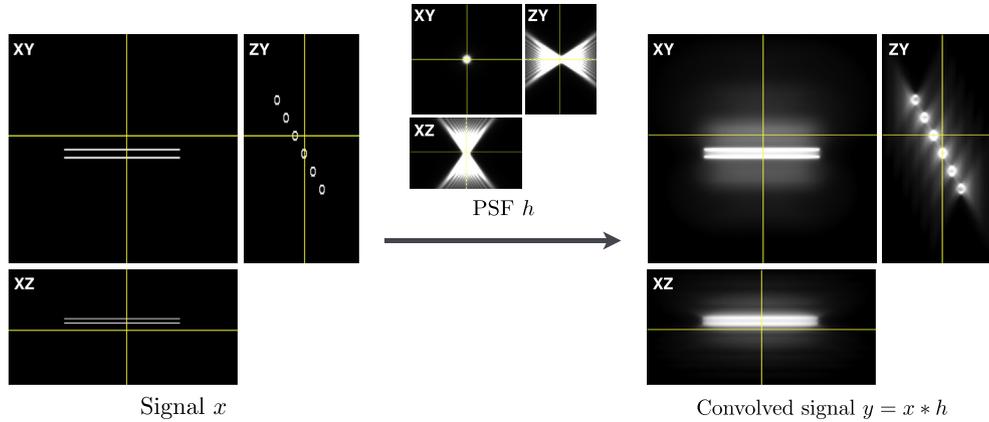
- Naive inverse filtering (NIF, Section 3.2);
- Tikhonov regularization (TR, Section 3.3);
- Regularized inverse filtering (RIF, Section 3.4);
- Landweber (LW, Section 3.5);
- Tikhonov–Miller (TM, Section 3.6);
- Fast iterative soft-thresholding (FISTA, Section 3.7);
- Richardson–Lucy (RL, Section 3.8);
- Richardson–Lucy with total-variation regularization (RL-TV, Section 3.9).

New algorithms are easily pluggable into the framework of *DeconvolutionLab2*. The source code is written in Java 1.6, as closely as possible to the text-book definition of the algorithms.

<sup>3</sup> <https://sites.google.com/site/piotrwendykier/software/deconvolution/paralleliterativedeconvolution>.

<sup>4</sup> <https://mitiv.univ-lyon1.fr/>.

<sup>5</sup> <http://bigwww.epfl.ch/deconvolution/>.



**Fig. 2.** Visualization of the convolution of simulated tubes with a PSF defined by the Born & Wolf model.

Inquisitive minds inclined to peruse the code will find it fosters the understanding of deconvolution.

Our goal with `DeconvolutionLab2` is to make deconvolution broadly accessible to the community of all those who show interest in this technique. To achieve such a goal, we provide a user-friendly front-end interface that also accommodates non-experts. Our software package is able to process large volumes on a mid-range desktop computer, or even on a laptop computer.

To experiment with the software, we share test data on the `DeconvolutionLab2` website. These data include synthetic and real cases to help benchmarking algorithms. `DeconvolutionLab2` can act not only as a didactic tool equipped with a simulator (convolution and noise generator), but also as a validation module that gives access to the signal-to-noise ratio between a ground-truth image and the output of every algorithm.

Like `DeconvolutionLab`, `DeconvolutionLab2` is able to process data relevant to real biological applications. However, and contrarily to the commercial software packages, our tools are restricted to deconvolution alone. We intentionally apply neither pre-processing nor postprocessing. Compared to `DeconvolutionLab`, `DeconvolutionLab2` includes new fast Fourier transform (FFT) libraries (see [Appendix A.1](#)), a recordable macro for ImageJ, new apodization functions, new padding schemes, and new switchable constraints in the space domain.

### 2.2.1. Practical details

`DeconvolutionLab2` is delivered as a plugin for ImageJ [20], for Fiji [21], and for the new bio-imaging platform Icy [22]. Since it is a Java class, it is also callable from the MATLAB command line and runnable as a standalone application through a Java Virtual Machine. For batch processing, we recommend calling `DeconvolutionLab2` from an ImageJ macro. This key feature enables one to handle time-lapse images or multiple channels, which need to be processed individually, in sequence.

Deconvolution is a heavy computational task in terms of running time and memory usage. In `DeconvolutionLab2`, we tried to find the best tradeoff between computational efficiency and code readability. The deconvolution is implemented in the discrete Fourier domain, so that the most time-consuming task is the FFT. Some iterative algorithms may require several FFT at every iteration, which can consume more than nine tenth of the runtime. Therefore, it is of utmost importance to rely on efficient FFT libraries.

## 3. Deconvolution algorithms

In this section, we recall the basic principles of image formation in fluorescence microscopy and give a brief technical description of

the algorithms implemented in `DeconvolutionLab2`. We focus on the impact of the underlying models and the influence of the parameters. For an in-depth understanding and a more complete overview of the deconvolution field, we refer to the reviews [3,5,6,23] that cover most of the methods described here.

### 3.1. Image-formation model

Fluorescence microscopes are often assumed to be shift-invariant, which means that the response of the system does not depend on the position in the image. Therefore, they can be characterized by a PSF which approximates the distortions of the signal in the optical system. More elaborated approximations (e.g. spatially varying PSF) are described in Section 6.2). From a signal-processing point of view, the acquisition of images is modeled as the convolution of the fluorophore distribution  $x$  in the observed volume with the PSF  $h$ , followed by a degradation by noise. The convolution operation is defined at a given 3D location  $p \in \mathbb{R}^3$  by

$$(x * h)(\mathbf{p}) = \int_{\mathbb{R}^3} x(\mathbf{r})h(\mathbf{p} - \mathbf{r}) d\mathbf{r}. \quad (1)$$

In epifluorescence microscopy, the shape of the PSF in the image domain, shown in Fig. 2 with the Born and Wolf model [24,25], is typically such that it produces an anisotropic blurring of the signal. The resolution of the convolved signal is usually three times lower in the axial direction than in the lateral plane.

From now on we consider a discretized model. We denote by  $\mathbf{y} \in \mathbb{R}^N$  the observed volume in vector form,  $\mathbf{x} \in \mathbb{R}^K$  the underlying fluorescence signal, and  $\mathbf{H} \in \mathbb{R}^{N \times K}$  the PSF matrix defined such that the discretization of the convolution defined in Eq. 1 writes as the matrix multiplication  $\mathbf{H}\mathbf{x}$ . Possibly, we may want to perform the reconstruction at an output resolution that differs from the input resolution, or to handle carefully border effects by estimating an image  $\mathbf{x}$  with larger size, whereby  $K \neq N$ .

For a circulant and shift invariant discrete PSF with  $K = N$ , the matrix-vector multiplication  $\mathbf{H}\mathbf{x}$  becomes an element-wise multiplication in the Fourier domain:  $\hat{\mathbf{y}} = \hat{\mathbf{h}} \times \hat{\mathbf{x}}$  where  $\hat{\mathbf{y}}$  and  $\hat{\mathbf{x}}$  are the discrete Fourier transform coefficients of  $\mathbf{y}$  and  $\mathbf{x}$ , and  $\hat{\mathbf{h}}$  are the coefficients of the discrete Fourier transform of the PSF. This permits efficient computation of  $\mathbf{H}\mathbf{x}$ , both in terms of speed and memory requirements through the use of the fast Fourier transform (FFT) algorithms. Every deconvolution algorithm we present in this paper relies on this technique.

The discrete image acquisition model is then

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (2)$$

with  $\mathbf{n} \in \mathbb{R}^N$  an additive noise component. The acquired images are affected by several sources of noise, which are often modeled by

two components. The first component is signal-dependent and models the fluctuation of the number of photons arriving at a given pixel. This so-called shot noise follows a Poisson distribution whose mean depends on the intensity of the incoming light. The second component accounts for various other distortions such as a background signal, read-out noise, or quantization noise, which are usually modeled as additive Gaussian noise. Note that in the case of Poisson noise, the variable  $\mathbf{n}$  depends on the data  $\mathbf{y}$  in Eq. (2). We decided to drop this dependency for the sake of clarity of the notations.

The aim of deconvolution algorithms is to invert the noisy convolution process defined in Eq. 2, thereby producing an estimated image  $\hat{\mathbf{x}}$  from the knowledge of  $\mathbf{y}$  and  $\mathbf{H}$ , and the assumptions about the noise  $\mathbf{n}$ .

### 3.2. Naive inverse filtering

The simplest approach to deconvolution consists in minimizing a least-squares cost function  $\mathcal{C}(\mathbf{x})$  that measures the similarity between the observation  $\mathbf{y}$  and the current estimate  $\mathbf{H}\mathbf{x}$ , so that

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \mathcal{C}(\mathbf{x}) \quad (3)$$

$$\text{with } \mathcal{C}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \quad (4)$$

We call it naive inverse filtering. It corresponds to maximum-likelihood estimation in the presence of Gaussian noise. The solution can be computed efficiently in the Fourier domain and amounts to the coefficient-wise division

$$\hat{\mathbf{x}} = \frac{\hat{\mathbf{y}}}{\max(\hat{\mathbf{h}}, \epsilon)}, \quad (5)$$

where  $\max$  denotes the element-wise maximum and  $\epsilon$  is a small constant to avoid divisions by zero. The final solution is then obtained by taking the inverse Fourier transform of  $\hat{\mathbf{x}}$ .

The method is parameter-free and the direct inversion in the Fourier domain leads to fast computations. Unfortunately, the NIF tends to amplify measurement noise, resulting in spurious high-frequency oscillations.

### 3.3. Tikhonov regularization

A way to avoid the noise amplification of NIF is to add to the cost function (4) the regularization term  $\|\mathbf{x}\|_2^2$  to penalize high values of the solution [26]. This leads to

$$\mathcal{C}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda\|\mathbf{x}\|_2^2, \quad (6)$$

where  $\lambda$  is a regularization parameter that balances the contribution of the two terms. The explicit minimizer of (6) is

$$\mathbf{x} = (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{H}^T\mathbf{y}, \quad (7)$$

where  $\mathbf{I}$  is the identity matrix, and  $\mathbf{H}^T$  denotes the adjoint of  $\mathbf{H}$ . As for NIF, the solution (7) can be computed directly in the Fourier domain. This formulation can also be interpreted as a *maximum a posteriori* model. There, the regularization introduces prior information about the signal to guide the estimation.

### 3.4. Regularized inverse filtering

Other types of regularizations than TR can be used. A popular approach that performs well is to impose smoothness on  $\mathbf{x}$  by penalizing the energy of its derivative. The resulting cost function is

$$\mathcal{C}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda\|\mathbf{L}\mathbf{x}\|_2^2, \quad (8)$$

where  $\mathbf{L}$  is a matrix that corresponds to the discretization of a differential operator. In `deconvolutionLab2`, we use the Laplacian operator. The explicit minimizer of (8) is given by

$$\mathbf{x} = (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{H}^T\mathbf{y}. \quad (9)$$

When the filtering by  $\mathbf{L}^T\mathbf{L}$  has a whitening effect on  $\mathbf{x}$  and  $\lambda$  is defined as the inverse of the noise variance, RIF amounts to Wiener filtering [27].

### 3.5. Landweber

The LW algorithm minimizes the same least-squares cost function than NIF. But, instead of expressing the solution through direct inversion, it resorts to an iterative gradient-descent approach [28]. In `DeconvolutionLab2`, we take advantage of the iterative nature of LW to impose a nonnegativity constraint at each iteration. Each update indexed by  $k$  can be written as

$$\mathbf{x}^{(k+1)} = \mathcal{P}_{(\mathbb{R}^+)^K} \left\{ \mathbf{x}^{(k)} + \gamma\mathbf{H}^T(\mathbf{y} - \mathbf{H}\mathbf{x}^{(k)}) \right\}, \quad (10)$$

where  $\gamma$  is a step size parameter and  $\mathcal{P}_{(\mathbb{R}^+)^K} \{\mathbf{x}\} = \max(\mathbf{x}, 0)$  is the component-wise projection of  $\mathbf{x}$  onto the set  $(\mathbb{R}^+)^K$ .

Minimizing the energy (4) only enforces data fidelity of the result. The consequence is that the solution at convergence of iterations (10) tends to produce an over-fitting of the noise in the input data. However, one may obtain a satisfactory tradeoff between deconvolution and noise amplification if the algorithm is stopped early. In fact, the number of iterations is made to act as a pseudo regularization parameter. This phenomenon occurs for all maximum-likelihood based algorithms.

### 3.6. Tikhonov–Miller

Similarly with the LW method, the *TM* algorithm uses iterative gradient descent to minimize the regularized inverse filter cost (8). The iterative procedure is

$$\mathbf{x}^{(k+1)} = \mathcal{P}_{(\mathbb{R}^+)^K} \left\{ \mathbf{x}^{(k)} + \gamma(\mathbf{H}^T\mathbf{y} - (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{L}^T\mathbf{L})\mathbf{x}^{(k)}) \right\}. \quad (11)$$

When iterative projections onto the set  $(\mathbb{R}^+)^K$  are performed, the method is sometimes referred to as iteratively constrained Tikhonov–Miller (ICTM).

### 3.7. Fast iterative soft-thresholding

Alternative regularization terms to the one in (8) can be considered. In particular, sparsity constraints in the wavelet domain have proven to yield better preservation of image details and discontinuities. The associated cost function is

$$\mathcal{C}(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda\|\mathbf{W}\mathbf{x}\|_1, \quad (12)$$

where  $\mathbf{W}$  represents a wavelet transform. Due to the non-smoothness of the  $\ell_1$  norm, gradient-descent algorithms cannot be used. However, the problem can be solved efficiently by fast iterative soft-thresholding [11] with the following iterations:

$$\mathbf{z}^{(k+1)} = \mathbf{s}^{(k)} - \gamma\mathbf{H}^T(\mathbf{H}\mathbf{s}^{(k)} - \mathbf{y}) \quad (13)$$

$$\mathbf{x}^{(k+1)} = \mathbf{W}^T \mathcal{F}(\mathbf{W}\mathbf{z}^{(k+1)}, \gamma\lambda), \quad (14)$$

$$p^{(k+1)} = \frac{1}{2} \left( 1 + \sqrt{1 + 4p^{(k)2}} \right) \quad (15)$$

$$\mathbf{s}^{(k+1)} = \mathbf{x}^{(k+1)} + \frac{p^{(k)} - 1}{p^{(k+1)}} (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}). \quad (16)$$

There,  $\gamma$  is a step size that can be determined explicitly to ensure convergence [11], and  $\mathcal{F}(\cdot, \tau)$  is a soft-thresholding operator with threshold  $\tau$ .

### 3.8. Richardson–Lucy

The RL method [29,30] is a maximum-likelihood approach, like NIF. The difference is that RL assumes that the noise follows a Poisson distribution, which leads to

$$\mathcal{C}(\mathbf{x}) = \mathbf{1}^T \mathbf{H}\mathbf{x} - \mathbf{y}^T \log(\mathbf{H}\mathbf{x}), \quad (17)$$

where the log operation is applied component-wise and  $\mathbf{1} = (1, \dots, 1) \in \mathbb{N}^N$ . The iterative minimization of (17) can be understood as a multiplicative gradient descent and writes

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} \times \mathbf{H}^T \left( \frac{\mathbf{y}}{\mathbf{H}\mathbf{x}^{(k)}} \right), \quad (18)$$

where the multiplication  $\times$  and the division  $\mathbf{y}/(\mathbf{H}\mathbf{x}^{(k)})$  are understood to be component-wise.

Since the updates of  $\mathbf{x}$  are multiplicative, nonnegativity is naturally ensured by the algorithm for any nonnegative starting point. As a maximum-likelihood method, the solution of RL is subject to the same noise-amplification problem as NIF and LW. Thus, the optimal number of iterations should be heuristically set to stop the algorithm before convergence.

### 3.9. Richardson–Lucy with total-variation regularization

To counterbalance the noise amplification effect of RL, a regularization term can be added to (17) [12]. The total-variation (TV) regularizer penalizes the  $\ell_1$  norm of the gradient of the signal, with

$$\mathcal{C}(\mathbf{x}) = \mathbf{1}^T \mathbf{H}\mathbf{x} - \mathbf{y}^T \log(\mathbf{H}\mathbf{x}) + \lambda \|\mathbf{D}\mathbf{x}\|_1. \quad (19)$$

There,  $\mathbf{D}$  is the finite-difference matrix for first-order derivatives. In [12], a differentiable approximation of the  $\ell_1$  norm is used and the multiplicative iterations are expressed as

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} \times \mathbf{H}^T \left[ \frac{\mathbf{y}}{\mathbf{H}\mathbf{x}^{(k)}} \right] \times \frac{1}{1 + \lambda \mathbf{g}^{(k)}}, \quad (20)$$

where  $\mathbf{g}^{(k)}$  is the derivative of a regularized version of the  $\ell_1$  norm of  $\mathbf{D}\mathbf{x}^{(k)}$ .

Compared to the  $\ell_2$  penalization used in (8), the  $\ell_1$  norm yields piecewise-constant results that better preserve image discontinuities.

## 4. Deconvolution in practice

### 4.1. Image acquisition

The preparation of samples and the design of the imaging system are of paramount importance to a successful deconvolution. In particular, it is critical to take into account elements of the imaging system such as calibration, sampling, and noise level. These practical issues have been well considered in the literature [4,31]. Specifically for deconvolution, it is also recommended to validate the acquisition and the further processing of *known* samples to avoid false interpretations, especially in the context of quantitative imaging assays [32,33].

### 4.2. Point-spread function

The quality of the deconvolution relies on the accuracy of the 3D PSF, which is the optical signature of an (ideally infinitesimally small) point. It is affected mostly by the objective, the medium, and

**Table 1**  
Important deconvolution parameters per method.

Method	Parameters	Section
NIF	–	3.2
TR	$\lambda$	3.3
RIF	$\lambda$	3.4
L	$M_{\text{iter}}, \gamma$	3.5
TM	$M_{\text{iter}}, \gamma, \lambda$	3.6
FISTA	$M_{\text{iter}}, \lambda, \gamma$	3.7
RL	$M_{\text{iter}}$	3.8
RL-TV	$M_{\text{iter}}, \lambda$	3.9

the coverslip. A PSF can be obtained either experimentally or theoretically.

An experimental PSF can be deduced from the acquisition of the z-stack of a sparse field of spherical beads of very small diameter (e.g., 100 nm). Regions of interest are cut in the data centered around several well-contrasted beads and averaged. Microscopists generally agree that the experimental PSF captures well the aberrations of the microscope, but that the resolution of an experimental PSF is tied to the resolution of the acquisition. Unless super-resolution methods are deployed, this enforces  $N = K$  (see Section 3.1).

By contrast, a theoretical PSF can be computed from a mathematical model. In addition to being able to lift the restriction on resolution, microscopists generally appreciate the convenience of software packages like `PSFGenerator`<sup>6</sup> that allow them to tune freely microscope parameters such as numerical aperture (NA), wavelength, and pixelsize [25].

### 4.3. Setting of parameters

A few important parameters are shared by groups of deconvolution methods. In this section, we give practical hints about the meaning and the impact of the main parameters for each group. We provide the parameters-per-method associations in Table 1.

**Regularization parameter  $\lambda$ .** When the cost function contains a regularization term weighted by  $\lambda$ , the value of  $\lambda$  balances the contribution of data fidelity and regularization. For algorithms with Tikhonov-type regularization, higher values of  $\lambda$  result in smoother images. Finally, by setting  $\lambda = 0$ , TR and RIF become equivalent to NIF, and RL-TV becomes equivalent to RL.

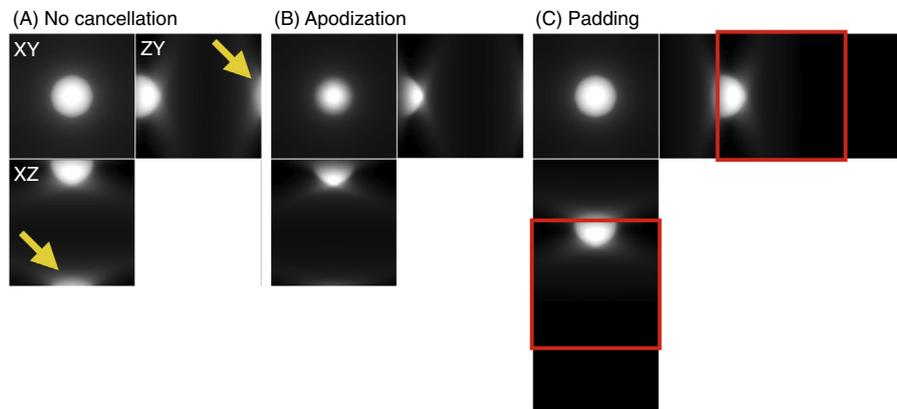
**Number  $M_{\text{iter}}$  of iterations.** For all iterative methods,  $M_{\text{iter}}$  puts a cap on the number of iterations. How to set  $M_{\text{iter}}$  follows one of two rules: either the deconvolution method is known to reach the desired solution at convergence, in which case  $M_{\text{iter}}$  has to be chosen large enough; or noise amplification happens during convergence, in which case  $M_{\text{iter}}$  has to be chosen small enough so that the deconvolution procedure stops before noise dominates. In the latter case, the choice of the appropriate  $M_{\text{iter}}$  has a crucial impact on the result.

**Stepsize.** Methods based on gradient descent rely on a stepsize  $\gamma \in (0, 1]$  which determines the speed of convergence. Small values of  $\gamma$  encourage safe but slow convergence.

#### 4.3.1. Ghosts and ringing

Every deconvolution algorithm presented in Section 3 relies partly on circular convolutions computed through FFT. Compared to space-based approaches, Fourier-based approaches reduce the computational cost of handling a PSF that would have a wide spatial support. The downside is the appearance of Fourier-related artifacts such as ghosts and ringing.

<sup>6</sup> <http://bigwww.epfl.ch/algorithms/psfgenerator/>.



**Fig. 3.** Illustration of border artifacts after a deconvolution operation on a bead placed on the top of the volume of size  $128 \times 128 \times 128$  pixels. The illustrations are presented as orthogonal sections. (A) Deconvolution without artifact-cancellation processing was applied on the signal; the arrow shows the impact of ghosting. (B) Deconvolution with Hann apodization along the axial direction. (C) Deconvolution with a zero-padding extension to  $(128 \times 128 \times 256)$  pixels (only the red surrounding of the signal will be kept).

Data subjected to a FFT must necessarily be assumed to be periodic. This implies that borders at opposite sides of the image are implicitly abutting once periodization is taken into account. Consequently, structures near the border of an image, once processed, will spill over the opposite border, letting ghosts appear.

Data subjected to a FFT must necessarily be assumed to be bandlimited. This implies that the sharp transitions of intensities found in an image (i.e., the edges), once processed, will incur local overshoots and undershoots of intensity. This mechanism is called ringing. Nonnegativity constraints may help cancel this artifact, but only with regard to undershoots, and only for those undershoots that would otherwise result in negative values. Nonnegativity, commonly positivity, therefore makes a lot of sense in fluorescence microscopy.

Inconveniently, Fourier-related artifacts frequently appear, particularly in the axial direction since this direction is often sampled to a lesser extent than the lateral ones. For instance, if a biological cell physically extends outside of the bottom of the acquired volume and is thus virtually cropped at acquisition time, then a reverse ghost of the cell will appear on the top part of the volume after deconvolution. At the same time, ringing artifacts will reveal themselves as waves in the background and as Gibbs phenomena in the high-contrast areas.

To attenuate these artifacts, we have implemented two countermeasures in *DeconvolutionLab2*: apodization and zero-padding. Apodization consists in multiplying the input data by a window function that gradually sets the signal to zero near the borders of the image. Depending on the window specifics, the central part of the data may or may not remain pristine. In *DeconvolutionLab2*, we have made available the five classical apodization functions referred to as Cosine, Hamming, Hann, Tukey, and Welch. They can be applied independently along the axial and the lateral directions. As shown in Fig. 3(B), apodization succeeds in cancelling the ghost object, but also reduces the intensity of the data.

While it modifies the data, apodization proceeds without a change in the image dimensions. Conversely, zero-padding maintains the data intact but modifies the dimensions of the image by extending its border with zero values. For practical reasons related to the computational efficiency of the FFT, the width of the extension is generally chosen such that the size of the extended image is highly decomposable as a product of small prime numbers. To facilitate adherence to this constraint, *DeconvolutionLab2* automatically proposes extensions to the next even number, to the next multiple of 2 and 3, to the next multiple of 2, 3, and 5, and to the next power of 2, independently in the axial and the lat-

eral directions. As shown in Fig. 3(C), zero-padding succeeds in cancelling the ghost object, but does so at an increased computational cost compared to apodization.

## 5. Experimental illustrations

We now illustrate the performance of *DeconvolutionLab2* and its built-in algorithms by restoring various types of degraded 3D images (i.e., synthetic volumes, beads, and real volumes). Visualizations of the deconvolution results are provided and quantitative measurements are reported when available. The data, as well as the corresponding model of the theoretical PSF, are available online<sup>7</sup>.

### 5.1. Synthetic data

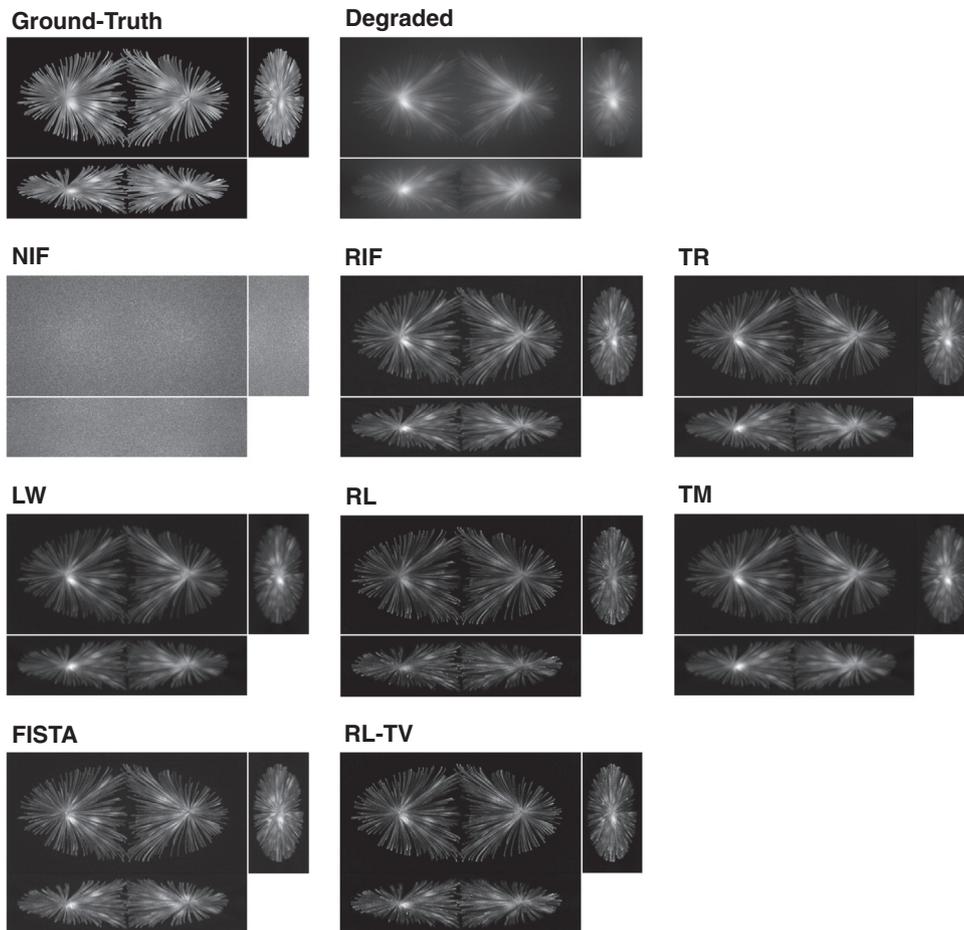
We applied all *DeconvolutionLab2* algorithms on a synthetically degraded volume. The ground-truth data consisted of a stack of 128 axial views of size  $512 \times 256$  pixels depicting cellular microtubules. To mimic the acquisition artifacts of classical wide-field microscopes, blurring and noise were generated on the ground-truth volume through the *Convolution* tool of *DeconvolutionLab2*. More precisely, the 3D data was convolved with a theoretical PSF and a mixture of Gaussian and Poisson noise was added to the volume.

The effect of the deconvolution algorithms is illustrated in Figs. 4 and 5, while the quantitative measurements after deconvolution are reported in Table 2. The visual and quantitative outputs lead to similar observations.

Firstly and most obviously, the severe artifacts introduced by the NIF algorithm lead to non-exploitable results. The introduction of regularization (TR, RIF) enables decent deconvolution results, but the presence of undesirable ringing artifacts still hinder correct visualization of the imaged structure. As supported by Table 2, the beneficial effect of deconvolution increases when classical iterative algorithms (LW, RL, TM) are applied. However, the cost of doing so translates into an augmentation of the required computational resources.

Finally, the more advanced methods (FISTA, RL-TV) were also applied to the data. Interestingly, although RL-TV is theoretically more sophisticated than RL, the algorithm yields similar deconvolution results when applied to the present data. This can be explained by the fact that the structure of the considered object

<sup>7</sup> <http://bigwww.epfl.ch/deconvolution/>.



**Fig. 4.** Orthogonal sections of the maximum intensity projection (MIP) of a degraded 3D synthetic volume after its deconvolution by *DeconvolutionLab2* algorithms. From top left to bottom right: Ground-truth volume, Degraded volume (*i.e.*, simulated acquisition), Naive Inverse Filter, Tikhonov regularization (low regularization), Regularized Inverse Filter (low regularization), Landweber ( $\tau = 1.0$ , 2000 iterations), Richardson–Lucy (150 iterations), Tikhonov–Miller (low regularization,  $\tau = 1.5$ , 150 iterations), FISTA (low regularization,  $\tau = 1.5$ , 50 iterations), Richardson–Lucy with TV (low regularization, 150 iterations). The data, as well as the corresponding PSF, are available online. A non-negativity constraint was used for all algorithms. The setting of the optimal parameters for each deconvolution algorithm was performed through visual assessment.

imposes a negligible level of regularization during deconvolution. Indeed, the synthetic sample harbors thin filament-like structures which are difficult to recover through a TV regularizer, since TV tends to promote piece-wise constant surfaces. This illustrates the fact that the efficiency of a certain deconvolution algorithm may vary with the type of the data being processed. Thus, one cannot straightforwardly use the results presented above as a direct indicator of the individual performance of each deconvolution algorithm. Moreover, depending on the data size, time constraints and the available computational resources, some less advanced methods may be better suited for the deconvolution task at hand.

### 5.2. Isolated bead

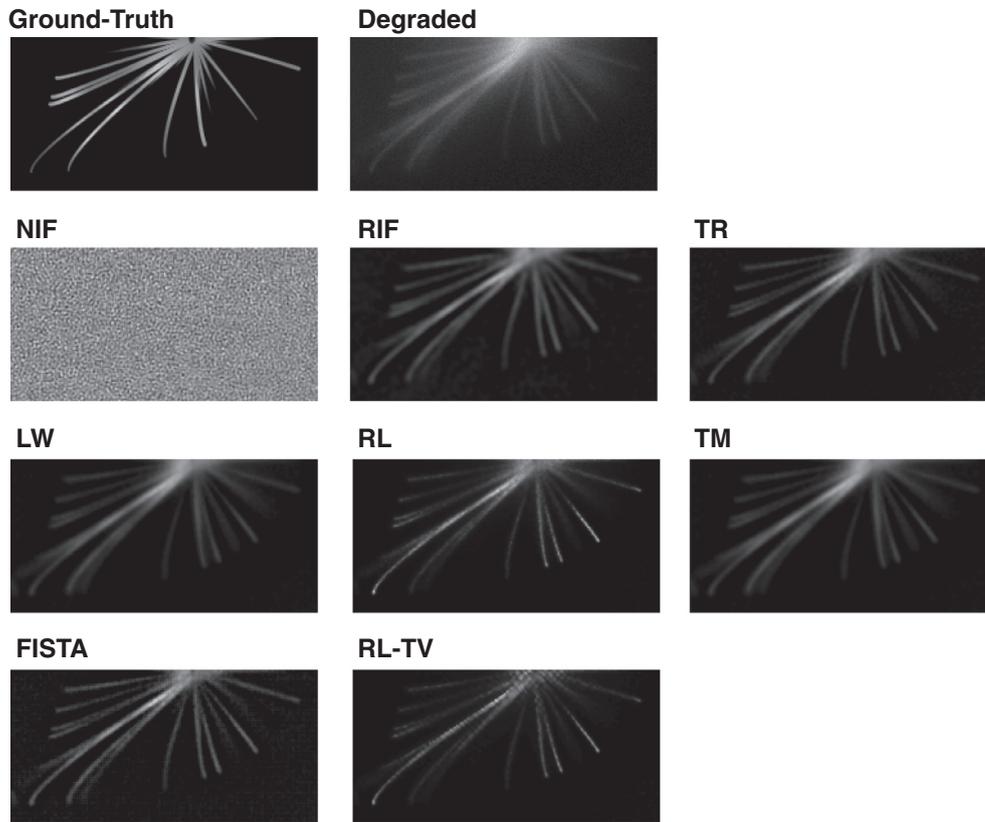
We apply several algorithms of *DeconvolutionLab2* on a z-stack called “Bead” [9]. The volume displays a single fluorescent bead, which corresponds to a sphere with known diameter of 2.5  $\mu\text{m}$ . The z-stack was acquired on a standard widefield microscope ( $\lambda = 530 \text{ nm}$ ,  $\text{NA} = 1.4$ ); the lateral pixel size is 64.5 nm and the stepsize in the axial direction is 160 nm.

The effect of the deconvolution algorithms is illustrated in Fig. 6, while the measurements of the full width at half maximum (FWHM) of the bead in the lateral and axial directions after deconvolution are reported in Table 3. We first observe that the NIF algo-

gorithm is not able to recover the bead. For the RIF algorithm, the effect of regularization on the deconvolution process becomes evident. Blurred images and overestimated dimensions are observed when the RIF regularization factor is overly increased, while setting it too low generates ringings. For the LW algorithm, the best results are obtained with 64 iterations. When the number of iterations is insufficiency, the effect of deconvolution is imperceptible. By contrast, using a too large number of iterations leads to high frequency artifacts appearing near the contour of the bead. This simple dataset thus illustrates the importance of the selection of the parameters for a given deconvolution method.

### 5.3. Widefield data

Finally, we briefly illustrate how *DeconvolutionLab2* may be used in a practical application to efficiently deconvolve real bio-microscopy data. We work with a 3D visualization of a *C. elegans* embryo which was acquired on a standard wide-field microscope ( $\lambda = 542 \text{ nm}$ ,  $\text{NA} = 1.4$ ); the lateral pixel size is 64.5 nm and the stepsize in the axial direction is 160 nm. As shown in Fig. 7, our 3D measurement displays some non-desirable visual features, such as a relatively poor contrast or an indistinguishability of certain neighboring centrosomes.



**Fig. 5.** Zooms on XY-views of a degraded synthetic volume after its deconvolution by `DeconvolutionLab2` algorithms. From top left to bottom right: Ground-truth volume, Degraded volume (i.e., simulated acquisition), Naive Inverse Filter, Regularized Inverse Filter (low regularization), Tikhonov regularization (low regularization), Landweber ( $\tau = 1.0$ , 2000 iterations), Richardson–Lucy (150 iterations), Tikhonov–Miller (low regularization,  $\tau = 1.5$ , 150 iterations), FISTA (low regularization,  $\tau = 1.5$ , 50 iterations), Richardson–Lucy with TV (low regularization, 150 iterations). The data, as well as the corresponding PSF, are available online. The zoom corresponds to a cropping with positions (244, 128, 238, 119) on the 64th z-slice. A non-negativity constraint was used for all algorithms. The setting of the optimal parameters for each deconvolution algorithm was performed through visual assessment.

**Table 2**

Quality and computational efficiency of the `DeconvolutionLab2` algorithms for the deconvolution of degraded 3D synthetic data. For comparison, the results of a widely-used commercial software (Huygens) and L2D-A3D (the “Learn 2D, Apply 3D” method [34] that won “3D Deconvolution Microscopy” challenge) were also added into this table. To assess the deconvolution performance, the signal-to-noise ratio (SNR), the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) were computed after an initial normalization of the volumes in ImageJ. Indications of the computation time and the memory ratio values are reported to allow for comparison of the computational complexity of the available algorithms. The “Required RAM” is the peak of allocated memory to run the algorithm on an input dataset of 16,000,000 voxels. The “Memory Ratio” corresponds to the ratio between the required memory and the number of voxels of the input dataset. The deconvolution tasks were performed on a Mac OS X  $2 \times 3.06$  GHz 6-Core Intel Xeon for all algorithms except for the Huygens software that was run on a 48-core server on Linux Red hat Enterprise

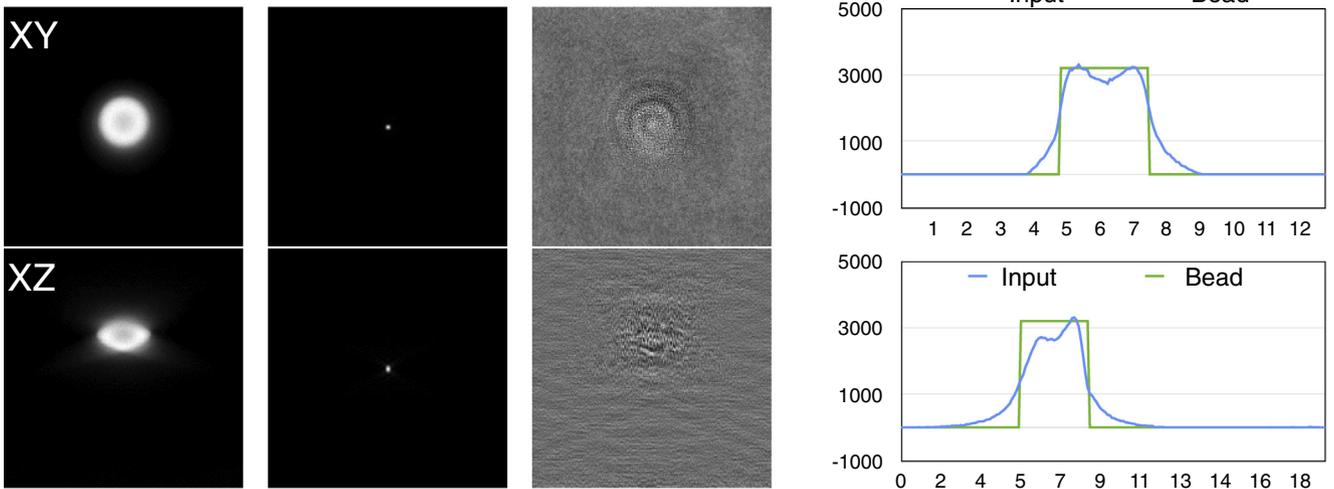
Algorithm	SNR [dB]	PSNR [dB]	SSIM [-]	Time [s]	Required RAM [Mb]	Memory Ratio
NIF	-75.45	-49.79	6.29e-9	7.6	258	$\times 16.1$
RIF	3.47	29.13	3.41e-2	7.0	322	$\times 20.1$
TR	2.78	28.45	2.48e-2	6.4	258	$\times 16.1$
LW	2.57	28.23	2.06e-2	2107	888	$\times 55.5$
FISTA	3.37	29.04	3.87e-2	1400	599	$\times 37.4$
TM	2.56	28.22	2.05e-2	2128	1016	$\times 63.5$
RL	3.66	29.33	3.30e-2	1661	258	$\times 16.1$
RL-TV	3.36	29.03	3.34e-2	2759	621	$\times 38.8$
Huygens (CMLE)	2.47	28.13	1.84e-2	180	n/a	n/a
L2D-A3D	7.27	32.94	6.73e-2	7200	n/a	n/a

To enhance the visual condition of this measurement, we apply three distinct deconvolution algorithms (TR, LW, RL) to it. The results after deconvolution are shown in Fig. 7. Globally, we observe similar effects than with the previous data sets. For all algorithms, the deconvolution permits a notable increase of the sharpness of the imaged structures and reduces out-of-focus blurring. Moreover, the iterative algorithms (LW, RL) yield better results than basic methods (RIF) at the cost of a more expensive computational need.

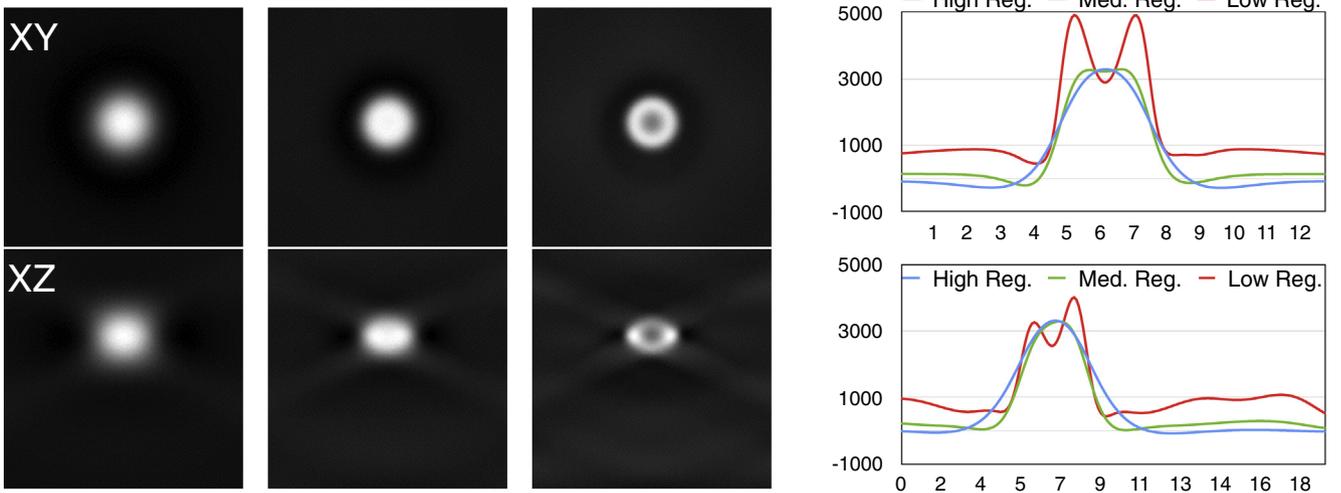
## 6. Discussion: trends in deconvolution

Similarly to many inverse problems, deconvolution requires one to express and minimize a cost function. As exemplified in Eqs. (6), (8), (12) and (19), the common form taken by this cost function is composed of a data-fidelity term that measures how well the model  $\mathbf{H}\mathbf{x}$  represents the data  $\mathbf{y}$ , and a regularization function that enforces some priors. Deconvolution methods are thus characterized by three ingredients: (i) data-fidelity measure; (ii)

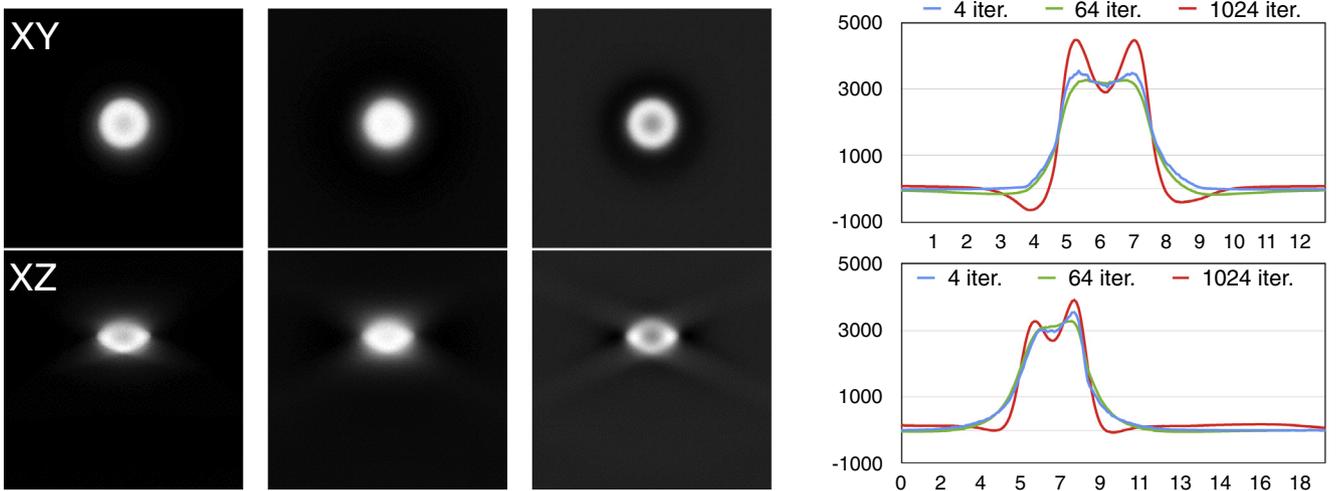
(A) Input / PSF / Naive Inverse Filter



(B) Regularized Inverse Filter



(C) Landweber

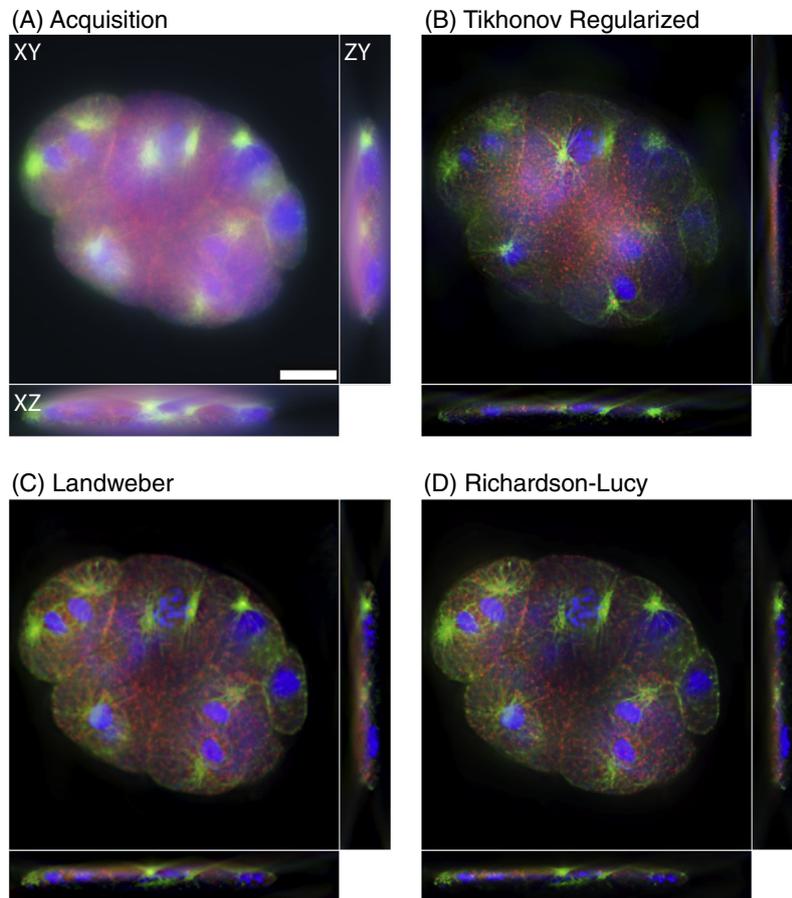


**Fig. 6.** Two orthogonal sections (XY and XZ) of the volumetric data before and after deconvolution. The plots show intensity profiles, the upper plot of a panel is the lateral profile through the bead; the lower plot is the axial profile. The unit is  $\mu\text{m}$ . (A) From left to right: input image, PSF, and the result of the NIF algorithm. Plots show the intensity profile of the input (blue line) and theoretical shape of the bead (green line). (B) Results of the RIF algorithm with various settings. From left to right: low level of regularization (Low Reg.), medium level of regularization (Med Reg.), and high level of regularization (High Reg.). (C) Results of the Landweber algorithm with various numbers of iterations. From left to right: 4 iterations, 64 iterations, and 1024 iterations.

**Table 3**

Lateral FWHM and axial FWHM of the bead measure on line profiles for the input image (upper row) and for various algorithms and settings.

Algorithm	Settings	Lateral FWHM [nm]	Axial FWHM [nm]
Acquisition		2695.33	3979.46
RIF	Reg: Low	2630	5909
RIF	Reg: Medium	2616	4881
RIF	Reg: High	2716	4900
Landweber	4 iterations	2714	4624
Landweber	64 iterations	2711	4777
Landweber	1024 iterations	2605	4449



**Fig. 7.** Orthogonal sections of the *C. Elegans* volume (size:  $672 \times 712 \times 104$  voxels). For better visualization, a Gamma correction have been applied to the images. Scale bar is  $10 \mu\text{m}$ . The data are available online. (A) Acquisition. (B) Tikhonov Regularized. (C) Landweber deconvolution (200 iterations). (D) Richardson Lucy deconvolution (200 iterations). Advanced iterative algorithms permit better distinction between two neighboring centrosomes.

regularization prior; and (iii) minimization algorithm. The impact of each block is quite independent, so that improvements can be devised separately. Typically, one can:

- upgrade the data-fidelity term by devising a more precise image-formation model and by gaining and taking advantage of a deeper knowledge of the statistics of the measurement noise;
- use prior-promoting regularizers that fit the object better;
- deploy robust and faster optimization schemes.

These three topics are shared by all inverse problems. Deconvolution microscopy can benefit from every improvements in this currently very active field of research.

The priors introduced by the regularizer must be chosen carefully to retain usefulness while avoiding the pitfall of overfitting. During the last decade, the compressive-sensing and the sparsity theories gave theoretical grounds to the observation that the  $\ell_1$ -based regularizers of (8) and (19) in Sections 3.7 and 3.9, respectively, always perform better than the  $\ell_2$ -based regularizers of Eqs. (6) and (8) in Sections 3.3 and 3.4, respectively.

Out of a dozen of competing methods, the methods that ranked first [34] and third [35] in the “2014 International Challenges on 3D Deconvolution Microscopy” took advantage of regularizers that were based on recent advances in signal processing<sup>8</sup>. The authors of

<sup>8</sup> <http://bigwww.epfl.ch/deconvolution/challenge/>.

[35] bring to fruition a second-order total-variation regularizer called a Schatten norm, while the method “Learn 2D Apply 3D” of [34] did exploit the fact that the resolution is much better within the lateral sections than along the axial direction. Assuming that the structures of interest are isotropic, it learned from the lateral sections of the acquired volume a dictionary of 2D high-resolution features that are used as priors to enhance the resolution along the axial direction. Approaches where the priors are learned appear to be very efficient; we surmise that the recent successes of deep neural networks in machine learning will soon lead to improved deconvolution algorithms [36] in microscopy. Finally, many modern deconvolution methods rely on state-of-the-art optimization schemes that can deal with non-differentiable  $\ell_1$  functions, for instance on proximal algorithms such as the alternating-direction method of multipliers [37].

Up to now in this paper, we have assumed that the PSF was known, either through ancillary measurements or through modeling. Moreover, we have assumed spatial shift-invariance of the systems. We now present approaches that have been recently developed to handle imaging situations in which these assumptions are not met.

### 6.1. Blind deconvolution

Blind deconvolution attempts to jointly estimate the object  $x$  and the PSF  $h$  from the data alone, without relying on ancillary measurements. It is a challenging, strongly ill-posed, and nonlinear problem. As an example, among other degeneracy issues [38], it must address that of scale, characterized by  $(\alpha h) * (\frac{1}{\alpha} x) = h * x$  for any non-vanishing  $\alpha$ . As it turns out, setting a meaningful value to  $\|h\|$  is highly nontrivial. Some proposed methods are explicitly designed to overcome degeneracies (scale and others) using an optically motivated parameterization of the PSF [39–41] or estimating the PSF from a dictionary [42]. Currently, the trend followed by all blind-deconvolution algorithms for fluorescence microscopy is to resort to iteratively alternating between the deconvolution and the estimation of the PSF [39–45].

### 6.2. Space-varying deconvolution

The deconvolution of large micrographs faces an important issue: in practice, the PSF varies across the field of view. In particular, a depth-varying PSF is often induced by a refractive index mismatch between the immersion medium and the specimen. In this case, the PSF suffers of spherical aberrations that get stronger as the focal plan is deeper inside the sample. This effect can be clearly seen on the Fig. 7(A) where the image is sharper at the bottom where the objective is closer to the sample.

Space-varying deconvolution raises two important problems. First, the assumption that the PSF varies across the field of view implies that the blurring process can no longer be modeled as a *convolution*. Hence, space-varying deconvolution is an oxymoron. As a consequence, FFT-based algorithm can no longer be used. The computational cost of space-varying deconvolution tends to rise as the square of the number of voxels. However using some approximations, several fast methods to model space-varying convolution have been proposed (see [46] for a review). In the refractive index mismatch case, as the size of the data along the depth axis is usually much smaller than along lateral axes, a depth only varying deblurring algorithm is much more tractable and several methods have been proposed in that case [47–50].

The second issue raised by the space varying deconvolution is how to estimate the PSF variation across the 3D object. With the exception of the case of refractive index mismatch where the PSF depth variation can be analytically known, one has to infer the PSFs

from the data in a space varying blind deconvolution algorithm. Up to now, only one attempt [41] has been done in that direction.

## Acknowledgement

The authors thank warmly Dr. Philippe Thévenaz for his useful feedback on the paper. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no 692726 “GlobalBioIm: Global integrative framework for computational bio-imaging”).

## Appendix A. Implementation aspects

### A.1. FFT Libraries

The algorithms that we have proposed made an extensive usage of the fast Fourier transform (FFT). For instance, one iteration of the Richardson–Lucy algorithm is composed of two multiplications in the Fourier domain (74 ms), a division in the space domain (51 ms), an application of the non-negativity constraint (6 ms), and two FFT/FFT<sup>-1</sup> (2’520 ms). The FFT/FFT<sup>-1</sup> are representing 95% of the computational time of this algorithm see (Table 4). `DeconvolutionLab2` has a Java wrapper for three FFT libraries.

- `AcademicFFT`<sup>9</sup>. This is pure Java library running on any platform. The source code of `AcademicFFT` is bundled with `DeconvolutionLab2`. It handles arbitrary data lengths, memory permitting. It is standalone; no external library is required.
- `JTransforms`<sup>10</sup>. This is the first, open-source, fast multithreaded FFT library written in pure Java. It is bundled with `Fiji` and `Icy`, but `JTransforms` is not part of the classical distribution of `ImageJ`.
- `FFTW 2.0`<sup>11</sup> [51]. `FFTW` is a C routine library for computing the fast Fourier transform in several dimensions, of arbitrary input size, and of both real and complex data. `FFTW` is one of the fastest FFT library. `DeconvolutionLab2` includes a wrapper for `FFTW` version 2.0 and it includes pre-compiled binaries for Mac OS X and Windows 32-bits and 64-bits.

### A.2. Dissection of an algorithm

We present a complete iterative algorithm from its mathematical formula to its Java snippet. Here, we choose to detail the Landweber algorithm Section 3.5.

We first reformulate the iteration to reduce the number of operations in the discrete Fourier domain and to limit the memory consumption.

#### A.2.1. Implementation of the Landweber algorithm

The original formulation is reduced to one multiplication and one addition in the discrete Fourier domain for every iterations.

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \gamma \mathbf{H}^T (\mathbf{y} - \mathbf{H} \mathbf{x}^k) \quad (21)$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \gamma \mathbf{H}^T \mathbf{H} \mathbf{x}^k + \gamma \mathbf{H}^T \mathbf{y} \quad (22)$$

$$\mathbf{x}^{k+1} = (\mathbf{I} - \gamma \mathbf{H}^T \mathbf{H}) \mathbf{x}^k + \gamma \mathbf{H}^T \mathbf{y} \quad (23)$$

$$\mathbf{x}^{k+1} = \mathbf{A} \mathbf{x}^k + \mathbf{g} \quad (24)$$

Using this expression, the variables  $\mathbf{A}$  and  $\mathbf{g}$  can be pre-computed.

<sup>9</sup> <http://bigwww.epfl.ch/thevenaz/academicfft/>.

<sup>10</sup> <https://sites.google.com/site/piotrwendykier/software/jtransforms>.

<sup>11</sup> <http://www.fftw.org/>.

**Table 4**

Computation time for a FFT and  $\text{FFT}^{-1}$  for a volume of size  $N \times N \times N$ . This experiment was performed on a Mac OS X 2.5 GHz Intel Core i7.

N (size)	FFT2 [ms]	JTransforms [ms]	AcademicFFT [ms]
32 × 32 × 32	1.5	9.8	11.5
37 × 37 × 37	13.3	30.3	17.2
56 × 56 × 56	9.6	12.8	34.8
64 × 64 × 64	17.1	23.5	38.6
74 × 74 × 74	101.2	61.8	111.0
111 × 111 × 111	353.9	189.1	324.0
128 × 128 × 128	247.4	151.9	577.9
147 × 147 × 147	347.4	243.3	620.8
223 × 223 × 223	7200.0	1615.4	4910.0
256 × 256 × 256	2937.7	1743.9	7860.0
294 × 294 × 294	3090.0	2197.7	11,200.0
446 × 446 × 446	62,200.0	46,700.0	61,100.0
512 × 512 × 512	35,000.0	25,900.0	141,000.0

$$\mathbf{A} = (\mathbf{I} - \gamma \mathbf{H}^T \mathbf{H}) \quad (25)$$

$$\mathbf{g} = \gamma \mathbf{H}^T \mathbf{y} \quad (26)$$

### A.2.2. Java snippet of Landweber

We choose the Java code of the Landweber algorithm. The iteration mechanism is handled by the object `controller` which is an instance of the class `Controller`. The instance of the Java FFT wrapper class is `fft` that contains two methods `transform()` and `inverse()`. The Java classes `ComplexSignal` and `RealSignal` are two classes of `DeconvolutionLab2` to store complex 3D signals and real 3D signals, respectively. The input variables are the two `RealSignal` objects, `input` and `psf` and the scalar parameter `gamma` which is the step parameter of the Landweber algorithm.

#### Landweber algorithm

```
// RealSignal y: this is the input volume to deconvolve
// RealSignal h: this is the PSF volume
// RealSignal x: this is the output deconvolved volume
// Operations.delta() is a high-level method to compute
// (I - gamma Ht H) public RealSignal call() {ComplexSignal Y
= fft.transform(y);
ComplexSignal H = fft.transform(h);
ComplexSignal A = Operations.delta(gamma, H);
ComplexSignal G = Operations.multiplyConjugate(gamma,
H, Y);
ComplexSignal X = G.duplicate();
while(!controller.ends(X)) {X.times(A);
X.plus(G);
constraint(X);}
RealSignal x = fft.inverse(X);
return x;}
```

## References

- [1] S. Inoué, Foundations of confocal scanned imaging in light microscopy, Handbook of biological confocal microscopy, Springer, 2006, pp. 1–19.
- [2] D.A. Agard, J.W. Sedat, Three-dimensional architecture of a polytene nucleus, Nature 302 (5910) (1983) 676–681.
- [3] J.G. McNally, T. Karpova, J. Cooper, J.A. Conchello, Three-dimensional imaging by deconvolution microscopy, Methods 19 (3) (1999) 373–385.
- [4] W. Wallace, L.H. Schaefer, J.R. Swedlow, A workingperson's guide to deconvolution in light microscopy, Biotechniques 31 (5) (2001) 1076–1097.
- [5] J. Sibarita, Deconvolution microscopy, Microsc. Tech. (2005) 1288–1291.
- [6] P. Sarder, A. Nehorai, Deconvolution methods for 3D fluorescence microscopy images, IEEE Signal Processing Mag. 23 (3) (2006) 32–45.
- [7] M. Bertero, P. Boccacci, G. Desiderà, G. Vicidomini, Image deblurring with poisson data: from cells to galaxies, Inverse Prob. 25 (12) (2009) 123006.
- [8] A. Griffa, N. Garin, D. Sage, Comparison of deconvolution software in 3D microscopy. a user point of view part 1, G.I.T., Imag. Microsc. 1 (2010) 43–45.
- [9] A. Griffa, N. Garin, D. Sage, Comparison of deconvolution software in 3D microscopy. a user point of view part 2, G.I.T., Imag. Microsc. 1 (2010) 41–43.
- [10] A. Ponti, P. Schwarb, A. Gulati, V. Bäcker, Huygens remote manager, Imag. Microsc. 9 (2) (2007) 57–58.
- [11] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci. 2 (1) (2009) 183–202.
- [12] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, P. Roux, J. Olivo-Marin, J. Zerubia, Richardson–Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution, Microsc. Res. Tech. 69 (2006) 260–266.
- [13] C. Vonesch, M. Unser, A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution, IEEE Trans. Image Process. 17 (4) (2008) 539–549.
- [14] V. Směkalová, I. Luptovčiak, G. Komis, O. Šamajová, M. Ovečka, A. Doskočilová, T. Takáč, P. Vadovič, O. Novák, T. Pechan, et al., Involvement of YODA and mitogen activated protein kinase 6 in arabidopsis post-embryonic root development through auxin up-regulation and cell division plane orientation, New Phytologist 203 (4) (2014) 1175–1193.
- [15] R. Chen, J. Zhang, Y. Wu, D. Wang, G. Feng, Y.-P. Tang, Z. Teng, C. Chen, Monoacylglycerol lipase is a therapeutic target for alzheimer's disease, Cell Rep. 2 (5) (2012) 1329–1339.
- [16] T. Deguchi, M.H. Alanne, E. Fazeli, K.M. Fagerlund, P. Pennanen, P. Lehenkari, P. E. Hänninen, J. Peltonen, T. Närejoja, In vitro model of bone to facilitate measurement of adhesion forces and super-resolution imaging of osteoclasts, Sci. Rep. 6 (2016) 22585.
- [17] M.M. Dobbin, R. Madabhushi, L. Pan, Y. Chen, D. Kim, J. Gao, B. Ahanonu, P.-C. Pao, Y. Qiu, Y. Zhao, et al., SIRT1 collaborates with ATM and HDAC1 to maintain genomic stability in neurons, Nat. Neurosci. 16 (8) (2013) 1008–1015.
- [18] M. Kicia, N. Janeczko, J. Lewicka, A.B. Hendrich, Comparison of the effects of subinhibitory concentrations of ciprofloxacin and colistin on the morphology of cardiolipin domains in escherichia coli membranes, J. Med. Microbiol. 61 (4) (2012) 520–524.
- [19] M. Padilla, B. Michl, B. Thaidigsmann, W. Warta, M. Schubert, Short-circuit current density mapping for solar cells, Solar Energy Mater. Solar Cells 120 (2014) 282–288.
- [20] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nat. Methods 9 (7) (2012) 671–675.
- [21] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., Fiji: an open-source platform for biological-image analysis, Nat. Methods 9 (7) (2012) 676–682.
- [22] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, et al., Icy: an open bioimage informatics platform for extended reproducible research, Nat. Methods 9 (7) (2012) 690–696.
- [23] J. Starck, E. Pantin, F. Murtagh, Deconvolution in astronomy: A review, Publ. Astron. Soc. Pac. 114 (800) (2002) 1051.
- [24] M. Born, E. Wolf, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, Pergamon Press, 1956.
- [25] H. Kirshner, F. Aguet, D. Sage, M. Unser, 3-D PSF fitting for fluorescence microscopy: implementation and localization application, J. Microsc. 249 (1) (2013) 13–25.
- [26] A. Tikhonov, Solution of incorrectly formulated problems and the regularization method, Soviet Mathematics Dokl., vol. 5, 1963, pp. 1035–1038.
- [27] N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, vol. 2, MIT press Cambridge, 1949.
- [28] L. Landweber, An iteration formula for fredholm integral equations of the first kind, Am. J. Math. 73 (3) (1951) 615–624.
- [29] W.H. Richardson, Bayesian-based iterative method of image restoration, J. Optical Soc. Am. 62 (1972) 55–59.
- [30] L.B. Lucy, An iterative technique for the rectification of observed distributions, Astrophys. J. 79 (6) (1974) 745–754.
- [31] J.R. Swedlow, M. Platani, Live cell imaging using wide-field microscopy and deconvolution, Cell Struct. Funct. 27 (5) (2002) 335–341.
- [32] J.R. Swedlow, Quantitative fluorescence microscopy and image deconvolution, Methods Cell Biol. 81 (2007) 447–465.
- [33] J.-S. Lee, T. Wee, C.M. Brown, Calibration of wide-field deconvolution microscopy for quantitative fluorescence imaging, J. Biomol. Technol. 25 (2014) 31–40.
- [34] F. Soulez, A learn 2D, apply 3D method for 3D deconvolution microscopy, IEEE 11th International Symposium on Biomedical Imaging (ISBI), IEEE, 2014, pp. 1075–1078.
- [35] S. Lefkimmatis, M. Unser, Poisson image reconstruction with hessian Schatten-norm regularization, IEEE Trans. Image Process. 22 (11) (2013) 4314–4327.
- [36] L. Xu, J.S. Ren, C. Liu, J. Jia, Deep convolutional neural network for image deconvolution, Advances in Neural Information Processing Systems, 2014, pp. 1790–1798.
- [37] N. Parikh, S.P. Boyd, Proximal algorithms, Found. Trends Optim. 1 (3) (2014) 127–239.
- [38] F. Soulez, M. Unser, Superresolution with optically motivated blind deconvolution, in: Computational Optical Sensing and Imaging, Optical Soc. Am., 2016, pp. JT3A–38.
- [39] J. Markham, J.-A. Conchello, Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur, J. Opt. Soc. Am. A 16 (10) (1999) 2377–2391.

- [40] F. Soulez, L. Denis, Y. Tourneur, E. Thiébaud, Blind deconvolution of 3D data in wide field fluorescence microscopy, 9th IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2012, pp. 1735–1738.
- [41] B. Kim, T. Naemura, Blind depth-variant deconvolution of 3D data in wide-field fluorescence microscopy, *Sci. Rep.* 5 (2015). 9894 EP.
- [42] T. Kenig, Z. Kam, A. Feuer, Blind image deconvolution using machine learning for three-dimensional microscopy, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (12) (2010) 2191–2204.
- [43] L.M. Mugnier, T. Fusco, J.-M. Conan, MISTRAL: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images, *J. Opt. Soc. Am. A* 21 (10) (2004) 1841–1854.
- [44] E.F.Y. Hom, F. Marchis, T.K. Lee, S. Haase, D.A. Agard, J.W. Sedat, AIDA: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data, *J. Opt. Soc. Am. A* 24 (6) (2007) 1580–1600.
- [45] M. Keuper, T. Schmidt, M. Temerinac-Ott, J. Padeken, P. Heun, O. Ronneberger, T. Brox, Blind deconvolution of widefield fluorescence microscopic data by regularization of the optical transfer function (otf), *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 2179–2186.
- [46] L. Denis, E. Thiébaud, F. Soulez, J.-M. Becker, R. Mourya, Fast approximations of shift-variant blur, *Int. J. Comput. Vision* 115 (3) (2015) 253–278.
- [47] M. Arigovindan, J. Shaevitz, J. McGowan, J.W. Sedat, D.A. Agard, A parallel product convolution approach for representing the depth varying point spread functions in 3D widefield microscopy based on principal component analysis, *Opt. Express* 18 (7) (2010) 6461–6476.
- [48] N. Patwary, C. Preza, Image restoration for three-dimensional fluorescence microscopy using an orthonormal basis for efficient representation of depth-variant point-spread functions, *Biomed. Opt. Express* 6 (10) (2015) 3826–3841.
- [49] E. Maalouf, B. Colicchio, A. Dieterlen, Fluorescence microscopy three-dimensional depth variant point spread function interpolation using zernike moments, *J. Opt. Soc. Am. A* 28 (9) (2011) 1864–1870.
- [50] S. Ben Hadj, L. Blanc-Féraud, G. Aubert, Space variant blind image restoration, *SIAM J. Imag. Sci.* 7 (4) (2014) 2196–2225.
- [51] M. Frigo, S.G. Johnson, The design and implementation of FFTW3, *Proc. IEEE* 93 (2) (2005) 216–231. special issue on Program Generation, Optimization, and Platform Adaptation.



# DiAna, an ImageJ tool for object-based 3D co-localization and distance analysis



Jean-François Gilles<sup>a</sup>, Marc Dos Santos<sup>b</sup>, Thomas Boudier<sup>c,d</sup>, Susanne Bolte<sup>a,1</sup>, Nicolas Heck<sup>b,\*,1</sup>

<sup>a</sup> Sorbonne Universités, UPMC Univ Paris 06, CNRS, Core Facility – Institut de Biologie Paris Seine (IBPS), 75005 Paris, France

<sup>b</sup> Sorbonne Universités, UPMC Univ Paris 06, INSERM, CNRS, Neurosciences Paris Seine – Institut de Biologie Paris Seine (NPS – IBPS), 75005 Paris, France

<sup>c</sup> Sorbonne Universités, UPMC Univ Paris 06, IPAL, CNRS, Singapore

<sup>d</sup> Bioinformatics Institute, Agency for Science, Technology and Research (A\*STAR), Singapore

## ARTICLE INFO

### Article history:

Received 4 August 2016

Received in revised form 22 November 2016

Accepted 22 November 2016

Available online 24 November 2016

### Keywords:

Image segmentation

Image co-localization

Fluorescence microscopy

Deconvolution

Spatial analysis

3D imaging

## ABSTRACT

We present a new plugin for ImageJ called DiAna, for Distance Analysis, which comes with a user-friendly interface. DiAna proposes robust and accurate 3D segmentation for object extraction. The plugin performs automated object-based co-localization and distance analysis. DiAna offers an in-depth analysis of co-localization between objects and retrieves 3D measurements including co-localizing volumes and surfaces of contact. It also computes the distribution of distances between objects in 3D. With DiAna, we furthermore introduce an original method, which allows for estimating the statistical significance of object co-localization. DiAna offers a complete and intuitive 3D image analysis tool for biologists.

© 2016 Elsevier Inc. All rights reserved.

## 1. Introduction

Many biological and physiological studies depend on the analysis of the distribution and spatial relationship between biomarkers in a cell or tissue. Various methods, from immunodetection to transgene-driven expression of fluorescent proteins, allow to observe positive cells in whole specimen or tissue sections at the histological level, or to detect protein sub-cellular localization at the cellular level. The 3D-organization and relationship of these biomarkers can be investigated using fluorescence microscopy techniques that allow optical sectioning, such as confocal microscopy or multiphoton microscopy. These should be combined with appropriate image analysis methods. Biologists often investigate the spatial overlap of pairs of biomolecules in a cell or in a sub-cellular compartment by means of co-localization analysis, before inferring biological interaction and drawing functional conclusions.

Co-localization analysis can be carried out using two different, complementary methods: a pixel based approach [1–2] or an object-based approach [3–5], as well as a combination of the two

approaches [6]. These methods have been reviewed in detail [5,7–8].

In the pixel-based approach, the linear relationship between a pair of biomolecules is calculated statistically between two fluorescent channels [2,9] without taking into account the positional information. Statistical significance of this correlation coefficient may be ensured by computing co-localization after scrambling pixel coordinates [10] or by shifting images pixel wise [11]. Pixel-based approaches are included in most commercially available image analysis softwares because they are easy to implement. However, co-localization analysis by pixel-based approaches is affected by the inherent noise of fluorescent images and thus not always applicable [5,8]. Furthermore, pixel-based approaches do not give information about the spatial relationship between objects.

In the object-based approach, spatial information is used to quantify the degree of co-localization between objects in the image. It is thus indispensable to perform segmentation prior to co-localization analysis in order to identify and delineate the objects of interest. Intensity thresholding is a simple method, but its easiness comes with limitations as the threshold is globally applied to the whole image. Recently, more sophisticated 3D, local segmentation methods have been developed [12,13], including spot segmentation and iterative thresholding, which we describe in this study.

\* Corresponding author at: Neuroscience Paris Seine, 7 Quai Saint-Bernard, Building A 3rd Floor, 75005 Paris, France.

E-mail address: [nicolas.heck@upmc.fr](mailto:nicolas.heck@upmc.fr) (N. Heck).

<sup>1</sup> Those authors equally contributed to his work.

The segmentation allows determining the volume that an object occupies in the 3D space as well as the localisation of the geometrical centroid or centre of mass of the object. The degree of co-localization can then be calculated in different ways, depending on the optical resolution limit and the size of objects investigated. In one approach co-localization is detected if the distance between the centres of the objects in two fluorescent channels is lower than image resolution. In another approach, co-localization can be inferred if the centre of one object falls into the volume occupied by the other object. These two paradigms are used in the JACOP plugin from ImageJ [5]. More recently, a plugin for object-based co-localization analysis has been implemented in the software Icy [8,14]. In this approach, a spherical region is computed around the object centre, and co-localization is deduced if the centre of another object falls into this spherical region. Several drawbacks are inherent to the aforementioned approaches. Indeed, the object position determined by its centroid or centre of mass is not sufficient to fully represent the object. Furthermore, no information, neither on the extent of co-localization for each object, nor on the property of the co-localizing objects (such as intensity or volume) can be obtained.

Here, we first summarize the critical steps for optimized image acquisition and present a new plugin, DiAna -for Distance Analysis. This plugin allows to segment the objects as well as to perform in-depth analysis of co-localization and distance between objects. The applicability of the various tools provided by DiAna is illustrated with images of neuronal synaptic markers, since they are objects with known co-localization patterns. The analysis of those pre- and post-synaptic elements that form a synapse is representative of the sub-cellular co-localization analysis which can be performed with DiAna and can be transferred to any kind of biological structures.

## 2. Material and methods

### 2.1. Animal care

Animal care was conducted in accordance with standard ethical guidelines (NIH publication No. 85-23, revised 1985 and European Committee Guidelines on the Care and Use of Laboratory Animals 86/609/EEC), and the experiments were approved by the local ethic committee. Male mice weighing 22–24 gm were housed 5 per cage and acclimatized to laboratory conditions (12 h light/dark cycle,  $21 \pm 1$  °C room temperature) with *ad libitum* access to food and water. VGLUT1-venus knock-in mice express the Vesicular Glutamate Transporter 1 (VGLUT1) fused to the fluorescent protein Venus under VGLUT1 endogenous promoter [15].

### 2.2. Sample preparation

Mice brains were fixed by intracardiac perfusion of 4% paraformaldehyde in 0.1 M  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  (phosphate buffer, PBS), pH 7.4. Brains were dissected and post-fixed overnight at 4 °C. Coronal sections of 50  $\mu\text{m}$  thickness were cut with a vibratome (Leica).

### 2.3. Immunofluorescence

Sections were permeabilized for 30 min in PBS containing 0.1% Triton X-100 and 3% bovine serum albumin (BSA). The sections were incubated with the primary antibody in PBS with 3% BSA overnight at 4 °C. The antibodies used were mouse monoclonal directed against synaptophysin (1/1000, Sigma), mouse monoclonal directed against bassoon (1/1000, Stressgene), rabbit polyclonal directed against tyrosine hydroxylase (1/1000, Sigma).

Following three washing steps of 15 min with PBS, secondary antibody conjugated with the fluorochrome alexa-561 or alexa-488 directed against either mouse or rabbit (1/500, Invitrogen) was incubated in PBS with 3% BSA for 2 h at room temperature. After washing, sections were rinsed before mounting in Prolong Gold.

### 2.4. Confocal image acquisition and deconvolution

Images stacks were taken with a Confocal Laser Scanning Microscope (TCS SP5, Leica Microsystems, Germany) equipped with a 1.4 NA objective (oil immersion, Leica) with pinhole aperture set to 1 Airy Unit, pixel size of 60 nm and z-step of 200 nm. Excitation wavelength was 488, 514 or 561, and emission range was 500–550, 525–540 or 570–620 nm, for detection of alexa-488, venus or alexa-561, respectively. Laser intensity and photomultiplier tube gain were set so the image occupies the full dynamic range of the detector. Deconvolution was performed using an experimental Point Spread Function obtained from fluorescent beads and Maximum Likelihood Estimation algorithm (Huygens software, Scientific Volume Imaging, Netherlands). 150 iterations were applied in classical mode, background intensity was averaged from the voxels with lowest intensity, and signal to noise ratio values were set to a value of 15.

### 2.5. Segmentation

Three segmentation procedures are implemented in the plugin. The first one is global intensity thresholding of the image. The second one is based on spot segmentation [16]. Local maxima are computed in the image, and a user-defined threshold allows selecting the local maxima belonging to objects. Then the 3D radial distribution of the voxel intensities around each local maximum is computed and a threshold is estimated for the border of each object. When the border intensity threshold has been found, the voxels around the local maximum are examined and successively included in the segmented object through the following algorithm: each 3D neighbour voxel is examined and included in the object if 1: their intensity is above the threshold 2: if their intensity is lower than the voxel previously added to the object 3: if the other neighbours would be added to the object as well. The third segmentation procedure is based on an iterative thresholding process, it is a simplified version of the algorithm published by Gul-Mohammed et al. [17], and is based on the idea of max-trees and MSER technique [18,19]. An interval of volumes is fixed and the image will undergo thresholding at each possible threshold and segmented objects having a volume in the defined range will be extracted from the image. The extracted objects are then organized into a hierarchy. Since a same object can be extracted with different thresholds, the corresponding extracted objects will be stored in different branches of the hierarchy. In case an object will split into two objects at higher threshold, a branch division will be created. Then on all final branches, corresponding to higher thresholds, the thresholds yielding to most stable objects, in term of volume, will be computed and the corresponding object will be displayed as the best object on this branch, all other instances of this object will then be discarded.

Unless stated differently, segmentation was performed in 3D using the spot segmentation procedure using the following parameters: Maxima detection: radius in xy-axis = 4, in z-axis = 3, noise parameter set to zero; Threshold for maxima selection was set to 5000; Parameters for Gaussian fit and threshold calculation were Radius maximum = 10, S.D. value = 1.5.

The validation of the segmentation was estimated by calculation of F-measure. The accuracy was classically measured by  $F = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$ . Precision and Recall

are estimates of false positives and false negatives rate, respectively.

### 2.6. Distance analysis

Distance analyses are based on classical euclidean distance computation. We implemented centre-to-centre distances, centre-to-edge distances and edge-to-edge distances. In order to be computationally efficient the objects are defined as the list of the voxels comprising the object. For edge-to-edge distance analysis, the list of contour voxels are extracted and put into a KD-tree for efficient distance computation [20]. The edge-to-edge distance is hence the smallest distance between the two objects, and will be equal to 0 if the two objects intersect. The closest objects are also computed using a KD-tree approach for efficiency.

### 2.7. Co-localization and contact surface analysis

The computation of co-localized voxels is based on the analysis of the corresponding labeled images of the objects [21]. The object 1 is labeled with values 1 in the first image and the object 2 with values 2 in the second image, the two images are then summed up. The number of voxels having a value 3 will hence correspond to the number of co-localized voxels between the two objects. From two non-co-localized objects we implemented a contact surface computation. We define a minimum distance between the borders of the two objects and compute the number of border voxels from one object having border voxels from the other object below the defined distance.

### 2.8. Statistical computation

In order to compute robust co-localization analysis, we need to assess the statistical significance of the co-localization. Based on the idea of the randomization of pixels described by Costes, we carry out a randomization of object's positions. A new position is randomly assigned to the objects, while ensuring the objects remain in the surrounding structure and do not intersect with other objects. Based on this randomization we used the framework defined by Andrey et al. [22] to define an index describing the statistical significance of the co-localization. We first compute the cumulated distribution function (cdf) of all distances between the centres of objects of the first channel to the centre of the closest object in the second channel for the observed data. We then compute the same cdf for  $n$  randomized data and rank the observed data among the  $n$  randomized data. For a 5% interval, if the observed data falls into the first or last 2.5% of the randomized

data, we can then reject the hypothesis than the co-localization is only due to chance, as this may happen in high density objects populations.

## 3. Results and discussion

### 3.1. Segmentation: determination of the objects in the image

Segmentation is a process that allows the identification of objects in digital images. Binarization assigns a value of either 0 or 1 to each pixel. Pixels belonging to objects have a value of 1, background pixels have a value of 0. In a second step, the objects are labeled, which means that all neighbouring pixels with a value of 1 are grouped to define an object. Segmentation is thus a process by which a gray-scale image becomes a space in which objects are located. In the most classical segmentation procedure, an intensity threshold is used to binarize the image. Before thresholding, smoothing of the image with an appropriate filter can be performed with the plugin DiAna, if the image has previously not been deconvolved and noise filtered. Indeed, proper image acquisition and pre-processing facilitates subsequent object extraction by image segmentation for reliable analysis. The preparation of the sample and image acquisition have been extensively discussed elsewhere [5,23–25], and the benefits of deconvolution have been well demonstrated [23,26–28] Noise filtering and global thresholding are implemented in DiAna, so objects can be extracted before analysis. The major problem with intensity thresholding lies in the fact that this is a global approach. If the image contains bright and faint objects, the application of a single threshold to the image will end with the bright objects being too big or the faint objects being too small. Hence, a local approach for segmentation is often more useful. Therefore, we included two 3D segmentation procedures in the DiAna plugin, which allow for segmentation of image containing objects with different sizes and intensities (Fig. 1). The “spot segmentation” procedure is based on detection of objects with local maxima and 3D analysis of the intensity distribution around the maxima [16] (Fig. 1A,B). This procedure brings the advantage that objects with different intensities will be properly segmented. It should, however, be noted that each local maximum is a seed that marks an object, it is thus advised to deconvolve or apply smoothing filters before maxima detection, and to carefully set the intensity threshold for maxima selection. The plugin DiAna offers visualization the local maxima so best parameters can be found before performing segmentation. Finally, it is noteworthy that the method is best suited for spot-like objects since in case of irregular objects several local maxima may be found within the same object which would then be split. Therefore, we intro-

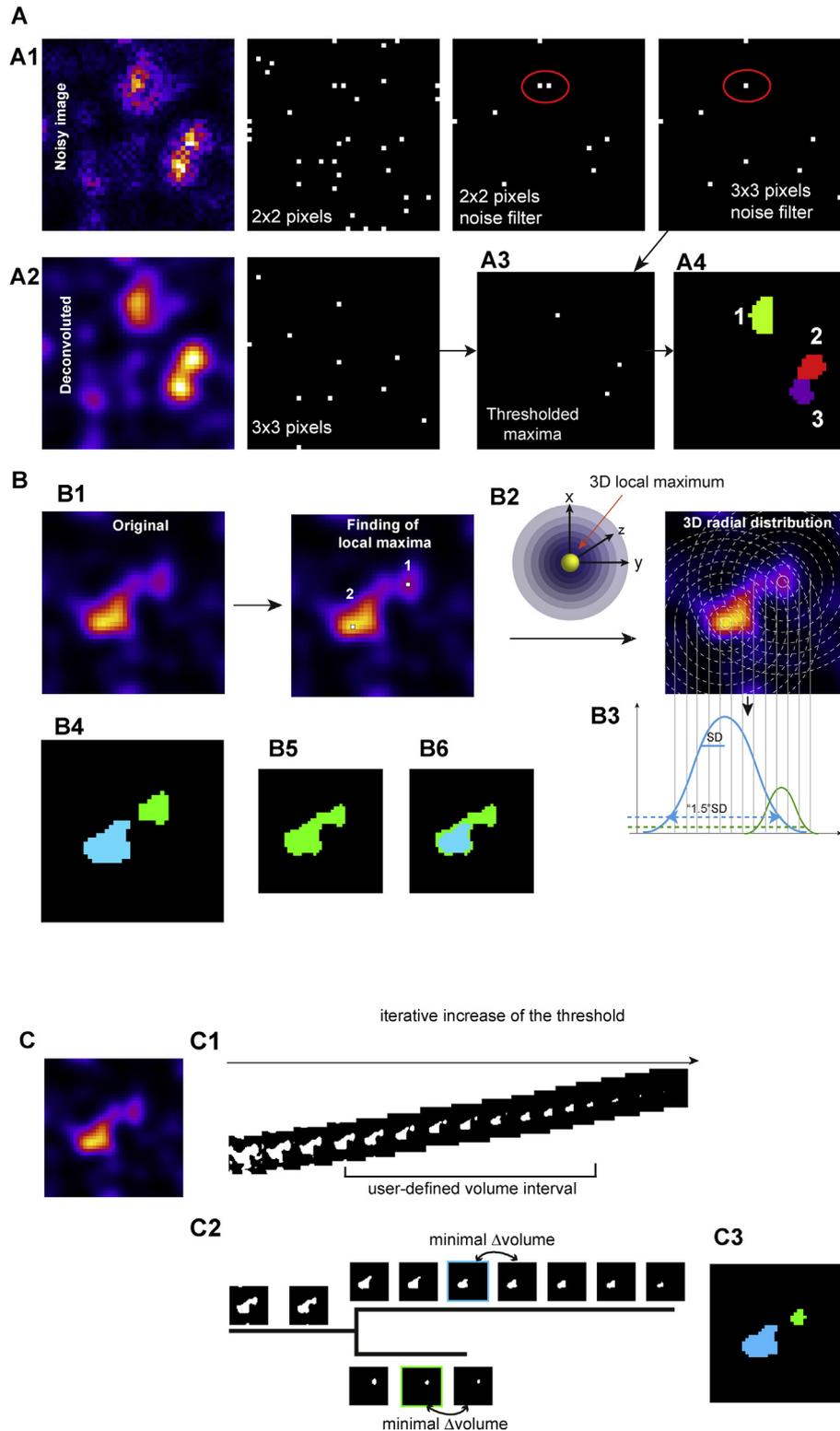
---

**Fig. 1.** Principles of the segmentation procedures performed by the plugin DiAna. A. Detection of local maxima for segmentation of the objects with the spot segmentation procedure. A1. In noisy images, the “noise” parameter allows the local maxima outside the objects to be discarded. For the radius parameter, a low radius ( $2 \times 2$ ) retrieves several maxima within the same object. A bigger radius ( $3 \times 3$ ) ensures minimal distance between local maxima and avoids this problem. An example is encircled in red. A2. The maxima detection is more robust in deconvolved or properly noise-filtered images. A3. Prior segmentation, a user-defined threshold intensity value allows selecting the maxima belonging to the objects. A4. Result of the spot segmentation further explained in B. B. Workflow of the spot segmentation procedure. B1. Maxima are detected as shown in A. B2. 3D radial distribution of the intensity centred to the maxima is computed. The user should define a maximum value in voxel for the radius of the largest sphere. It is advised that it is bigger than the largest object expected in the image. B3. The distribution is plotted and fitted to a Gaussian curve. The user defines a factor applied to the standard deviation of the Gaussian curve, which sets the size of a horizontal line placed to cover the area defined by the Gaussian curve (i.e. the size of the object). Factors of 1.5 and 2 will cover 86.6 and 95.4% of the Gaussian curve, respectively. The corresponding intensity value on y-axis of the graphic is the threshold used for segmentation of the object. B4. The procedure allows finding a specific threshold for objects of different intensities. Voxels around the maxima are successively included in the segmented object through an algorithm with three criteria of acceptance: The first criterion is that the voxel intensity should be higher than the defined threshold. The second criterion is that their intensity should be lower than the intensity of the voxels previously included in the objects, which avoids merging of adjacent objects as shown in B5. The third criterion is that the voxel is included on the condition that neighbouring voxels are included as well, which avoids creation of filamentous structures extruding from the object as shown in B6. C. Workflow of the iterative segmentation method. C1. The image is segmented with different intensity thresholds. From low to high intensity thresholds, the objects become isolated and decrease in size. The segmented objects which fall in the volume interval defined by the user are stored in a hierarchy. C2. The selected objects are classified in a hierarchy showing the object separation and decrease in size along increasing threshold values. When an object becomes isolated it is stored in a new branch in the hierarchy. The final branches contain most disconnected objects, still having their volume in the specified interval. The best thresholds within the last branches are found with the maximum stable volume algorithm. The difference in the object volume between each threshold is computed and the threshold corresponding to the minimum difference is retained. C3. The iterative method determines a threshold specific to each object and reconstructs the segmented image.

duced a tool in the ROI manager for manually selecting and merging pairs of objects when necessary. The “iterative segmentation” procedure extracts each object by examining the output of thresholding at different intensity values [17] (Fig. 1C). A complete version of these segmentation procedures with more options can be found in the 3D Image Suite [29,30].

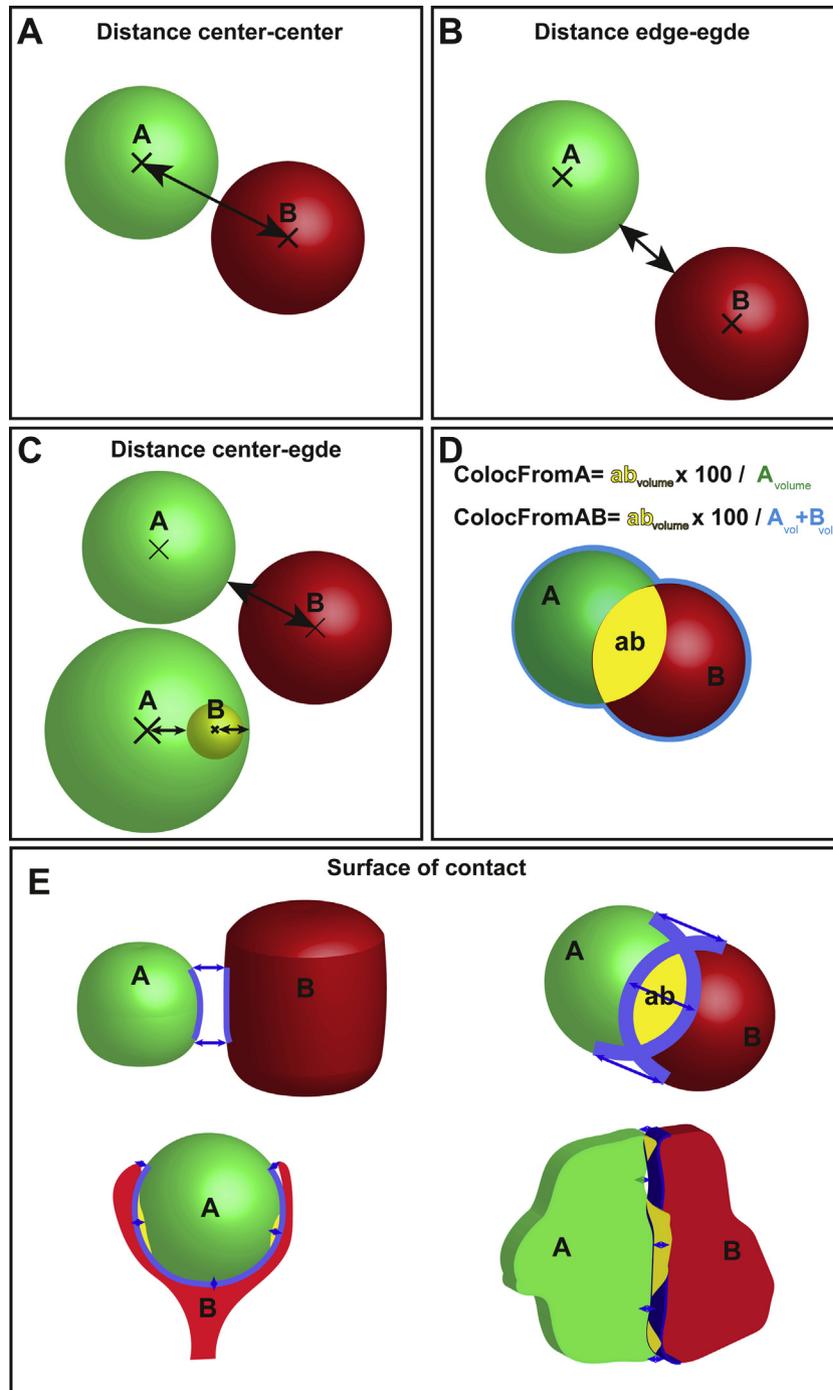
The segmentation procedures implemented in DiAna were validated for an image of neuronal presynaptic elements by

immunofluorescence labelling of tyrosine hydroxylase, acquired with confocal microscopy (Supplementary Fig. S1). In comparison to the ground truth (positive objects eye counted by experimenter), both segmentation procedures extracted objects accurately (F-measure of 0.992 and 0.987 for spot and iterative procedure, respectively). The output of segmentation was further tested on another dataset from similar objects to compare results from raw and deconvolved image. The segmentation was less pre-



cise on raw images but still efficient enough to apply co-localization analysis (F-measure for deconvolved images 0.966 and 1, and for raw images 0.934 and 0.966, for spot and iterative procedure, respectively). The algorithms in both spot and iterative segmentation procedures have been chosen so they are generally applicable to most biological images with good robustness. However, users are advised to test several values for the parameters and validate the segmentation by visual inspection of the segmentation result. It is a general rule that only the biologist can properly estimate whether the segmented objects correspond to what he/

she considers as biological objects. Images of similar biological objects often contain variable intensities. This may be due to differences in the efficacy of fluorescent labelling from one experiment to another, or from lack of stability of the excitation and detection system of the microscope. Therefore, the values determined for the parameters of the segmentation may not be applicable to two different sets of images. A way to circumvent this problem is to normalize the mean intensity of the whole image [23] before defining the segmentation parameters and reliably applying them to all images of the study.

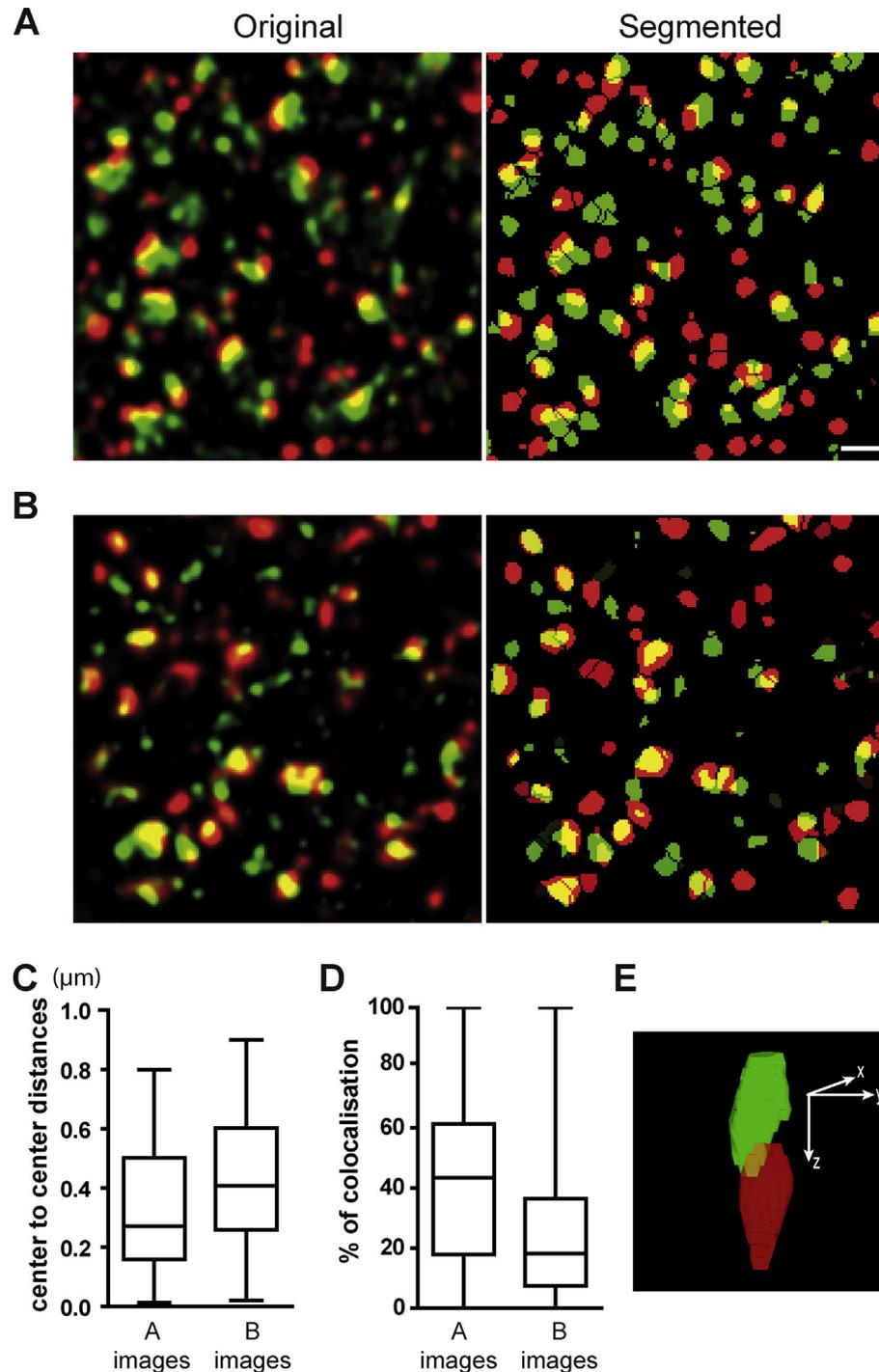


**Fig. 2.** Measurements performed in DiAna plugin. A–C. Distance analysis measurements performed by DiAna are centre-to-centre (A), edge-to-edge (B), centre-to-edge (C). D. Percentage of co-localizing volumes normalized to the volume of either one or both objects are computed by DiAna. E. Contact surface measurements for overlapping and distant objects. The user-defined maximum distance between objects (blue arrows) set the extent of the object contour for which the contact is defined. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

### 3.2. Three-dimensional measurements for object analysis

Following the application of the segmentation procedure described in Section 3.1, DiAna offers precise 3D measurements for co-localization and distance analysis. Of note, users can segment their images with any other method of their choice and

directly use the analysis tools of DiAna. The plugin quantifies several parameters for pairs of objects (Fig. 2). It also introduces measurement of surface of contact for distant and co-localizing objects, which requires a user-defined edge-to-edge distance (Fig. 2E). Those parameters are used for co-localization and distance analysis as described in Sections 3.3 and 3.4. Moreover, DiAna performs



**Fig. 3.** Detailed co-localization measurements using the plugin DiAna A–B. Close-ups of single sections from deconvolved and segmented image stacks showing two fluorescent channels. Images were obtained from brain sections immunolabeled for synaptic proteins (green: vesicular transporter VGLUT1, red in A: bassoon, red in B: synaptophysin). C. Distance measurements between centres of co-localizing objects reveal that the co-localized objects from image B are closer to each other than the co-localized objects from image A. D. Measurements of the percentage of the co-localizing volume for each object's pair reveal that the co-localizing volume is bigger for the objects from image B than for the objects from image A. E. 3D volume rendering from image stack showing an example of co-localization artefact due to optical smear. The percentage co-localization volume is lower than 5% and a cut-off can be defined to avoid such false positives. Images were segmented using spot segmentation protocol. Co-localization analysis was performed on image stacks containing more than 5000 objects in each channel. Scale bars in A–B: 1  $\mu\text{m}$ . Note that some spots are excluded from segmentation as they are not biologically relevant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measurements of several parameters such as volume, mean intensities, surface area, Feret's diameter, coordinates of centres of mass and centroids for all objects of the image.

### 3.3. Co-localization analysis with DiAna

Co-localization is determined by the detection of overlapping objects. DiAna computes distances between co-localizing objects (Fig. 2A–C) as well as measurements of co-localizing object volumes for each pair of objects (Fig. 2D). Those measurements allow for precise quantification of the co-localization. Fig. 3 shows two biological examples where images of synaptic markers with known differences in co-localization were analysed with DiAna. Note that some spots are excluded from segmentation because they are not biologically relevant. VGLUT1 can sparsely be found in the axon in between the presynaptic elements which are the objects to be extracted [31] and immunofluorescence for synaptophysin and bassoon retrieves low intensity background besides relevant signal. The co-localization of the vesicular transporter VGLUT1 with synaptophysin (Fig. 3A) and with Bassoon (Fig. 3B) was analysed by determining centre-to-centre distances and the percentage of the co-localization volume (Fig. 3C,D). A perfect co-localization between VGLUT1 and synaptophysin was revealed with a centre-to-centre distance smaller than optical resolution, while the co-localization between VGLUT1 and bassoon is not complete, as shown by a longer centre-to-centre distance and lower percentage of co-localization for each pair of objects. Those results fit the measurements performed from electron microscopy

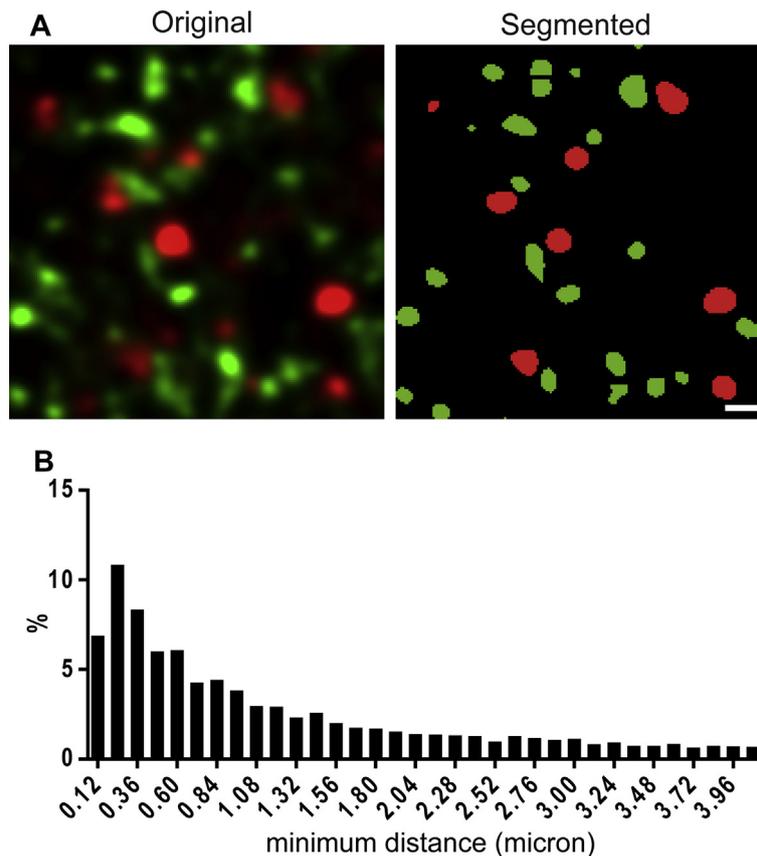
and STED microscopy data [32,33], as VGLUT1 and synaptophysin are found in the main vesicular pool while bassoon is restricted to the active zone next to the plasma membrane of the presynapse [34,35].

Interestingly, the percentage of the co-localizing object's volume given by DiAna allows using a cut-off to discard false positives, which could appear because the blur in z-axis can lead to small overlap between objects (Fig. 3E). This loss of axial resolution can be reduced by refractive index matching [36] and by deconvolution [27,28]. It is noteworthy that it was measured that deconvolution improves co-localization analysis [37–39].

Finally, in addition to co-localization measurements, DiAna retrieves tables with quantifications such as volume and mean intensity for each object of both images. Each object is identified in the co-localization and measurement's tables, ensuring that a correlation between chosen criteria can be drawn.

### 3.4. Distance analysis with DiAna

DiAna also offers a distance analysis for either co-localized objects or for all objects from the two images. For each object from one image, the centre to centre distances with all objects of the other image are computed in 3D. By default, the plugin identifies the closest, not co-localizing, object. The user can also select the rank of the objects to be included in the result table (first closest only, first and second closest, first to fifth etc.). Note that in this case the co-localizing objects will be included in the results as the first closest objects. The distances measured can be used to plot



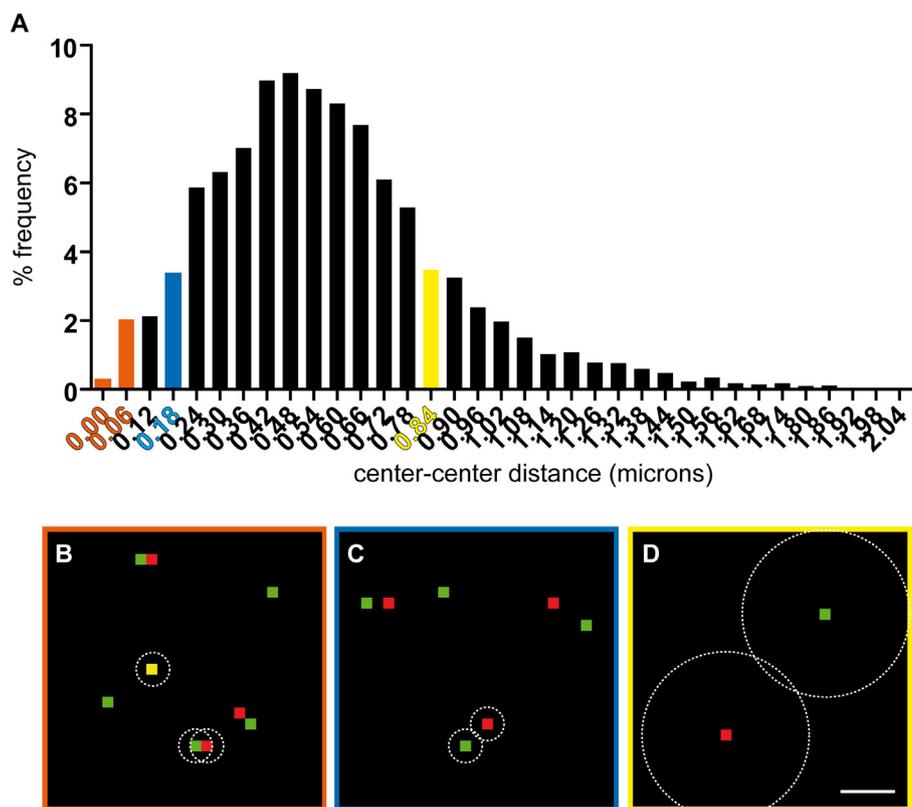
**Fig. 4.** Distance analysis of paired of non co-localizing objects with the plugin DiAna A. Close-up of a single slice from deconvolved and segmented image stack showing both channels. Images were obtained from brain sections immunolabeled for proteins expressed in different, hence not co-localizing, synaptic structures (green: vesicular transporter VGLUT1, red: tyrosine hydroxylase). B. Distance analysis retrieves centre-to-centre distance in between all objects from green channel and their closest neighbour in red channel. The histogram of distance distribution shows that 50% of the objects from the red channel are localized at less than 1  $\mu\text{m}$  from objects of green channel. Distance analysis was performed on image stacks containing more than 3000 objects in each channel. Scale bars in A–B: 1  $\mu\text{m}$ . Note that some spots are excluded from segmentation as they are not biologically relevant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a histogram of distances between neighbouring objects for spatial distribution analysis, and correlation between quantified parameters from the object pair can be assessed. Fig. 4 shows the spatial analysis in a biological example where the distances of two synaptic markers, the vesicular transporter VGLUT1 and tyrosine hydroxylase where analysed with DiAna. The analysis of the distances to the first closest object shows that objects from the red channel are preferentially localized in the vicinity of objects from the green channel, since 50% of all red objects are found in a distance below 1  $\mu\text{m}$  to a green object, which exactly fits with measurements performed on images obtained by electron microscopy [40].

The distance analysis can furthermore be used to assess co-localization in the case where objects are represented as single voxels. DiAna can be used to estimate co-localization from images of the centre of the objects. The centre-to-centre distance, given by the distance analysis, allows determination which objects do co-localize (Fig. 5). The measured distances can be ranked to set a threshold value, which will define which single-voxel objects do co-localize. Co-localization can be decided if the distance is smaller than the optical resolution. In this case the threshold distance can be either set to zero or more, depending on the resolution of the image (Fig. 5A–C). Alternatively, the co-localization can be decided if the distance value is lower than a threshold distance estimated on the basis of the knowledge of minimal size of the objects (Fig. 5A,D). It is noteworthy that this method can be applied for co-localization analysis in images of single particles, in which objects are represented as single voxels, often identified as the peak of the Gaussian distribution of objects intensity.

### 3.5. Assessing statistical robustness in object based co-localization analysis

In pixel-based co-localization analysis, statistical significance of co-localization may be estimated by comparing co-localization measurements before and after randomization of pixels. Costes et al. [10] confronted the Pearson coefficient (PC) of an image pair with correlation coefficients obtained between the green channel and randomized images of the red channel. To do so, they shuffled pixel blocks of one fluorescent channel in a randomized manner and measured the PC after each randomization round. They obtained a Gaussian distribution of PC after randomization and deduced that the PC obtained for the original image pair would be statistically significant if it is not included in the area of the Gaussian curve. With the DiAna plugin, we introduce a similar method, originating from spatial statistics analysis and described in Andrey et al. 2010 [22], applied to object-based co-localization (Fig. 6). First, objects from one image are randomly redistributed. The shuffle function allows to either redistribute the objects in a uniform manner within the whole image, or to import a binary image which defines regions in which redistribution is constrained (Fig. 6A–C). Shuffled images are generated, and for each of these images the centre-to-centre distances between objects of the randomized channel to the closest object in the second channel from the original image are computed. The cumulative distribution of the distances is plotted, and represented as the mean (Fig. 6D, red curves) flanked by 95% confidence intervals of the results (Fig. 6D, green curves). In parallel, observed centre-to-centre dis-



**Fig. 5.** Co-localization analysis based on localization of object's centres with the plugin DiAna A. Distance measurements between all points from two images allows the analysis of the distance distribution of objects. The coloured bars in the histogram illustrate threshold values, which can be used to define co-localization according to criteria defined in B–D. B. Co-localization can be defined for either a zero distance or a distance equal to voxel size. C. Co-localization can be defined for distances which are lower than the resolution of the imaging system. D. In case the minimal size of the biological objects are known, co-localization can be defined for distances which are lower than the sum of the minimum radii of objects A and B. Scale bars in B–D: 300 nm. Pixel size: 60 nm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tances between objects from the non-randomized original images are measured and plotted on the same graphic (Fig. 6D, blue curves). Statistical significance is assessed by the null hypothesis that the experimental data are due to randomness. If the distribution of the distances from experimental images falls outside the confidence interval of the distance distribution obtained for shuffled images in which object locations are random, one concludes that there is less than 5% chance ( $p < 0.05$ ) that the observed distribution is random and thus the co-localization is considered as statistically significant. In addition, the plugin calculates the rank of the observed distribution within  $n$  distributions obtained from shuffled images. A rank lower than 0.025 or higher than 0.975 indicates that the probability that the observed distribution is random is inferior to 5% and the co-localization is then considered signifi-

cant. In the example of the Fig. 6, the co-localization of VGLUT1 and synaptophysin was assessed and the analysis shows that the co-localization is significant, which is expected as these two proteins are both found in the vesicles from the presynaptic element of neurons [35,41].

Note that this methodology for assessing statistical robustness using a randomization procedure is quite generic and can be applied to other functions for co-localization analysis and generally to many other problems.

#### 4. Conclusion

We have developed an ImageJ-based tool named DiAna, allowing for spatial analysis in the three dimensions. In this tool we implemented two methods of 3D-segmentation, which show faithful and robust object extraction despite high variability of object size and intensity within the image. Furthermore, the tool allows executing extended object-based co-localization and distance analysis between objects in 3D. After co-localization or distance analysis, quantifications for each object are possible. Finally, we introduce a new method for the estimation of statistical significance of object-based co-localization. The algorithms we developed are implemented in a user-friendly plugin, which allows for complete but intuitive 3D image analysis, applicable to a large variety of biological objects.

The plugin with instructions for use can be found at [[http://imagejdocu.tudor.lu/doku.php?id=plugin:analysis:distance\\_analysis\\_diana\\_2d\\_3d\\_start](http://imagejdocu.tudor.lu/doku.php?id=plugin:analysis:distance_analysis_diana_2d_3d_start)].

#### Declaration of interest

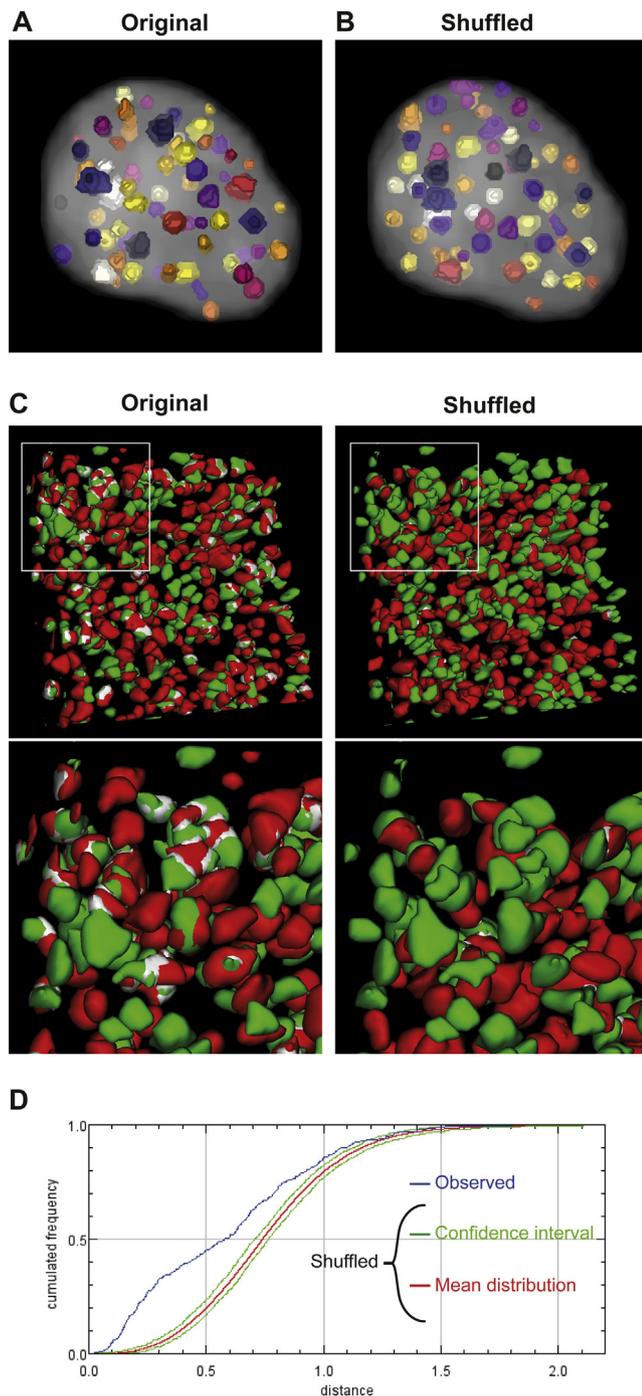
The authors declare no competing financial interests.

#### Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. TB's international mobility was partly funded by the programme "Investissements d'avenir" ANR-11-IDEX-0004-02.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jymeth.2016.11.016>.



**Fig. 6.** Statistical significance of the object based co-localization. A. DiAna can perform randomization of objects localization within a selected region in the image. 3D surfacing rendering of objects segmented within a mask. Each spot corresponds to an object and the mask volume appears in grey. B. View of the same objects as in A following the application of a shuffle procedure. The positions of the objects are randomly redistributed within the corresponding mask volume. C. 3D surface rendering of segmented objects from two images. Images were obtained from brain sections immunolabeled for synaptic proteins (green: vesicular transporter VGLUT1, red: synaptophysin). In the shuffled image, red objects are randomly redistributed. The co-localizing volumes appear in white. Note that their occurrence is lower after shuffling. Lower panels shows close-ups from the regions of the upper panel. The objects have an average diameter of 600 nm. D. The DiAna plugin provides a graphic, which represents the cumulative distribution of the minimum centre-to-centre distances between objects from two images. The blue curve shows the distribution for the experimental images shown in B. The red curve shows the mean distribution of distances between objects from the experimental green images and from 100 red images obtained by the shuffle procedure. The green curve represents the 2.5 and 97.5% confidence intervals around the mean. The experimental curve (in blue) is localized outside the 95% confidence interval (in green) of the distance analysis done after randomization, the co-localization is thus considered as statistically significant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

## References

- [1] Q. Li, A. Lau, T.J. Morris, L. Guo, C.B. Fordyce, E.F. Stanley, A syntaxin 1, Galpho, and N-type calcium channel complex at a presynaptic nerve terminal: analysis by quantitative immunocolocalization, *J. Neurosci.* 24 (2004) 4070–4081.
- [2] E. Manders, J. Stap, G. Brakenhoff, R. Van Driel, J. Aten, Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy, *J. Cell Sci.* 103 (1992) 857–862.
- [3] E. Lachmanovich, D.E. Shvartsman, Y. Malka, C. Botvin, Y.I. Henis, A.M. Weiss, Co-localization analysis of complex formation among membrane proteins by computerized fluorescence microscopy: application to immunofluorescence co-patching studies, *J. Microsc.* 212 (2003) 122–131.
- [4] B. Obara, A. Jabeen, N. Fernandez, P.P. Laissue, A novel method for quantified, super resolved, three-dimensional colocalisation of isotropic, fluorescent particles, *Histochem. Cell Biol.* 139 (2013) 391–402.
- [5] S. Bolte, F.P. Cordelieres, A guided tour into subcellular colocalization analysis in light microscopy, *J. Microsc.* 224 (2006) 213–232.
- [6] F. Jaskolski, C. Mulle, O.J. Manzoni, An automated method to quantify and visualize colocalized fluorescent signals, *J. Neurosci. Methods* 146 (2006) 42–49.
- [7] F.P. Cordelières, S. Bolte, Experimenters' guide to co-localization studies: finding a way through indicators and quantifiers, in practice, *Methods in Cell Biology* 123 (2014) ISSN 0091–679X, chapter 21.
- [8] T. Lagache, N. Sauvonnnet, L. Danglot, J.C. Olivo-Marin, Statistical analysis of molecule colocalization in bioimaging, *Cytometry A* 87 (2015) 568–579.
- [9] E. Manders, A. Visser, A. Koppen, W. De Leeuw, R. Van Lie, G. Brakenhoff, R. Van Driel, Four-dimensional imaging of chromatin dynamics during the assembly of the interphase nucleus, *Chromosome Res.* 11 (2003) 537–547.
- [10] S.V. Costes, D. Daelemans, E.H. Cho, Z. Dobbin, G. Pavlakis, S. Lockett, Automatic and quantitative measurement of protein–protein colocalization in live cells, *Biophys. J.* 86 (2004) 3993–4003.
- [11] B. Van Steensel, E. Van Binnendijk, C. Hornsby, H. Van Der Voort, Z. Krozowski, E. De Kloet, R. Van Driel, Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons, *J. Cell Sci.* 109 (1996) 787–792.
- [12] E. Meijering, Cell segmentation: 50 years down the road, *IEEE Signal Process. Mag.* 29 (2012) 140–145.
- [13] V. Wiesmann, D. Franz, C. Held, C. Münzenmayer, R. Palmisano, T. Wittenberg, Review of free software tools for image analysis of fluorescence cell micrographs, *J. Microsc.* 257 (2015) 39–53.
- [14] T. Lagache, V. Meas-Yedid, J.C. Olivo-Marin, A statistical analysis of spatial colocalization using ripley's k function, in: *IEEE International Symposium on Biomedical Imaging ISBI*, San Francisco, 2013. pp. 896–901.
- [15] E. Herzog, F. Nadrigny, K. Silm, C. Biesemann, I. Helling, T. Bersot, H. Steffens, R. Schwartzmann, U.V. Nägerl, S. El Mestikawy, J. Rhee, F. Kirchhoff, N. Brose, In vivo imaging of intersynaptic vesicle exchange using VGLUT1 Venus knock-in mice, *J. Neurosci.* 26 (2011) 15544–15559.
- [16] N. Heck, M. Dos Santos, B. Amairi, M. Salery, A. Besnard, E. Herzog, T. Boudier, P. Vanhoutte, J. Caboche, A new automated 3D detection of synaptic contacts reveals the formation of cortico-striatal synapses upon cocaine treatment in vivo, *Brain Struct. Funct.* 220 (2015) 2953–2966.
- [17] J. Gul-Mohammed, I. Arganda-Carreras, P. Andrey, V. Galy, T. Boudier, A generic classification-based method for segmentation of nuclei in 3D images of early embryos, *BMC Bioinf.* 14 (2014) 15–19.
- [18] J. Matas, O. Chum, M. Urban, T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, *Image Vis. Comput.* 22 (2004) 761–767.
- [19] L.A. Tavares, R.M. Souza, L. Rittner, R.C. Machado, R.A. Lotufo, Interactive max-tree visualization tool image processing and analysis, in: *Image Processing Theory, Tools and Applications* (2015) Conference article pp. 119–124.
- [20] J.L. Bentley, Multidimensional binary search tree used for associative searching, *Commun. ACM* 18 (1975) 509–517.
- [21] D.A. Randell, G. Landini, A. Galton, Discrete mereotopology for spatial reasoning in automated histological image analysis, *IEEE Trans. Pattern Recognit. Mach. Intell.* 35 (2013) 568–581.
- [22] P. Andrey, K. Kiêu, C. Kress, G. Lehmann, L. Tirichine, Z. Liu, E. Biot, P.G. Adenot, C. Hue-Beauvais, N. Houba-Hérin, V. Duranthon, E. Devinoy, N. Beaujean, V. Gaudin, Y. Maurin, P. Debey, Statistical analysis of 3D images detects regular spatial distributions of centromeres and chromocenters in animal and plant nuclei, *PLoS Comput. Biol.* 6 (7) (2010 Jul 8) e1000853.
- [23] O. Ronneberger, D. Baddeley, F. Scheipl, P.J. Vermeer, H. Burkhardt, C. Cremer, L. Fahrmeir, T. Cremer, B. Joffe, Spatial quantitative analysis of fluorescently labeled nuclear structures: problems, methods, pitfalls, *Chromosome Res.* 16 (2008) 523–562.
- [24] A.J. North, Seeing is believing? A beginners' guide to practical pitfalls in image acquisition, *J. Cell Biol.* 172 (2006) 9–18.
- [25] J.W. Lichtman, J.A. Conchello, Fluorescence microscopy, *Nat. Methods* 12 (2005) 910–919.
- [26] N. Heck, S. Betuing, P. Vanhoutte, J. Caboche, A deconvolution method to improve automated 3D-analysis of dendritic spines: application to a mouse model of Huntington's disease, *Brain Struct. Funct.* 217 (2012) 421–434.
- [27] J.G. McNally, T. Karpova, J. Cooper, J.A. Conchello, Threedimensional imaging by deconvolution microscopy, *Methods* 19 (1999) 373–385.
- [28] W. Wallace, L. Schaefer, J. Swedlow, A workingperson's guide to deconvolution in light microscopy, *Biotechniques* 31 (2001) 1076–1097.
- [29] J. Ollion, J. Cochenec, F. Loll, C. Escudé, T. Boudier, TANGO: A generic tool for High-throughput 3D Image Analysis for Studying Nuclear Organization, *Bioinformatics* 29 (2013) 1840–1841.
- [30] [http://imagejdocu.tudor.lu/doku.php?id=plugin:stacks:3d\\_ij\\_suite:start](http://imagejdocu.tudor.lu/doku.php?id=plugin:stacks:3d_ij_suite:start).
- [31] K. Staras, T. Branco, J.J. Burden, K. Pozo, K. Darcy, V. Marra, A. Ratnayaka, Y. Goda, A vesicle superpool spans multiple presynaptic terminals in hippocampal neurons, *Neuron* 66 (2010) 37–44.
- [32] C. Kempf, T. Staudt, P. Bingen, H. Horstmann, J. Engelhardt, S.W. Hell, T. Kuner, Tissue multicolor STED nanoscopy of presynaptic proteins in the calyx of Held, *PLoS One* 8 (2013) e62893.
- [33] C.C. Garner, S. Kindler, E.D. Gundelfinger, Molecular determinants of presynaptic active zones, *Curr. Opin. Neurobiol.* 10 (2000) 321–327.
- [34] F.E. Schweizer, T.A. Ryan, The synaptic vesicle: cycle of exocytosis and endocytosis, *Curr. Opin. Neurobiol.* 16 (2006) 298–304.
- [35] T.C. Südhof, R. Jahn, Proteins of synaptic vesicles involved in exocytosis and membrane recycling, *Neuron* 6 (1991) 665–677.
- [36] C. Fouquet, J.F. Gilles, N. Heck, M. Dos Santos, R. Schwartzmann, V. Cannaya, M. P. Morel, R.S. Davidson, A. Trembleau, S. Bolte, Improving axial resolution in confocal microscopy with new high refractive index mounting media, *PLoS ONE* 10 (2015) e0121096.
- [37] L. Landmann, P. Marbet, Colocalization analysis yields superior results after image restoration, *Microsc. Res. Tech.* 64 (2004) 103–112.
- [38] L. Landmann, Deconvolution improves colocalization analysis of multiple fluorochromes in 3D confocal data sets more than filtering techniques, *J. Microsc.* 208 (2002) 134–147.
- [39] T. Abraham, S.E. Allan, M.K. Levings, Deconvolution and chromatic aberration corrections in quantifying colocalization of a transcription factor in three-dimensional cellular space, *Micron* 41 (2010) 633–640.
- [40] J. Moss, J.P. Bolam, A dopaminergic axon lattice in the striatum and its relationship with cortical and thalamic terminals, *J. Neurosci.* 28 (2008) 11221–11230.
- [41] M. Grønberg, N.J. Pavlos, I. Brunk, J.J. Chua, A. Münster-Wandowski, D. Riedel, G. Ahnert-Hilger, H. Urlaub, R. Jahn, Quantitative comparison of glutamatergic and GABAergic synaptic vesicles unveils selectivity for few proteins including MAL2, a novel synaptic vesicle protein, *J. Neurosci.* 30 (2010) 2–12.



# Interactive Exploration for Continuously Expanding Neuron Databases



Zhongyu Li<sup>a</sup>, Dimitris N. Metaxas<sup>b</sup>, Aidong Lu<sup>a</sup>, Shaoting Zhang<sup>a,\*</sup>

<sup>a</sup> Department of Computer Science, University of North Carolina at Charlotte, USA

<sup>b</sup> Department of Computer Science, Rutgers University, USA

## ARTICLE INFO

### Article history:

Received 15 July 2016

Received in revised form 15 February 2017

Accepted 16 February 2017

Available online 17 February 2017

### Keywords:

Neuron Morphology  
Large-Scale Retrieval  
Binary Coding  
Online Updating  
User Interaction

## ABSTRACT

This paper proposes a novel framework to help biologists explore and analyze neurons based on retrieval of data from neuron morphological databases. In recent years, the continuously expanding neuron databases provide a rich source of information to associate neuronal morphologies with their functional properties. We design a coarse-to-fine framework for efficient and effective data retrieval from large-scale neuron databases. In the coarse-level, for efficiency in large-scale, we employ a binary coding method to compress morphological features into binary codes of tens of bits. Short binary codes allow for real-time similarity searching in Hamming space. Because the neuron databases are continuously expanding, it is inefficient to re-train the binary coding model from scratch when adding new neurons. To solve this problem, we extend binary coding with online updating schemes, which only considers the newly added neurons and update the model on-the-fly, without accessing the whole neuron databases. In the fine-grained level, we introduce domain experts/users in the framework, which can give relevance feedback for the binary coding based retrieval results. This interactive strategy can improve the retrieval performance through re-ranking the above coarse results, where we design a new similarity measure and take the feedback into account. Our framework is validated on more than 17,000 neuron cells, showing promising retrieval accuracy and efficiency. Moreover, we demonstrate its use case in assisting biologists to identify and explore unknown neurons.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

Analyzing single neuron properties, such as cell types, brain regions, functions, and development stages, is usually a fundamental task for understanding the nervous system and brain working mechanism. Given the huge numbers of neuron cells in the human brain, it is infeasible to understand every neuron's properties through traditional biological experimentation and quantitative computation. Generally, neuron morphology plays a major role in determining neuron's network connectivity, functional and physiological properties. It is therefore reasonable and essential to explore neuronal properties according to their morphologies. Recent developments in the frontiers of neuroscience (e.g., BigNeuron [1]) have greatly facilitated research in neuron morphology, and an expanding number of neurons are being reconstructed and added to the public repositories [3,2]. These fast-growing large databases provide a new avenue to help biologists explore and analyze neuronal properties [4,40,24]. Specifically, given an unknown neuron, we can retrieve neurons with similar morpholo-

gies in the databases. These retrieved similar neurons can be used to identify the unknown neuron and discover latent knowledge of their morphologies and properties.

Recently, researchers have been actively investigating this neuron morphological retrieval issue. For example, Costa et al. [9] first presented *NBLAST* to measure pairwise neuronal similarity. *NBLAST* considers both the position and local geometry, decomposing neurons into short segments and score matched segments to decide the similarity level among neurons. Subsequently, Wan et al. [49] designed *BlastNeuron* for automated comparison, retrieval, and clustering of 3D neuron morphologies. In the retrieval stage, *BlastNeuron* searches similar neurons via the normalization of the ranked scores in terms of the similarity of feature vectors. Despite their high accuracy, these two methods could be inefficient when handling large-scale neuron databases. Mesbah et al. [32] proposed a data-driven hashing scheme, i.e., *Hashing Forest*, to search among large neuron databases. By establishing multiple unsupervised random forests, 128 or more binary bits are generated to represent morphological features. The *Hashing Forest* has achieved efficient and accurate results in neuron retrieval [7,6]. Nonetheless, it usually needs a large number of binary bits (e.g., larger than 128), and its efficiency can be further improved with shorter codes. Accordingly, how to search similar neurons in large-scale

\* Corresponding author.

E-mail address: [rutgers.shaoting@gmail.com](mailto:rutgers.shaoting@gmail.com) (S. Zhang).

databases with high efficiency and accuracy is the main focus in neuron retrieval.

As described in [32], binary coding and hashing have been widely applied in content-based image retrieval (CBIR), which target efficient similarity search in large-scale databases [50,56,55,58]. Binary coding/hashing methods usually train a coding function from the batched image databases, and then employ the coding function to transform image features into short binary codes. Many representative methods have been proposed in recent years, including, but not limited to, Spectral Hashing (SH) [51], Iterative Quantization (ITQ) [15], and Asymmetric Inner-product Binary Coding (AIBC) [48]. Despite the good performance in CBIR problem, they may not be directly applicable to the neuron retrieval problem, since the neuron databases are continuously expanding. Due to the recently well-developed neuron tracing techniques, an increasing number of neuron cells are reconstructed and added to the databases in a streamed manner. For example, the NeuroMorpho database [3] usually releases 1000–2000 reconstructed neuron cells with each update. If we re-train the coding function every time from scratch, using both the original and the newly added neuron batch, it is very time-consuming and adversely affects the efficiency of exploration.

Besides the problems of efficiency, there are two limitations influencing the retrieval accuracy of the neuron data. First, the neuron databases lack supervised information, i.e., no sufficient annotations to label the class of every neuron [4,34]. Generally, supervised retrieval is more accurate than unsupervised retrieval, since it can bridge the gap between low-level image descriptors and high-level semantic meaning. In the neuron retrieval problem, if we only consider the low-level neuronal morphologies, the retrieval results may not be consistent with their functions and properties. Second, binary coding can only provide coarse retrieval results for neuron morphological data [27,59,57]. As each neuron cell has a tree-topological structure, the difference among neuron morphologies can be subtle. Moreover, coding functions do not have a one-to-one correspondence when mapping morphological features into binary codes. This may present a difficult question, in which some unrelated neurons are represented by the same binary codes. The above two problems should be addressed to achieve good retrieval performance in the neuron morphological data.

To alleviate these problems, we designed a novel framework to achieve accurate and efficient data retrieval from large-scale neuron morphological databases. Specifically, we employ a matrix sketching technique [25] for binary coding, which can learn coding functions from the sketched neuron data, significantly reducing the matrix size of neuron databases, and continues to give good approximations with an orthogonal matrix. To tackle the continuously expanding neuron databases, we extend the binary coding with an online updating scheme, where the coding function can be updated on-the-fly without accessing whole neuron databases. Subsequently, based on the coarse retrieval results from binary coding, we introduce domain experts/users in our framework, which can give relevance feedback to improve the retrieval accuracy. In our feedback model, domain experts/users are only required to label the relevant samples with respect to query neurons from top- $z$  results. Then, the similarity levels of the unlabeled neurons are re-ranked accordingly through our newly designed similarity measure. To the best of our knowledge, this is the first work that focuses on the interactive exploration of the continuously expanding neuron databases.

The remaining paper is organized as follows: Section 2 briefly reviews work related to 3D neuron morphology and content-based image retrieval. Section 3 provides the details of binary coding and interactive neuron retrieval. Followed by experimental results and discussion of potential use cases in Section 4. Finally, Section 5 concludes the paper and presents future work.

## 2. Preliminaries

This paper pursues neuron exploration via the retrieval of morphological data, which has a strong multidisciplinary component that involves a nexus of ideas from neuroscience, machine learning and information retrieval. In this section, we provide an introduction to 3D neuron morphology and content-based image retrieval.

### 2.1. 3D neuron morphology

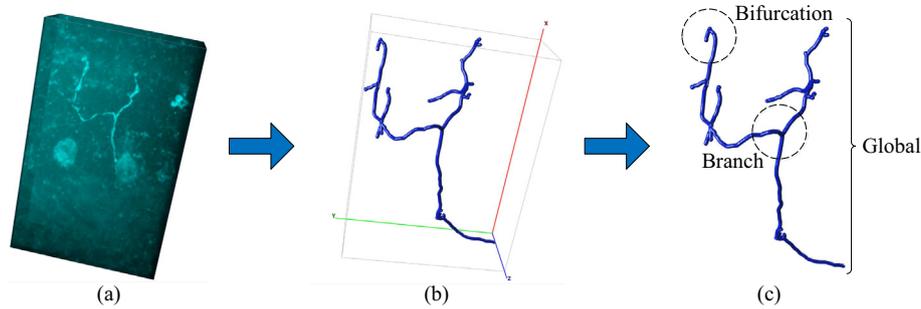
Benefiting from recent advances in microscopy imaging and neuron reconstruction [42,61,38], researchers have gradually elucidated the 3D neuron morphology. Currently, two kinds of microscopy imaging techniques can be used to obtain neuron images, i.e., light microscopy (LM) and electron microscopy (EM). LM images have long been used as a fundamental tool for neuroscientists. EM images can usually achieve higher resolution and magnification compared with LM images, but also have a higher cost and require harsher specimen processing. Researchers have developed a variety of methods for the reconstruction of LM and EM images [10,37,54]. Generally, from the original 2D microscopy images to the 3D morphological data, a neuron tracing system consists of several processing steps, including image preprocessing (e.g., noise reduction, deconvolution, mosaicing), segmentation (e.g., soma, dendritic trees, spines, axons segmentation), reconstruction and connection ([5,31,52,16,62,11,18]). Fig. 1(a) illustrates a neuron slice [33] which includes microscopy images acquired from multiple view points, and Fig. 1(b) presents the corresponding 3D reconstructed neuron morphological data through the neuron tracing software *Vaa3D* [38]. Accordingly, 3D morphological data can provide precise descriptions of neuronal shape and structure.

Given the 3D neuron morphological data, one critical problem is how to mathematically represent these neurons for the further retrieval and analysis, i.e., extracting feature vectors for each neuron. Unlike 2D biological images which can extract features with many well-studied algorithms, extracting good features from 3D neuron data is still a challenging problem. Costa et al. [8] first proposed the concept of neuromorphological space, which introduced many quantitative measurements (e.g., neuronal height, number of branches) based on the neuron's tree-topological structure. Subsequently, many researchers have employed these quantitative measurements as morphological features to represent each neuron [9,49,32]. For example, Wan et al. [49] employed several global measurements to reveal the overall morphology of neurons.

In this paper, we also utilize these quantitative measurements as neuron morphological features. Specifically, as illustrated in Fig. 1(c), we compute three levels of measurements to reflect neuron morphologies comprehensively, i.e., bifurcation, branch and global [23]. For example, the features in branch level indicate the measurements regarding neuron branches that are directly connected to soma, such as the branch length, the Euclidean distance from compartments to somas. The above three levels of measurements are invariant to translation, rotation and the tracing resolution. In the following sections, we employ the assembled measurements as feature vectors to represent each neuron cell for retrieval and analysis.

### 2.2. Content-based image retrieval

Content-based image retrieval (CBIR) is a long-term research topic which aims at searching similar images by their content. As a comprehensive application in computer vision and machine learning, CBIR has developed many branches for different concerns and targets. In this subsection, we introduce its two branches, i.e., hashing and human interaction, which are the main focus in our neuron retrieval framework.



**Fig. 1.** From original microscopy slice to 3D neuron morphology: (a) original microscopy slice; (b) 3D reconstructed neuron morphology; (c) three levels of measurements for feature extraction.

Numerous methods of binary coding/hashing have been proposed in recent years. In CBIR, by compressing long feature vectors into short binary codes, similarity search will be much more efficient in binary Hamming space compared to high dimensional feature space. The key challenge is how to obtain coding functions which can not only transform feature vectors via binary codes, but also keep similarity and diversity among the original data. One major taxonomy of hashing methods is whether they need specified training data to obtain the coding functions, i.e., data-independent and data-dependent. Locality-Sensitive Hashing (LSH) and its variants [14,20,41] are one of the most representative data-independent methods. Despite the fact that these methods can generalize coding functions to compact any given datasets, they usually need long bits of code to ensure good performance. For the data-dependent category, a large number of methods have been proposed in recent years. While these methods can only learn coding functions for given datasets, they are usually more accurate and efficient. Representative methods include Iterative Quantization (ITQ) [15], AGH [29], Isotropic Hashing (IsoHash) [19], Minimal Loss Hashing (MLH) [35,36], FastHash [26], etc. Despite the good performance that these methods have achieved, all of them are batch based methods, which learn coding functions from one batch of training data. For the continuously expanding neuron databases, as neurons are released in streamed fashion, we cannot directly adopt the above methods for the neuron retrieval problem.

In order to improve retrieval performance and reduce the semantic gap, some CBIR systems introduce domain experts/users in the loop, which can interactively provide relevance feedback for the previous retrieval results. Generally, for an image query and its coarse retrieval results, there are three kinds models to give relevance feedback: (1) positive feedback, where the users only need to select relevant images; (2) positive–negative feedback, where the users need to select both relevant and irrelevant images; (3) positive–neutral–negative feedback, where the users need to specify the degree of relevance for all the retrieved images. A comprehensive review of the early work on relevance feedback for CBIR is presented in [60]. Most of the early approaches use the marked images as individual queries and combine the retrieval results to refine the similarity weights of relevant images [45,44]. In recent years, many algorithms in the machine learning field have been used for the interactive CBIR problem, e.g., random forests [39], graph-cut [46], random walk [21], manifold learning [17]. All these methods can improve the retrieval performance with several rounds of interactive feedback. However, it is still a challenging issue when applying user interaction in large-scale neuron databases.

### 3. Methodology

This section presents the methodological details of our neuron retrieval framework, including binary coding with online updating and interactive neuron retrieval.

#### 3.1. Overview

Fig. 2 shows an overview of our neuron retrieval framework. The first part is training binary coding model (i.e., coding functions) for the neuron databases. We apply matrix sketching method on the feature vectors which are extracted from the original neuron databases, generating the initial values (including data sketch and virtual sample) for subsequent online updating. When a new neuron batch becomes available, we combine it with the aforementioned virtual sample. This combination can overcome the mean-varying problem in continuously expanding databases. The sketching result can subsequently be used for binary coding, updating the current coding function. The new data sketch and virtual sample are also stored for the next update. In the interactive neuron retrieval part, feature vectors of query neuron and all neurons in current databases can be compressed into short binary codes based on the updated coding functions. Then the similarity search between query neuron and neurons in databases are transformed into the Hamming distance ranking between their binary codes. To further improve the retrieval performance, we present top ranked neurons to users. Users will give relevance feedback to clarify which is similar with the query neuron. Our framework can process the above feedback and re-rank the retrieval results. Finally, the refined retrieved neurons can be used to help biologists to explore and analyze the query neuron.

#### 3.2. Binary coding with online updating

**Matrix Sketching based Binary Coding:** The goal of binary coding is to compress feature vectors into short binary codes, keeping diversities and similarities among original data. Denote a training neuron database  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n\} \subset \mathbb{R}^{n \times d}$ , which includes  $n$  neurons, and each neuron has  $d$  dimension of features. We aim to learn a coding function  $\mathbf{W} \in \mathbb{R}^{d \times r}$  that every normalized neuron feature in  $\mathbf{X}$  can be transformed into  $r$  bits of binary codes, i.e.,  $h(\mathbf{x}_i) = \text{sgn}((\mathbf{x}_i - \bar{\mathbf{X}})\mathbf{W})$ , where  $\bar{\mathbf{X}}$  is the mean value of  $\mathbf{X}$ . Note that feature normalization with zero mean is a crucial step in binary coding, especially for neuron data, because each dimension of features has their physical meaning. To learn effective binary codes, usually two requirements should be satisfied: (1) binary bits are uncorrelated and their variances are maximal; (2) numbers of 0 and 1 are roughly equal in learned binary codes of  $\mathbf{X}$ . Same as [50], the requirements are satisfied by maximizing the following objective function:

$$J(\mathbf{W}) = \frac{1}{n} \text{trace}(\mathbf{W}^T (\mathbf{X} - \bar{\mathbf{X}})^T (\mathbf{X} - \bar{\mathbf{X}}) \mathbf{W}), \quad \text{s.t. } \mathbf{W}^T \mathbf{W} = \mathbf{I}_{r \times r} \quad (1)$$

Instead of directly optimizing the above objective function, we apply the matrix sketching technique [25,12,22] on the training database to learn coding function from data sketch. Sketching is

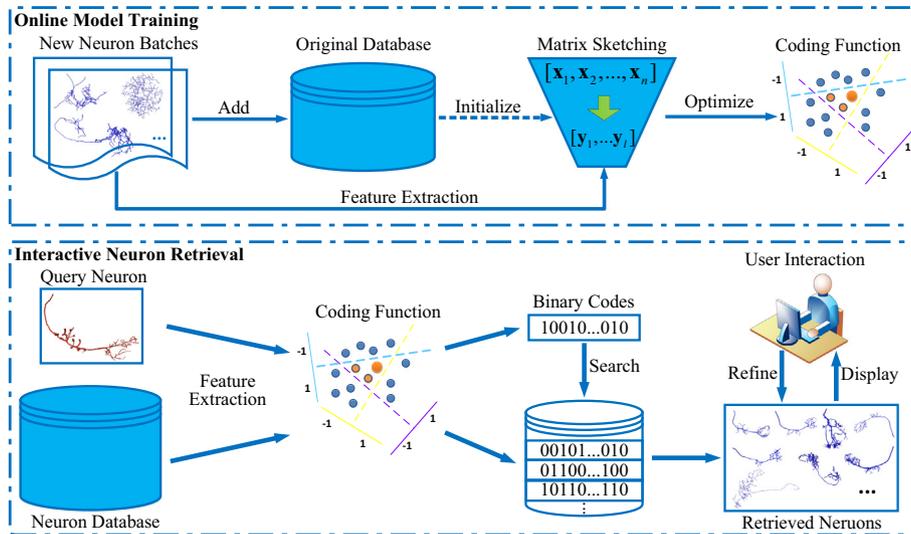


Fig. 2. Overview of our proposed framework, including online model training and interactive neuron retrieval.

a data compression technique which can significantly reduce the data size, without losing much data properties. Specifically, for the neuron database  $\mathbf{X}$ , we denote its matrix sketch as  $\mathbf{Y} \in \mathbb{R}^{l \times d}$ , which has the property  $\mathbf{Y}^T \mathbf{Y} \approx (\mathbf{X} - \bar{\mathbf{X}})^T (\mathbf{X} - \bar{\mathbf{X}})$ . We employ the *Frequent-directions* (FD) algorithm [13] to compute  $\mathbf{Y}$ , as this algorithm can effectively keep the property of matrix sketch. More importantly, FD is a streaming algorithm which can sequentially process the training data. In other words, when a new data batch comes, the FD algorithm will update the current sketch which only consider the new data one by one, without accessing the previously processed data. We show the advantage of such a streaming strategy for the online updated neuron data in the next subsection.

Given the data matrix  $\mathbf{X}$ , the FD algorithm can obtain its sketch  $\mathbf{Y}$  with a much smaller data size ( $l \ll n$ ). Then, the objective function of binary coding can be re-written as:

$$J(\mathbf{W}) \approx \frac{1}{n} \text{trace}(\mathbf{W}^T \mathbf{Y}^T \mathbf{Y} \mathbf{W}), \quad \text{s.t. } \mathbf{W}^T \mathbf{W} = \mathbf{I}_{r \times r} \quad (2)$$

This objective function is exactly the same as that of Principle Component Analysis (PCA). The optimal coding function  $\mathbf{W}$  can be obtained by taking the top  $r$  eigenvectors of the data covariance matrix  $\mathbf{Y}^T \mathbf{Y}$  [15]. In addition, to alleviate the unbalance of different dimensions in neuron data, we adopt orthogonal rotation  $\mathbf{R}$  for the above coding function, where  $\mathbf{W} = \mathbf{W}\mathbf{R}$ . However, since the sketch  $\mathbf{Y}$  is much smaller than the whole training data  $\mathbf{X}$ , we cannot learn the optimized  $\mathbf{R}$  as ITQ [15], which relies on all training data. Instead, we generate a random orthogonal rotation matrix, which achieves promising accuracy and efficiency in our experiments.

**Online Coding Function Updating:** When new neuron batches are added to the database, we need to update the coding function accordingly to maintain the retrieval performance. Re-training the coding function from scratch is very time-consuming, and sometimes infeasible when the existing neuron database is too large to load into memory. Considering that the FD algorithm can compute the data sketch in a streaming manner, an intuitive solution is to set the previous database sketch as the initial value, and then employ the FD algorithm to compute the sketch for newly added data. The coding function can be also updated with the newly computed data sketch. Unfortunately, this approach is impractical because of the aforementioned feature normalization requirement in binary coding. As the neuron database is continuously changing, the mean value for normalization is also changed.

How to overcome this mean-varying problem is a critical step to online update the coding function.

Assume  $\mathbf{B}_k$  is the newly added batch at round  $k$  which include  $m_k$  neurons, and the current database is denoted as  $\mathbf{X}_k = \{\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_k\}$ , where  $\mathbf{B}_0$  is the original neuron database. Then, the mean value of  $\mathbf{X}_k$  can be computed as:

$$\bar{\mathbf{X}}_k = \frac{\bar{\mathbf{X}}_{k-1} \cdot n_{k-1} + \bar{\mathbf{B}}_k \cdot m_k}{n_k} \quad (3)$$

where  $\bar{\mathbf{B}}_k$  is the mean value of  $\mathbf{B}_k$  and  $n_k = \sum_{i=0}^k m_i$ . Obviously, the mean value of the neuron database is changed in each update. To solve this problem, we introduce a virtual sample  $\mathfrak{V}_k$ , which considers the difference of mean value between the previous database and the current batch [43,22]:

$$\mathfrak{V}_k = \sqrt{\frac{n_{k-1} m_k}{n_k}} (\bar{\mathbf{B}}_k - \bar{\mathbf{X}}_{k-1}) \quad (4)$$

Combining this virtual sample with the currently added neuron batch, we can obtain a data set  $\hat{\mathbf{B}}_k$ :

$$\hat{\mathbf{B}}_k = [\mathbf{B}_k - \bar{\mathbf{B}}_k, \mathfrak{V}_k] \quad (5)$$

At round  $k$ , we have a new set of data  $\hat{\mathbf{X}}_k = \{\mathbf{B}_0 - \bar{\mathbf{B}}_0, \hat{\mathbf{B}}_1, \dots, \hat{\mathbf{B}}_k\}$ . According to [43,22], in each update,  $\hat{\mathbf{X}}_k$  takes the shift of mean into account and corrects such a shift by the virtual sample in  $\hat{\mathbf{B}}_k$ . More importantly, combining with Eq. (3), we find that  $\hat{\mathbf{X}}_k^T \hat{\mathbf{X}}_k = (\mathbf{X}_k - \bar{\mathbf{X}}_k)^T (\mathbf{X}_k - \bar{\mathbf{X}}_k)$ . This property indicates that the data sketch of  $\hat{\mathbf{X}}_k$  and  $\mathbf{X}_k - \bar{\mathbf{X}}_k$  is the same. Since  $\hat{\mathbf{X}}_k$  has no mean-varying problem, we can employ the aforementioned FD algorithm to sketch the continuously updated neuron data, i.e., sketch  $\hat{\mathbf{B}}_k$  to obtain  $\mathbf{Y}_k$ , initialized by the previous data sketch  $\mathbf{Y}_{k-1}$ . Then the coding function can be also updated on-the-fly via the matrix sketching based binary coding. In each update, without accessing the entire neuron databases, we only need to keep the mean value  $\bar{\mathbf{X}}_k$  (through Eq. 3), data size  $n_k$  (through  $n_k = n_{k-1} + m_k$ ) and the data sketch  $\mathbf{Y}_k$ . Therefore, this binary coding with online updating scheme can efficiently tackle the continuously expanding neuron databases for further retrieval and analysis.

### 3.3. Interactive neuron retrieval

According to the above binary coding method, as illustrated in Fig. 2, we can compress the query neuron and all neurons in the current database into short binary codes, through the learnt coding functions. Then the similar neurons can be retrieved based on their Hamming distance ranking with the query neuron. As discussed in Section 1, binary coding can only provide coarse retrieval results for the neuron morphological data. Therefore, given the coarsely retrieved neurons (e.g., neurons with top- $Z$  minimum Hamming distance), we propose to introduce domain experts/users in the framework, which can interactively provide relevance feedback to refine the retrieval results.

Fig. 3 presents an illustration of our user interaction interface. For a query neuron, it first searches similar neurons based on the aforementioned binary coding method. Then we display the top- $z$  ( $z = 16$  in Fig. 3) retrieved neurons to users, and users will compare and observe these neurons to decide whether they are relevant to the query neuron. This feedback scheme is easily implemented since it requires users to give only one-click inputs. Unlike many interactive models which require users to specify the class of the retrieval results [39,53,30], our strategy is particularly suitable for neuron databases which have insufficient annotations to classify every neuron.

After receiving the interactive feedback from users, our framework is able to process this feedback to improve the retrieval performance. Benefiting from the binary coding step which can efficiently provide the coarse retrieval results, we first define the outer scope size  $Z$ , where most similar neurons are included in the top- $Z$  coarse results. Subsequently, we define the inner scope size  $z$ , which represents the number of neurons that should be presented to the users at each feedback round. In practice,  $Z$  is larger than  $z$  but much smaller than the size of the whole neuron database. During the interactive neuron retrieval phase, we focus on re-ranking these  $Z$  coarse neurons to obtain fine-grained results.

In the  $t$ th round of relevance feedback,  $\mathbf{V}^{(t)}$  is the set of labeled similar samples from users which include  $m^{(t)}$  neurons. As neurons in  $\mathbf{V}^{(t)}$  are all similar with the query neuron, we can assemble them together to interpret and re-rank the similarities for the rest of the unlabeled neurons. Denoting  $E(\mathbf{x}_i, \mathbf{x}_j)$  as the similarity measure between two neurons  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , for one unlabeled neuron in  $t$ th round of relevance feedback  $\mathbf{x}_i^{(t)}$ , we re-define its similarity with the query neuron as follows:

$$E_i^{(t)} = \lambda E(\mathbf{x}_i^{(t)}, \mathbf{x}_q) + (1 - \lambda) \frac{1}{m^{(t)}} \sum_{j=1}^{m^{(t)}} E(\mathbf{x}_i^{(t)}, \mathbf{x}_j^{(t)}) \quad (6)$$

where  $\mathbf{x}_q$  is the query neuron and  $\mathbf{x}_j^{(t)}$  is the neuron in  $\mathbf{V}^{(t)}$ . The above objective function indicates that if an unlabeled neuron has similarity with the query neuron, it should be similar with the labeled neurons as well to some extent under the trade-off parameter  $\lambda$ .

According to Eq. (6), how to compute the similarity measure is a critical issue for re-ranking the unlabeled neurons. In most scenarios of image retrieval, the similarity measure between two images is defined as the Euclidean distance of their feature vectors. However, this similarity measure cannot be directly applied for the neuron morphological data, since each dimension of neuron features are distinct quantitative measurements which have different levels of representation. As discussed in Section 2.1, we compute three levels of measurements as features based on the neuron's tree-topological structure, i.e., global, branch and bifurcation. Accordingly, we propose to group features into these three levels and assign them with different representative weights (i.e.,  $\omega_{gl}$ ,  $\omega_{br}$ ,  $\omega_{bi}$ ) to compute a more accurate similarity measure. For

the two neurons  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , their new similarity measure can be formulated as:

$$E(\mathbf{x}_i, \mathbf{x}_j) = \omega_{gl} D_{gl}(\mathbf{x}_i, \mathbf{x}_j) + \omega_{br} D_{br}(\mathbf{x}_i, \mathbf{x}_j) + \omega_{bi} D_{bi}(\mathbf{x}_i, \mathbf{x}_j) \quad (7)$$

where  $D_{gl}$ ,  $D_{br}$ ,  $D_{bi}$  denotes the normalized Euclidean distance of global, branch and bifurcation features respectively. This similarity measure is specifically designed for neuron morphological data. In practice, the representative weights are determined by the neuronal tree-topological structure, and we will discuss it in the experiment.

Finally, with the newly defined similarity measure, we re-rank all the unlabeled neurons in ascending order based on their results in Eq. (6), and present the updated top- $z$  results to users. Users can iteratively give relevance feedback for these  $z$  neurons until they are satisfied with the retrieval results.

### 3.4. Implementation details

Given the query neuron  $\mathbf{x}_q$ , and the continuously expanding neuron database  $\{\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_k\}$ , our neuron retrieval method can efficiently obtain the similar neurons based on online binary coding and interactive feedback. We outlined the framework in Algorithm 1.

---

#### Algorithm 1. Neuron Retrieval based on Online Binary Coding and Interactive Feedback.

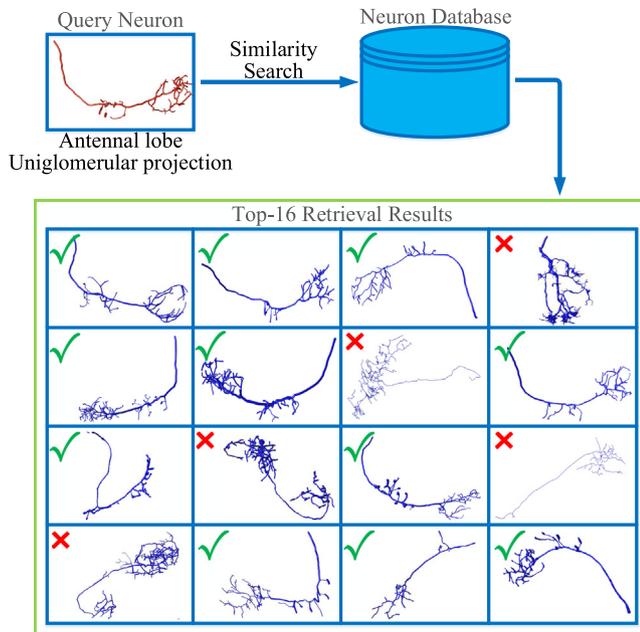
---

**Input:** Continuously added neuron database  $\{\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_k\}$ ;  
Query neuron  $\mathbf{x}_q$ .

**Output:** top- $S$  retrieved neurons.

- 1: Sketch  $\mathbf{B}_0 - \bar{\mathbf{B}}_0$  into  $\mathbf{Y}_0$ ;
  - 2: Initialize data size  $n_0 = m_0$ , mean value  $\bar{\mathbf{X}}_0 = \bar{\mathbf{B}}_0$ ;
  - 3: **for**  $i = 1 \rightarrow k$  **do**
  - 4: Sketch  $\hat{\mathbf{B}}_i = [\mathbf{B}_i - \bar{\mathbf{B}}_i, \mathfrak{I}_i]$  into  $\mathbf{Y}_i$ , initialize by  $\mathbf{Y}_{i-1}$ ;
  - 5: Update  $\mathbf{W}_i$  through Eq. (2);
  - 6: Update data size  $n_i = n_{i-1} + m_i$ ;
  - 7: Update mean value through Eq. (3);
  - 8: **end for**
  - 9: **if** retrieve  $\mathbf{x}_q$  is required **then**
  - 10: Binary encoding  $\mathbf{x}_q$  and  $\{\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_k\}$  through  $\mathbf{W}_i$ ;
  - 11: Rank the Hamming distance in ascending order;
  - 12:  $t = 0$ ;
  - 13: **while** users are not satisfied with the retrieval results **do**
  - 14:  $t = t + 1$ ; ( $t$ -th round of user feedback)
  - 15: Compute  $E^{(t)}$  of unlabeled neuron through Eq. (6);
  - 16: Re-rank the similarity based on  $E^{(t)}$ ;
  - 17: **end while**
  - 18: Present the top- $S$  retrieved neurons.
  - 19: **end if**
- 

In the above neuron retrieval algorithm, we combine binary coding and interactive retrieval to efficiently obtain the fine-grained retrieval results. The binary coding part can efficiently tackle the large-scale and continuously expanding neuron databases, which update the coding function on-the-fly every time a new neuron batch is added. For a query neuron, this part provides the coarse retrieval results and significantly reduces the search scope from tens of thousands to a few hundred (e.g., providing neurons with top- $Z$  minimum Hamming distance). In the interactive retrieval part, users can give relevance feedback for the top- $z$  ranked neurons. Then the proposed algorithm will re-rank the similarity of unlabeled neurons based on the results of Eq. (6). Finally, the  $Z$  coarse neurons will be re-ranked repeatedly until users are



**Fig. 3.** An illustration of our user interaction interface, users will give feedback by one-click inputs.

satisfied with the retrieval performance, and we will provide the top- $S$  retrieved neurons to biologists to assist the exploration and analysis of the query neuron.

## 4. Experiment

This section presents the evaluation of our framework for neuron retrieval. We first validate the performance for both the binary coding with online updating and interactive neuron retrieval. Then, we demonstrate an example of its use in neuron exploration and analysis.

### 4.1. Experimental setting

Our experiments are carried out on the NeuroMorpho [3], which has the largest collection of publicly accessible 3D reconstructed neuron data. Specifically, we use the entire 17,107 *Drosophila Melanogaster* neurons to evaluate the retrieval performance. Following the convention, we employ L-measure toolbox to extract 38 quantitative measurements as morphological features for each neuron [47], including 15 global, 10 branch and 13 bifurcation features respectively. All the experiments are conducted on a 3.6 GHz CPU with 4 cores and 32G RAM, in a MATLAB implementation.

To evaluate the retrieval performance, we select projection neurons as queries for which the brain region is the olfactory antennal lobe, and the cell classes are principal cell and uniglomerular projection (233 such projection neurons in total). We denote such projection neurons as uPNs. Despite that there are ten thousands of neurons in NeuroMorpho database, most of them are not well classified due to the lack of sufficient annotations, i.e., they are not identified in the finest level. We therefore select uPNs, since these are one of the best identified classes in the drosophila brain. This setting is also consistent with [9,49]. Then we evaluate the performance by computing the retrieval precision, which is defined as:

$$\text{Precision} = \frac{|\{uPNs\} \cap \{\text{Retrieved Neurons}\}|}{|\{\text{Retrieved Neurons}\}|} \quad (8)$$

where  $|\cdot|$  denote to count the number of samples inside. In the experiments, we compute the average precisions obtained over all queries.

In the binary coding with online updating, due to the requirement of the FD algorithm [13], the sketched data size  $l$  should be no larger than the feature dimension  $d$  ( $d = 38$  as discussed above). Since the feature dimension of neuron data is not high, we set  $l = 38$  to preserve the information from the original database as much as possible. In the interactive neuron retrieval, we adopt an outer scope size of  $Z = 300$  and an inner scope size of  $z = 30$ . The trade-off parameter  $\lambda$  is set as 0.3. For the defined similarity measure in Eq. (7), we assign three levels of features with different representation weights. Generally, global features can only represent neurons at a coarse level, and bifurcation features are so subtle that even neurons in same class are different. In practice, we empirically set global, branch and bifurcation features with the weights ratio of 1:2:1, which reflect their respective representative levels.

### 4.2. Evaluation of binary coding with online updating

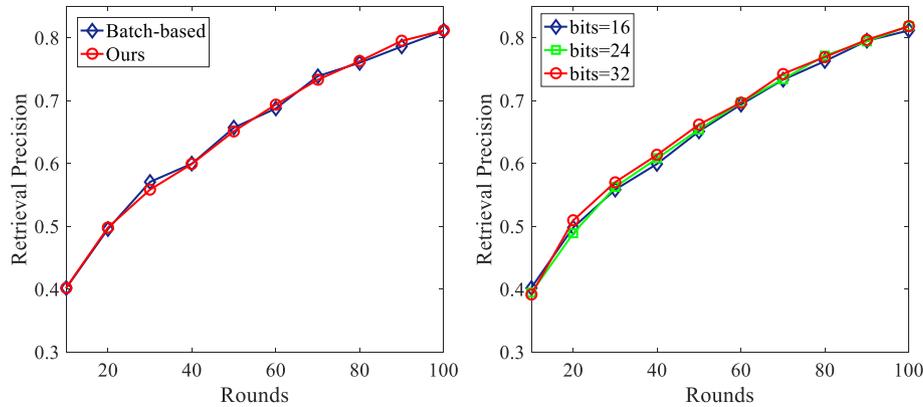
In this experiment, we aim to demonstrate that the binary coding part can attain promising performance with the continuously expanding neuron database. We randomly split the 17,107 *Drosophila Melanogaster* neurons into two parts. The first 1,107 neurons are used as the original database, and the remaining 16,000 neurons are equally divided into 100 batches (160 neurons in each), which are sequentially added to simulate the expanding size of the neuron database. Our online binary coding method is compared with the batch-based method. The batch-based method corresponds to the first part in Section 3.2, i.e., matrix sketching based binary coding. For each update, it needs to learn the coding function from scratch, using all neurons in the database. To overcome randomness, we repeat the experiments a hundred times to report the average.

Table 1 presents the accumulated training time of our method and batch-based method at the 20th, 60th and 100th update rounds respectively. Compared with the batch-based method, our binary coding with online updating shows great superiority in computational efficiency, and the superiority becomes more obvious with more rounds of updates, e.g., 51.11s versus 1.88s for one hundred updates. When new neuron batches are added to the database, our method only need to consider these newly added neurons and update the coding function on-the-fly, while the batch-based method needs to take all the neurons into account to re-train the coding function. The merit of this online binary coding method is particularly beneficial in the future, since an increasing number of neurons are reconstructed and added to the databases through the recently well-developed neuron tracing techniques.

Besides the superiority in computational efficiency, our binary coding with online updating also demonstrates its comparable performance in retrieval precision. Fig. 4(a) shows the average retrieval precision of two competitive methods, taking their top-10 retrieved neurons into accounts. The learned coding function compact the feature vectors into 32 bits of binary codes in this experiment. According to Fig. 4(a), our online method is able to achieve similar retrieval precision as the batch-based method. Therefore, the binary coding with online updating can significantly improve the computational efficiency without sacrificing the retrieval precision. This is mostly improved by (1) the employed FD algorithm,

**Table 1**  
Training time comparison (in second).

Rounds	20	60	100
Batch-based	1.12	13.24	51.11
Ours	<b>0.26</b>	<b>0.84</b>	<b>1.88</b>



**Fig. 4.** Evaluation of the retrieval precision in 100 rounds of update: (a) Comparison of our method with the batch-based method; (b) Comparison of our method using different bits of binary codes.

which can compute the data sketch in a streaming manner; and (2) the newly introduced batch of data with virtual samples, which can overcome the mean-varying problem in a continuously expanding neuron database. Regarding the parameter, Fig. 4(b) shows the retrieval precision of our method when using 16, 24 and 32 bits of binary codes in each update. The online method can always achieve good performance using different bits of binary codes. These results verify that the online updated coding function can generate effective and representative binary codes for neuron morphological features.

#### 4.3. Evaluation of interactive neuron retrieval

In this part, we aim to validate that the interactive strategy can actually achieve good performance for the neuron retrieval problem. We will re-fine the coarse retrieved neurons from previous binary coding results, where the neurons are retrieved from the entire 17,107 *Drosophila Melanogaster* neuron database after 100 rounds of database updating. Particularly, we only consider to refining neurons within top- $Z$  minimum Hamming distance (neurons in outer scope size). For the user interaction, users will give feedback for the top- $z$  unlabeled neurons (neurons in inner scope size).

To evaluate the performance of interactive neuron retrieval, at each feedback round, the top- $z$  unlabeled neurons will be automatically labeled using the ground truth in order to simulate the user's feedback. Since our interactive method only requires users to give one-click inputs (relevant or non-relevant), the ground truth can be easily achieved by checking whether the unlabeled neurons are uPNs or not.

We compare our neuron retrieval method with three state-of-the-art methods, i.e., ITQ [15], AGH [29] and MIPS [24], which were all proposed to tackle the retrieval problem for large-scale databases. ITQ [15] is a very effective binary coding method for most natural image retrieval problems. AGH [29] has already achieved excellent retrieval performance in mammogram data [28], and MIPS is specially designed for the neuron morphological retrieval problem. The above three methods are batch based method, which can not process the continuously expanding neuron databases. Thus, for fair comparison, the retrieval model of these methods are trained through the currently entire 17,107 *Drosophila Melanogaster* neuron database. As with the previous experiment setting, we employ the 233 uPNs as queries to validate the retrieval performance.

Table 2 reports the average retrieval precision of four competitive methods under different number of retrieved neurons. For our

**Table 2**

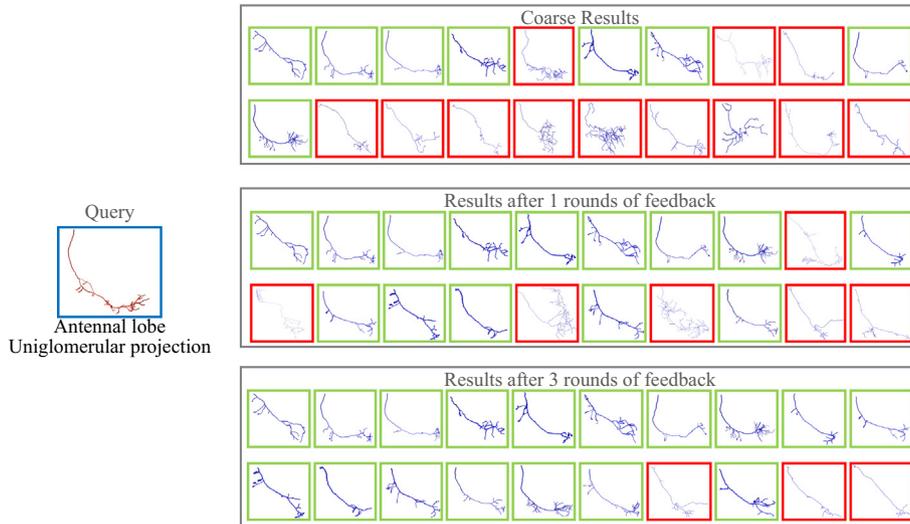
Retrieval precision of four methods under different number of retrieved neurons.

	top20	top30	top40	top50
ITQ [15]	0.7673	0.7249	0.6948	0.6614
AGH [29]	0.7589	0.7216	0.6951	0.6735
MIPS [24]	0.7923	0.7508	0.7088	0.6828
Ours	<b>0.9015</b>	<b>0.8550</b>	<b>0.7888</b>	<b>0.7092</b>

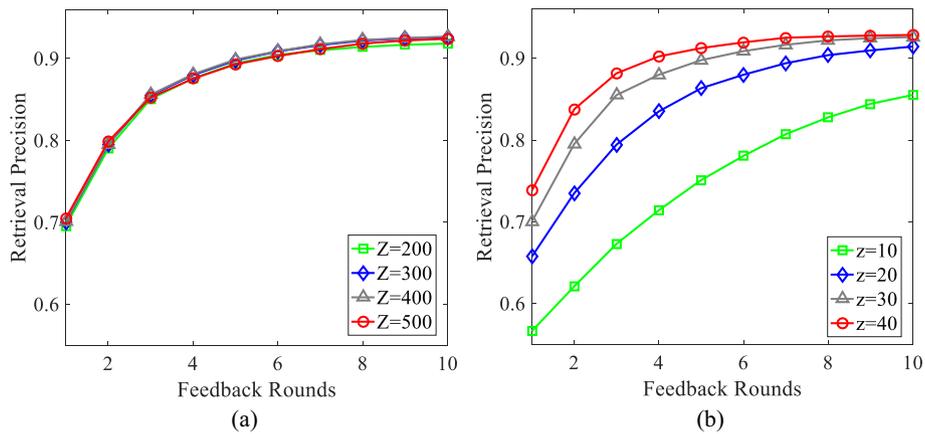
interactive method, the retrieval precision is recorded after 3 rounds of feedback. According to Table 2, our method can achieve the highest precision under different number of retrieved neurons. These results verify the proposed method is effective for the neuron retrieval problem. It mostly benefits from the interactive strategy which introduces users in the loop to give feedback for the coarse retrieval results. Specifically, based on the user's feedback, our method can re-rank unlabeled neurons by the newly designed similarity measure.

We randomly select a query neuron and present its top-20 retrieval results in Fig. 5 under different rounds of feedback. We employ *Vaa3D* [38] software to display these neurons. The neurons with green frames are relevant to the query, and neurons with red frames are not relevant to the query. Generally, the retrieval performance is improved greatly from coarse results to the results after user feedback, which verifies the effectiveness of the proposed interactive strategy. We also find that with the increased numbers of feedback rounds, the retrieval performance improves accordingly. This is because of the increasingly labeled neurons, providing more information for re-ranking. In addition, according to Fig. 5, many non-relevant neurons also present similar morphologies with the query, which is usually hard to distinguish through traditional retrieval methods. Thus, our interactive strategy is a good choice for the fine-grained neuron retrieval problem.

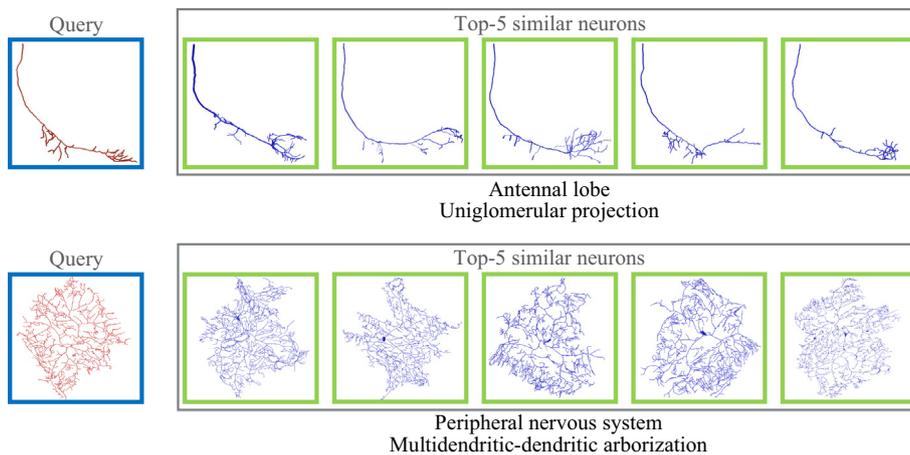
In our interactive neuron retrieval, two parameters may influence the final performance, i.e., the outer scope size  $Z$  and the inner scope size  $z$ . In the interactive part, we only consider the refinement of top- $Z$  ranked neurons from coarse retrieval results. Fig. 6 (a) shows the average retrieval precision with different outer scope sizes after the 1st to 10th feedback rounds, taking top-30 retrieved neurons into account. In Fig. 6(a), we find that with the outer scope size ranging from 200 to 500, the retrieval precision has not change too much. This is because the majority of relevant neurons are already included in a small sized outer scope (e.g.,  $Z = 300$ ). A larger outer scope may include more relevant neurons, but non-relevant and noisy neurons are also included which will



**Fig. 5.** Query example of the proposed method under different rounds of feedback: green framed neurons were relevant with the query, while red framed were non-relevant neurons.



**Fig. 6.** Retrieval performance with different parameter settings: (a) retrieval precision with different outer scope size after 1st to 10th feedback rounds; (b) retrieval precision with different inner scope size after 1st to 10th feedback rounds.



**Fig. 7.** Illustration of two unknown neurons and their top-5 retrieved neurons through the proposed method.

influence the retrieval performance. Moreover, the inner scope size  $z$  is the number of neurons we provide to users in each feedback round, which can also influence the final performance. Fig. 6(b)

presents the retrieval precision with different inner scope sizes after the 1st to 10th feedback rounds. According to Fig. 6(b), the larger  $z$  can achieve better performance compared with smaller  $z$

values. This is easy to understand since a larger  $z$  will contain more feedback information which can help us to re-rank the coarse retrieved neurons.

#### 4.4. Use cases and discussions

One important use case of our framework is the exploration and analysis for unknown neurons. Currently, despite the fact that an increasing number of neurons are reconstructed and added to the public databases, most of them are not well identified and lack basic annotations, such as cell classes and brain regions. Nanda et al. [34] proposed to annotate brain regions and cell classes for the NeuroMorpho [3] database. They employed the text-based query tool to search neurons with given lengths (e.g., 10, 20 microns) in each region to determine their brain regions, then they identified cell classes based on the brain regions invaded by the neurite terminals of every neuron. This method may be inefficient and unreliable, which the annotations are mainly obtained empirically. Identity of unknown neurons is an urgent demand in current neuron repositories.

Considering that neuron morphologies are associated with their properties, and our neuron retrieval framework can search similar neurons at a fine-grained level. It is reasonable to employ our framework to conduct neuron exploration via examining retrieved neurons, which have similar morphologies. To demonstrate this, we randomly select two query neurons from NeuroMorpho [3], whose neuron types are assumed to be unknown. After running our neuron retrieval framework (retrieve the entire 17,107 *Drosophila melanogaster* neurons, with 1 round of feedback), Fig. 7 illustrates their top-5 similar neurons respectively. For the two query neurons, we find that their corresponding top-5 similar neurons all have the same neuron types, i.e., in NeuroMorpho [3], the 5 neurons in first row are annotated as antennal lobe and uniglomerular projection (uPNs), the 5 neurons in the second row are annotated as peripheral nervous system and multidendritic-dendritic arborization. Therefore, we can infer that the two query neurons also have the same type with their top-5 retrieved neurons. The information provided in NeuroMorpho [3] also verifies our inference about the two query neurons. In practical situations, we can employ more retrieved neurons (e.g., top-30 similar neurons) to statistically identify and analyze query neurons.

The proposed method can efficiently achieve the above neuron retrieval task, as it is designed for the exploration of large-scale neuron databases. The method will be particularly suitable in the future since big data is one major direction in neuroscience [1]. Besides the efficiency, for some specific neuron databases which are not very large (e.g., considering neurons in some specific brain regions with only hundreds of neurons), exhaustive search and comparison can be applied to achieve more accurate results. In addition, extracting more representative features for the 3D neuro-morphological data will be also helpful to improve the neuron retrieval performance.

## 5. Conclusions

In this paper, we present a novel framework for neuron exploration and analysis, which interactively retrieves similar neurons in the continuously expanding neuron databases. Specifically, our framework achieves neuron morphological retrieval from coarse to fine-grained levels. In the coarse level, we introduce binary coding with online updating to tackle the large-scale and continuously expanding neuron databases. In each database update, coding functions are learned on-the-fly by only considering the newly added neuron data, and the coarse retrieval results are subse-

quently obtained in real-time. In the fine-grained level, we bring users in the loop, which interactively gives relevance feedback for the coarse results. By processing the feedback and re-ranking the coarse neurons, our framework finally obtains a set of fine-grained retrieval results. Experiments verify the efficacy of our neuron retrieval framework and also illustrates its application in neuron exploration. Based on the present work, we will develop a comprehensive tool for efficient and accurate neuron retrieval, which can help biologists to explore and analyze unknown neurons.

## Acknowledgement

We would like to acknowledge support for this project from the National Science Foundation (NSF Grant #1629913.).

## References

- [1] A. Bigneuron project, <http://www.alleninstitute.org/bigneuron/>, (Accessed June 28, 2016).
- [2] B. Bigneuron released data, <https://github.com/BigNeuron/Data/releases>, (Accessed June 28, 2016).
- [3] Neuromorpho repository, <http://neuromorpho.org/>, (Accessed June 28, 2016).
- [4] R. Armañanzas, G.A. Ascoli, Towards the automatic classification of neurons, *Trends Neurosci.* 38 (2015) 307–318.
- [5] H. Chen, H. Xiao, T. Liu, H. Peng, Smarttracing: self-learning-based neuron reconstruction, *Brain Inf.* 2 (2015) 135–144.
- [6] S. Conjeti, A. Katouzian, A. Kazi, S. Mesbah, D. Beymer, T.F. Syeda-Mahmood, N. Navab, Metric hashing forests, *Med. Image Anal.* 34 (2016) 13–29.
- [7] S. Conjeti, S. Mesbah, M. Negahdar, P.L. Rautenberg, S. Zhang, N. Navab, A. Katouzian, Neuron-miner: an advanced tool for morphological search and retrieval in neuroscientific image databases, *Neuroinformatics* 14 (2016) 369–385.
- [8] L.D.F. Costa, K. Zawadzki, M. Miazaki, M.P. Viana, S.N. Taraskin, Unveiling the neuromorphological space, *Front. Comput. Neurosci.* 4 (2010) 150–163.
- [9] M. Costa, A.D. Ostrovsky, J.D. Manton, S. Prohaska, G.S. Jefferis, Nblast: rapid, sensitive comparison of neuronal structure and construction of neuron family databases, *Neuron* 91 (2014) 293–311.
- [10] A. Fakhry, H. Peng, S. Ji, Deep models for brain em image segmentation: novel insights and improved performance, *Bioinformatics* 32 (2016) 2352–2358.
- [11] D. Feng, C. Lau, L. Ng, Y. Li, L. Kuan, S.M. Sunkin, C. Dang, M. Hawrylycz, Exploration and visualization of connectivity in the adult mouse brain, *Methods* 73 (2015) 90–97.
- [12] M. Ghashami, A. Desai, J.M. Phillips, Improved practical matrix sketching with guarantees, *Eur. Symp. Algorithms* (2014) 467–479.
- [13] M. Ghashami, E. Liberty, J.M. Phillips, D.P. Woodruff, Frequent directions: Simple and deterministic matrix sketching. *ArXiv Preprint:1501.01711*, 1–28, 2015.
- [14] A. Gionis, P. Indyk, R. Motwani, Similarity search in high dimensions via hashing, *Vldb* (1999) 518–529.
- [15] Y. Gong, S. Lazebnik, A. Gordo, F. Perronnin, Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 2916–2929.
- [16] S. Gulyanov, N. Sharifai, S. Bleykhan, E. Kelly, M. Kim, A. Chiba, G. Tschepnakis, Three-dimensional neurite tracing under globally varying contrast, *ISBI* (2015) 875–879.
- [17] J. He, M. Li, H.J. Zhang, H. Tong, C. Zhang, Manifold-ranking based image retrieval, *ACM MM* (2004) 9–16.
- [18] S. Ji, Computational genetic neuroanatomy of the developing mouse brain: dimensionality reduction, visualization, and clustering, *Bioinformatics* 14 (2013) 222–236.
- [19] W. Kong, W.J. Li, Isotropic hashing, *NIPS* (2012) 1646–1654.
- [20] B. Kulis, P. Jain, K. Grauman, Fast similarity search for learned metrics, *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (2009) 2143–2157.
- [21] B. Kulis, P. Jain, K. Grauman, Content-based image retrieval with relevance feedback using random walks, *Pattern Recogn.* 44 (2011) 2109–2122.
- [22] C. Leng, J. Wu, J. Cheng, X. Bai, H. Lu, Online sketching hashing, *CVPR* (2015) 2503–2511.
- [23] Z. Li, R. Fang, F. Shen, A. Katouzian, S. Zhang, Indexing and mining large-scale neuron databases using maximum inner product search, *Pattern Recogn.* 63 (2017) 680–688.
- [24] Z. Li, F. Shen, R. Fang, S. Conjeti, A. Katouzian, S. Zhang, Maximum inner product search for morphological retrieval of large-scale neuron data, *ISBI* (2016) 602–606.
- [25] E. Liberty, Simple and deterministic matrix sketching, *SIGKDD* (2013) 581–588.
- [26] G. Lin, C. Shen, Q. Shi, A. Hengel, D. Suter, Fast supervised hashing with decision trees for high-dimensional data, *CVPR* (2013) 1971–1978.
- [27] K. Lin, H.F. Yang, J.H. Hsiao, C.S. Chen, Deep learning of binary hash codes for fast image retrieval, *CVPR* (2015) 27–35.

- [28] J. Liu, S. Zhang, W. Liu, X. Zhang, D.N. Metaxas, Scalable mammogram retrieval using anchor graph hashing, *ISBI* (2014) 898–901.
- [29] W. Liu, J. Wang, S. Kumar, S.F. Chang, Hashing with graphs, *ICML* (2011) 1–8.
- [30] D. Markonis, R. Schaer, H. Müller, Evaluating multimodal relevance feedback techniques for medical image retrieval, *Inf. Retrieval J.* 19 (2016) 100–112.
- [31] E. Meijering, Neuron tracing in perspective, *Cytometry Part A* 77 (2010) 693–704.
- [32] S. Mesbah, S. Conjeti, A. Kumaraswamy, P. Rautenberg, N. Navab, A. Katouzian, Hashing forests for morphological search and retrieval in neuroscientific image databases, *MICCAI* (2015) 135–143.
- [33] S. Mukherjee, B. Condrón, S.T. Acton, Tubularity flow field technique for automatic neuron segmentation, *IEEE Trans. Image Process.* 24 (2015) 374–389.
- [34] S. Nanda, M. Allaham, M. Bergamino, S. Polavaram, R. Armañanzas, G. Ascoli, R. Parekh, Doubling up on the fly: neuromorpho. org meets big data, *Neuroinformatics* 1 (2015) 127–129.
- [35] M. Norouzi, D.M. Blei, Minimal loss hashing for compact binary codes, *ICML* (2011) 353–360.
- [36] M. Norouzi, D.M. Blei, R. Salakhutdinov, Hamming distance metric learning, *NIPS* (2012) 1061–1069.
- [37] H. Peng, M. Hawrylycz, J. Roskams, S. Hill, N. Spruston, E. Meijering, G.A. Ascoli, Bigneuron: large-scale 3d neuron reconstruction from optical microscopy images, *Neuron* 87 (2015) 252–256.
- [38] H. Peng, Z. Ruan, F. Long, J.H. Simpson, E.W. Myers, V3d enables real-time 3d visualization and quantitative analysis of large-scale biological image data sets, *Nat. Biotechnol.* 28 (2010) 348–353.
- [39] L. Peter, D. Mateus, P. Chatelain, N. Schworm, S. Stangl, G. Multhoff, N. Navab, Leveraging random forests for interactive exploration of large histological images, *MICCAI* (2014) 1–8.
- [40] S. Polavaram, T.A. Gillette, R. Parekh, G.A. Ascoli, Statistical analysis and data mining of digital reconstructions of dendritic morphologies, *Front. Neuroanat.* 8 (2014) 1–16.
- [41] M. Raginsky, S. Lazebnik, Locality-sensitive binary codes from shift-invariant kernels, *NIPS* (2009) 1509–1517.
- [42] E. Rebollo, K. Karkali, F. Mangione, E. Martín-Blanco, Live imaging in *Drosophila*: the optical and genetic toolkits, *Methods* 68 (2014) 48–59.
- [43] D.A. Ross, J. Lim, R.S. Lin, M.H. Yang, Incremental learning for robust visual tracking, *Int. J. Comput. Vision* 77 (2008) 125–141.
- [44] Y. Rui, T.S. Huang, S.F. Chang, Image retrieval: current techniques, promising directions, and open issues, *J. Vis. Commun. Image Represent.* 10 (1999) 39–62.
- [45] Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, Relevance feedback: a power tool for interactive content-based image retrieval, *IEEE Trans. Circuits Syst. Video Technol.* 8 (1998) 644–655.
- [46] H. Sahbi, J.Y. Audibert, R. Keriven, Graph-cut transducers for relevance feedback in content based image retrieval, *ICCV* (2007) 1–8.
- [47] R. Scorcioni, S. Polavaram, G.A. Ascoli, L-measure: a web-accessible tool for the analysis, comparison and search of digital reconstructions of neuronal morphologies, *Nat. Protoc.* 3 (2008) 866–876.
- [48] F. Shen, W. Liu, S. Zhang, Y. Yang, H. Tao Shen, Learning binary codes for maximum inner product search, *ICCV* (2015) 4148–4156.
- [49] Y. Wan, F. Long, L. Qu, H. Xiao, M. Hawrylycz, E.W. Myers, H. Peng, Blastneuron for automated comparison, retrieval and clustering of 3d neuron morphologies, *Neuroinformatics* (2015) 1–13.
- [50] J. Wang, S. Kumar, S.F. Chang, Semi-supervised hashing for large-scale search, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (2012) 2393–2406.
- [51] Y. Weiss, A. Torralba, R. Fergus, Spectral hashing, *NIPS* (2009) 1753–1760.
- [52] G. Wu, H. Jia, Q. Wang, D. Shen, Sharpmean: groupwise registration guided by sharp mean image and tree-based registration, *NeuroImage* 56 (2011) 1968–1981.
- [53] L. Zhang, L. Wang, W. Lin, Semisupervised biased maximum margin analysis for interactive image retrieval, *IEEE Trans. Image Process.* 21 (2012) 2294–2308.
- [54] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, D. Shen, Deep convolutional neural networks for multi-modality isointense infant brain image segmentation, *NeuroImage* 108 (2015) 214–224.
- [55] X. Zhang, H. Dou, T. Ju, J. Xu, S. Zhang, Fusing heterogeneous features from stacked sparse autoencoder for histopathological image analysis, *IEEE J. Biomed. Health Inf.* 20 (2016) 1377–1383.
- [56] X. Zhang, W. Liu, M. Dundar, S. Badve, S. Zhang, Towards large-scale histopathological image analysis: hashing-based image retrieval, *IEEE Trans. Med. Imaging* 34 (2015) 496–506.
- [57] X. Zhang, H. Su, L. Yang, S. Zhang, Fine-grained histopathological image analysis via robust segmentation and large-scale retrieval, *CVPR* (2015) 5361–5368.
- [58] X. Zhang, F. Xing, H. Su, L. Yang, S. Zhang, High-throughput histopathological image analysis via robust cell segmentation and hashing, *Med. Image Anal.* 26 (2015) 306–315.
- [59] X. Zhang, L. Yang, W. Liu, H. Su, S. Zhang, Mining histopathological images via composite hashing and online learning, *MICCAI* (2014) 479–486.
- [60] X.S. Zhou, T.S. Huang, Relevance feedback in image retrieval: a comprehensive review, *Multimedia Syst.* 8 (2003) 536–544.
- [61] Z. Zhou, X. Liu, B. Long, H. Peng, Tremap: automatic 3d neuron reconstruction based on tracing, reverse mapping and assembling of 2d projections, *Neuroinformatics* (2015) 1–10.
- [62] Z. Zhou, S. Sorensen, H. Zeng, M. Hawrylycz, H. Peng, Adaptive image enhancement for tracing 3d morphologies of neurons and brain vasculatures, *Neuroinformatics* (2014) 1–14.



# Investigation of podosome ring protein arrangement using localization microscopy images



Adela D. Staszowska, Patrick Fox-Roberts, Elizabeth Foxall, Gareth E. Jones, Susan Cox\*

Randall Division of Cell and Molecular Biophysics, King's College London, London, UK

## ARTICLE INFO

### Article history:

Received 15 August 2016  
Received in revised form 4 November 2016  
Accepted 7 November 2016  
Available online 10 November 2016

### Keywords:

Podosomes  
Super-resolution microscopy  
Localization microscopy  
Image analysis  
Pattern recognition

## ABSTRACT

Podosomes are adhesive structures formed on the plasma membrane abutting the extracellular matrix of macrophages, osteoclasts, and dendritic cells. They consist of an f-actin core and a ring structure composed of integrins and integrin-associated proteins. The podosome ring plays a major role in adhesion to the underlying extracellular matrix, but its detailed structure is poorly understood. Recently, it has become possible to study the nano-scale structure of podosome rings using localization microscopy. Unlike traditional microscopy images, localization microscopy images are reconstructed using discrete points, meaning that standard image analysis methods cannot be applied. Here, we present a pipeline for podosome identification, protein position calculation, and creating a podosome ring model for use with localization microscopy data.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Super-resolution microscopy is the term given to a class of techniques which can image structures smaller than the diffraction limit ( $\sim 250$  nm for visible light). Currently, localization microscopy is a highly popular method for achieving super-resolution due to its experimental simplicity. Localization microscopy is based on detection and localization of randomly activated single molecules in a sequence of images. These single molecule localizations are then used to reconstruct a super-resolution image [1,2]. This means that the reconstructed image is a collection of discrete points and should be considered as a data set rather than as an image.

Quantitative analysis of localization microscopy datasets requires identification of structures of interest in the data. Although pattern recognition is possible without a model of a structure, it is expensive computationally. Therefore it is easier to identify structures of interest either manually or by software using a set of rules, characteristics or a model of the structure. For example, a number of studies have discussed identification of the nuclear pore complex imaged with localization microscopy. This was done by creating an intensity profile from single molecule localizations and reconstructing. An average model was created either using localizations from many nuclear pore structures [3] or by convolving molecule localizations with a Gaussian to create a continuous image [4] and then identifying them.

Here, we propose a method using the Ransac algorithm [5] and an application of the Hough transform [6] to analyze localization microscopy data to identify podosomes, which are matrix adhesive structures formed on the cell surface of a number of cell types (for example macrophages, osteoclasts, dendritic cells, and Src-transformed fibroblasts [7]). Podosomes consist of an f-actin core and a ring structure formed of integrin and integrin-associated proteins (for example vinculin, paxillin, and talin) [7]. Podosomes are thought to be involved in adhesion, tissue transmigration and cancer metastasis [8]. Adhesion to the extracellular matrix, one of the functions of podosomes, is moderated and maintained mainly by the podosome ring [7].

Podosome rings in standard resolution fluorescence microscopy images are visible as circles surrounding the actin-rich podosome cores, and their radius varies between 0.5 and 1  $\mu\text{m}$  [9], with some evidence that diverse proteins are occupying discrete zones in the ring [10]. Recent studies using different super-resolution methods have presented two conflicting models of the podosome structure. The podosome ring was reported to have the shape of continuous hexagons in studies with high density localization microscopy methods using live and fixed samples expressing fluorescent proteins [11]. In contrast, when primary/secondary antibody labeled fixed samples were imaged with localization microscopy the ring was reported to be a collection of sparse protein clusters [10,12]. Results acquired using both of these approaches had suggested a specific protein arrangement in the ring. For example, a visual assessment of the images in [11] hinted that talin was closer to the podosome center than vinculin. This was partially confirmed by a

\* Corresponding author.

E-mail address: [susan.cox@kcl.ac.uk](mailto:susan.cox@kcl.ac.uk) (S. Cox).

confocal microscopy study presented in [10], which found talin formed an inner ring close to the core, with vinculin being uniformly distributed through the podosome ring. However, a comparison of the relative positions of different protein-classes found between the two experimental approaches has not been possible due to a lack of analysis methods suitable for localization microscopy data.

Our analysis method, described below, is tailored to be used with localization microscopy datasets. It identifies podosomes using a model of podosome ring structure, calculates the podosome ring protein positions and uses those positions to create a second, more refined model of the podosome ring. Podosomes were identified using an approximate, circular model of the ring structure and for each identification the center point and radius was measured. For identified podosomes we calculated the average position of the protein around the ring. As the ring size varies between different podosomes, we looked at the relative distance of one protein to another to amass reliable statistics. Relative protein distances can then be used to build a podosome ring model.

## 2. Methods

### 2.1. Sample preparation

The podosome samples were prepared using the protocol presented in [13]. Conjugated tandem dyes were supplied by Oleg Glebov and prepared using the protocol presented in [14].

### 2.2. Imaging

Localization microscopy imaging was performed using the Nikon STORM system, with an Eclipse Ti-E Inverted Nikon Microscope, Andor Ixon camera, laser and LED light sources (laser wavelengths are: 405 nm, 30 mW; 488 nm, 90 mW; 561 nm, 90 mW and 647 nm, 170 mW) and operated with NIS Elements software with the N-STORM module. The imaging was performed with TIRF, 100x, N.A. 1.49 objective. In two color STORM imaging the 647 nm laser was set to 25 mW, the 488 nm to 18 mW, and the 561 nm to 18 mW. The laser power was adjusted during the acquisition to acquire similar number of counts in every frame (as far as possible), up to around 80% of the maximal laser power (72 mW for 488 nm, 561 nm, and 136 mW for 647 nm laser). As podosome rings are rather flat structures positioned at the cell membrane, imaging was performed in TIRF (or near-TIRF) angle to reduce background and improve the signal-to-noise ratio.

Prior to imaging samples were immersed in an imaging buffer. The base buffer was made according to the Nikon Protocols for sample preparation [14] with  $\beta$ -Mercaptoethylamine (MEA, Sigma Aldrich, 30070-50G). To ensure better stability of dyes in the samples Cyclooctatetraene (COT, 98%, Sigma Aldrich, 138924-1G) was dissolved in DMSO (Sigma Aldrich, 472301-1L-D) and added to the base buffer to a final concentration of 2 mM [15].

In each series about 10,000 frames were acquired, at a rate of 30–50 frames per second. An epi-fluorescent image of the region of interest was also captured (with LED light, 488 nm and 561 nm) for each localization microscopy acquisition, which was later compared with the reconstructed image. The super-resolution images were reconstructed from the image sequences using QuickPALM [16]. The sample preparation, imaging and analysis for 3B microscopy was performed according to [11].

### 2.3. Podosome identification and protein position calculation

To identify podosomes, it was first necessary to create a model for them. We made a number of simplifications to the model which we used to identify the podosome ring structure. The ring was approximated to be circular (a circle's center and radius can be

found using coordinates of three points positioned on that circle). After the podosome identification the relative positions of proteins in the podosome ring were calculated. Because some of the podosomes were elongated, and the podosome size varies, the distance between the podosome center and proteins positioned in the ring is not constant between podosomes or even a single protein in one podosome ring. Thus the protein positions were calculated relative to each other by subtracting the average position of one protein from the average position of another.

The podosome rings were identified using a circular model. The equation of a circle passing through three points can be found using Eq. 1. Here we were interested in identifying the circle center and the radius of the fitted circle (see Section 3). The software randomly selected three points (separated by a small enough distance so they could confidently belong to a single podosome ring) from the localization data set (see step 1 in Fig. A.1) [5]. Then the circle center and radius were calculated (step 2 in Fig. A.1). The fitted circle was then examined to meet two criteria: the size of the fitted circle had to be similar to the size of actual podosomes (radius 0.5–1  $\mu\text{m}$ ) and the inside of the circle should have a very small number of localizations, because there should be no fluorophores present in the podosome cores. The acceptable number of localizations in the podosome center was established by finding the density of background localizations and then setting the threshold 50% higher (usually the same value was used for analysis datasets of the same type). Lastly, the overlap of the fitted circle with the actual podosome ring structure was evaluated, by examining the distribution of protein localizations surrounding the fitted circle in a proximity corresponding to an actual podosome ring thickness ( $\sim 400$  nm). The values used for this fitting step were established experimentally based on a fit to an average podosome ring from a localization microscopy dataset (step 3 in Fig. A.1). If these criteria were met the fitted circle parameters were saved (step 4 in Fig. A.1), if not they were discarded. Then a new set of three points was chosen and circle fitting steps were repeated (steps 1–3 in Fig. A.1). The circular model is optimal for the rounder podosomes. For elongated (elliptical) podosomes the fitted circle centers provided a wider range of possible podosome center points resulting in an elongated distribution (this can be seen for some podosomes in Fig. B.2b and c). Usually circles are fitted across the whole image, however, their density is much higher in areas with podosomes present.

The software ran for a set number of repetitions, which was arbitrarily selected to be around 100x higher than number of points in the data. The repetition number was smaller than the number of possible combinations of selecting three points, however it usually provided about 3000 circle fittings meeting the filtering criteria. For a data set containing up to twenty podosomes this provided a good range of possible podosome centers. The centers of fitted circles were saved and plotted in an image for reference (see Fig. B.2). The podosome centers were found using an application of the Hough transform [6]. The center points of the fitted circles were convolved with a Gaussian function. Then, the central points of actual podosomes were found by finding the local maxima (step 5 and 6 in Fig. A.1 and B.2). Compared with the visual assessment of the data the software had 80% success rate (the remaining 20% were false positives and negatives, which were removed during the protein distance calculation, see Fig. B.2).

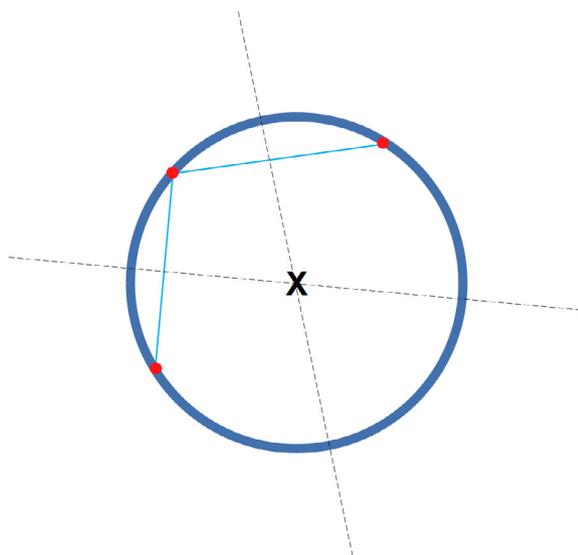
The protein positions in the podosome ring were calculated relative to each other, comparing the average position of one protein to the second. The podosome centers were used for the protein distance calculation. The outer boundary of the podosome ring was usually hard to define computationally because of background noise and close proximity between podosomes, although the rings boundaries are clearly visible to the human eye (see Fig. B.1). Thus a user input was required at this stage to create a mask image

marking podosome boundaries. These images were created in black-and-white, where white marked areas with podosomes (step 1 in Fig. A.2). Use of masking images had an additional advantage of excluding areas where podosome ring composition may be different. For example fragments of the ring shared by two joint podosomes and removing false positive podosome identifications (see Appendix B).

The podosome ring shape can sometimes be elongated, and so the proteins in the ring may not be equally spaced from the podosome center. Thus, the calculations of relative positions were performed for small sections of the ring the angular increment for each section was set to  $5.7^\circ$ . The increment value was selected to get on average about ten localizations of each protein in the segment (step 2 in Fig. A.2). The calculated average difference in positions of the two proteins in each segment was then weighted by the number of points in given segment – so the regions with small numbers of localizations would not influence strongly the end result (step 3 in Fig. A.2).

### 3. Equation of the circle passing through three points

There is exactly one circle passing through three points  $(x_1, y_1)$ ,  $(x_2, y_2)$ ,  $(x_3, y_3)$  which are not positioned on the same line. This



**Fig. 1.** A geometrical construction of a circle passing through three points. If we draw two perpendicular lines crossing the segments connecting points, then their crossing point will be the center of the circle.

circle can be constructed geometrically by drawing two perpendicular lines crossing the segments between points, their crossing point will be the center of the circle or arithmetically using equation of a circle:

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \quad (1)$$

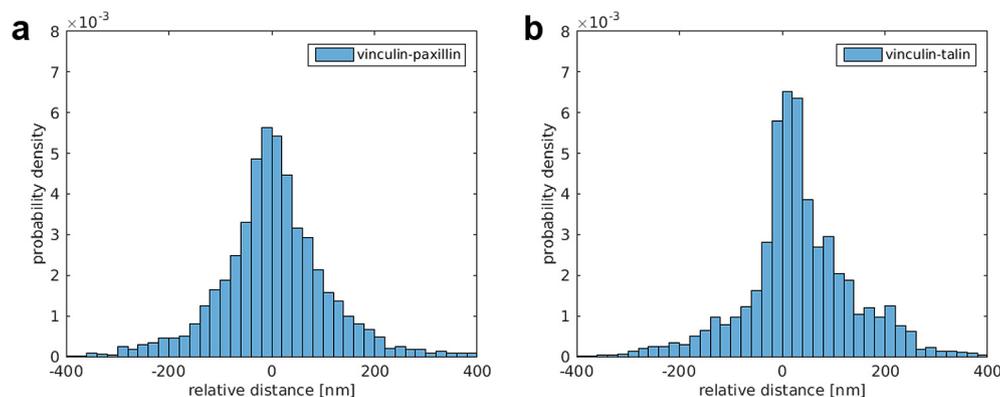
By solving Eq. 1 using three points positioned on the circle we found coordinates of the central point. The circle radius was later found by calculating distance between the circle center point and any of the three points on the circle (Fig. 1).

### 4. Results

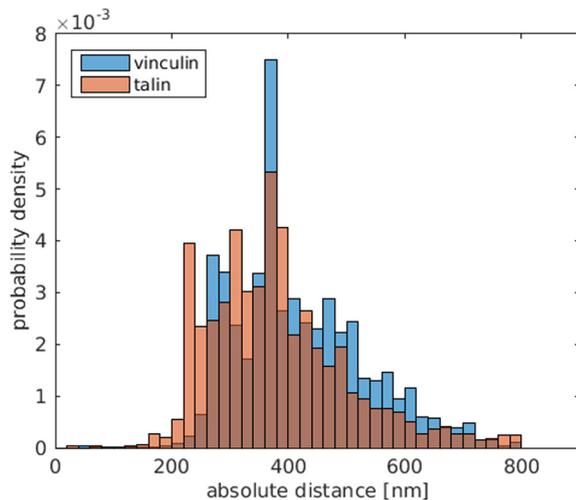
The methodology developed for podosome ring analysis was used to acquire preliminary results of protein positions in the podosome ring. The relative positions of vinculin-paxillin and vinculin-talin were calculated. Vinculin and paxillin data sets were collected using samples stained using tandem dye pairs and imaged using Nikon N-STORM system. The talin-vinculin pair was imaged using samples prepared with mCherry-talin construct and stained with Alexa Fluor 488 (see [11] for details). The relative positions of two proteins in each pair were calculated (see Section 2.3). The resulting distributions of relative distance measurement are shown in Fig. 2. For the vinculin-paxillin pair around 380 podosomes (from four samples, Fig. 1 a) and for the talin-vinculin pair 159 podosomes from a single data set were analyzed (see Fig. 1 b).

The results distribution for vinculin-paxillin is symmetric and it has a very sharp peak and weak tails. Mean, median, standard deviation and the 1st and 3rd quartiles were calculated for the results distribution (see Table 1). The central tendency of this distribution suggests a very small difference between positions of vinculin and paxillin. The mean value is equal to 4 nm and median 1 nm, suggesting that the paxillin is located further from the podosome core than vinculin. However, the mean and median values are still too small to provide a definitive answer about protein arrangement. The central tendency is smaller than the proteins size (minimal protein size was calculated using methodology presented in [17], see Table C.1).

Use of fluorescent proteins expressed directly by the protein of interest can remove the issue of the localization error due to distance between the detected fluorescent marker and marked protein, because they are expressed directly into the protein of interest. However, the quantum yield of the fluorescent proteins is smaller than that of organic dyes resulting in lower intensity and more dense data sets. Analysis of denser data sets requires more time and specially designed analysis methods for Example 3B [11]. Datasets with two of the podosome ring proteins vinculin



**Fig. 2.** Histograms of the position difference between two analyzed proteins from podosome rings. (a) The relative position was calculated as difference between the average position of paxillin from the vinculin position. For positive values paxillin is closer to the podosome core, for negative vinculin is closer. (Results for data sets acquired with standard localization microscopy.) (b) The relative positions of talin and vinculin (calculated by subtracting average talin position from vinculin position). For positive values talin is closer to the podosome center, where for negative vinculin is closer. (Results for 3B method analyzed images).



**Fig. 3.** Histograms of the absolute positions of talin and vinculin in podosome rings. Absolute positions were calculated from the podosome center.

stained with Alexa Fluor 488 and talin marked with mCherry was imaged and analyzed using 3B method [11]. Analysis of 159 podosomes was performed to measure relative distance of talin and vinculin (measured as a difference between average vinculin position and average talin position). The relative positions of talin in respect to vinculin (see Fig. 1b) and absolute protein positions (see Fig. 3) were calculated.

An analysis of more than 150 podosome rings provides an indication into talin and vinculin positions in the podosome ring. The relative position measurements indicated that on average talin is 33 nm closer to the podosome center than vinculin. When considering median value the relative distance is 24 nm. It is possible to analyze absolute distances of the protein positions in the ring, however, the podosome sizes in analyzed data should be similar.

Here, the absolute protein distances (measured from protein position to the podosome center) for talin and vinculin also suggest that talin is closer to the podosome core (see Fig. 3 and Table 2). We note that it is important to only compare measurements taken with the same technique as differences in the ring thickness could give rise to a bias in the measured radius (see Appendix D).

#### 4.1. Discussion

We have presented a methodology for podosome identification and calculation of the relative position of different proteins in the ring. The podosome rings were identified in localization microscopy data using a circular model of podosome structure. Our method provides a success rate of 80% (similar to one delivered

**Table 1**  
Statistical parameters of distributions of relative protein distance calculations for vinculin-paxillin and vinculin-talin protein pairs.

Protein pair	Mean relative distance [nm]	Median relative distance [nm]	Standard deviation	Quartiles
Vinculin-paxillin	4	1	110	(-40, 30)
Vinculin-talin	33	24	110	(0, 100)

**Table 2**  
Statistical parameters of results distributions of absolute positions of talin and vinculin.

Protein	mean distance [nm]	Median distance [nm]	Standard deviation	Quartiles
Vinculin	421	394	120	(360, 500)
Talin	388	366	131	(300, 440)

for podosome identification in confocal images [12]). Using podosome identifications the protein localized positions were used for calculations of relative protein distance. In order to build a complete model of the positions of different proteins the positions of vinculin, paxillin, and talin were compared with each other. This has enabled us to build up information about the average positions of these proteins in the podosome ring. Since we are looking at an average of the relative positions, this method allows us to use data from different localization techniques. The data also yield absolute values, which show some difference between the standard localization and 3B methods for the same protein (see Fig. D.3). Resolving the origin of these differences will require the comparison of multiple labelling and super-resolution techniques in the future.

The varying qualities of the images resulting from different localization microscopy methods highlights that labelling, imaging and image analysis methods can have a strong impact on data quality and interpretation. The localization error can be caused by a number of factors, starting with the method used to tag proteins. The primary/secondary antibody construct separates the detectable organic dye from the protein [18]. The organic dyes are usually around 4 nm in size and the antibody length is around 5 nm, which means that the labeled protein localization can differ up to 9 nm from the organic dye [19]. For this study the proteins of interest were stained with tandem dye pairs which could potentially introduce even bigger distance between the protein of interest and the organic probe. Another issue is that, although the monoclonal primary antibody used for staining does attach to a specific location on the protein, the information about the exact attachment sequence is not readily available. With primary/secondary labeling systems there can be clustering artefacts, as more than one secondary can attach to each primary. Clustering artefacts can also be caused by the reappearance of single molecules. Although this affects the final number of molecules detected [19], it is less important for the average position measurements as reappearances are thought to be equally likely for all molecules. Additionally, the thickness of the ring and density of labelling can also influence a measurement of protein positions (see Appendix D). Lastly, each localized molecule position is estimated using information delivered by photons coming from the molecule. Thus, the molecule position is estimated with uncertainty caused by a limited number of photons detected and it is approximately inversely proportional to the square root of the number of photons detected [20], while the exact uncertainty of a localization can be found from the Fisher information limit [21].

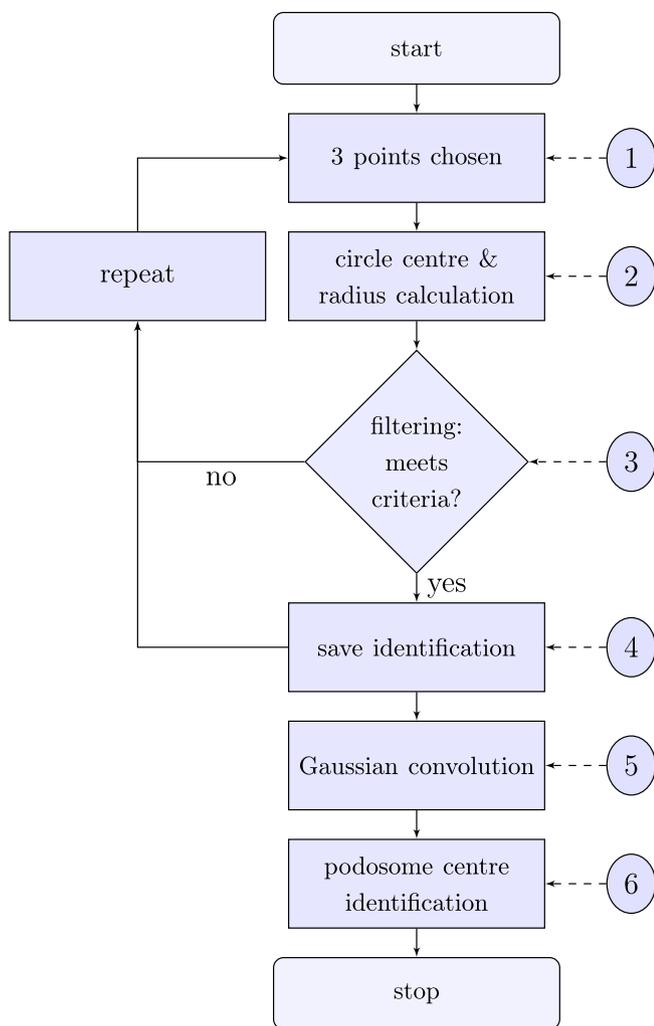
Overall, our analysis method provided more precise measurements for data sets acquired for samples where at least one protein was marked with fluorescent protein and the second with organic dye. Previously discussed factors limiting accuracy for primary-antibody staining are limited by use of only one dye to mark the protein of interest (here vinculin). Use of more than one fluorescent protein to mark proteins in the podosome ring could lead to even more precise measurement. The main source of localization error – use of two antibodies to attach a fluorescent marker – would be removed, because fluorescent markers would be synthesized directly in the desired protein.

Most importantly the pipeline presented here is ready to be used for a complete analysis of podosome ring protein positions. More data sets will be required to map positions of vinculin, paxillin, and talin to reference their relative distance. It is also possible to use software presented here for identification and analysis of a wide range of other structures, imaged with localization microscopy, which have circular ring shape, for example the nuclear pore complex. Lastly, the pipeline can be used to identify structures with different shapes after implementation of a different model of structure. The new model would have to be optimized to best describe the analyzed structure.

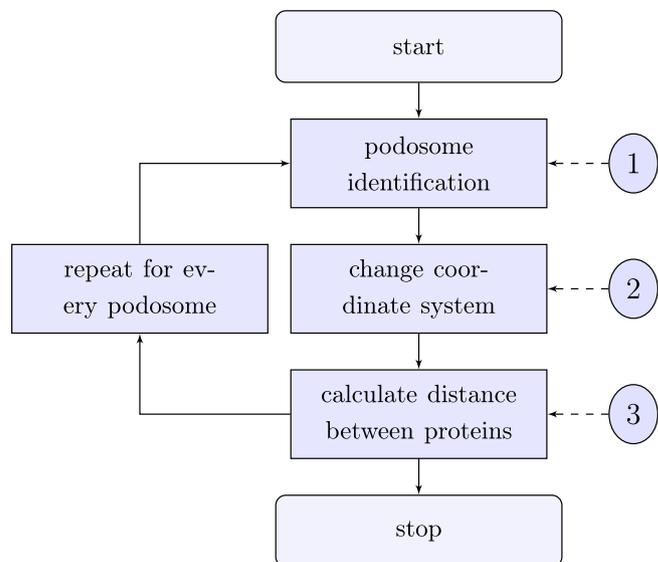
## Acknowledgements

The authors thank S. Mavadia, R. Szepietowski, and E. Rosten for early discussions, and O. Glebov for kindly supplying us with the tandem conjugated dyes. G.E.J and S.C were supported by the MRC (Next Generation Optical Microscopy Initiative Program Grant MR/K015664/1). S.C. by the BBSRC (Grant BB/K01563X/1) and S.C. is a Royal Society University Research Fellow. A.S. was funded by an EPSRC studentship and E.F. was funded by King's College studentship. The authors thank the Nikon Imaging Centre at King's College London, where localization microscopy data were acquired. The data supporting this research can be accessed using <http://doi.org/10.18742/RDM01-101>. For further information about the data please email the corresponding author or [adela.staszowska@kcl.ac.uk](mailto:adela.staszowska@kcl.ac.uk).

## Appendix A. Software flowcharts: Podosomes identification and protein distance calculation



**Fig. A.1.** The podosome localization software operation. (1) The three points were selected randomly from the data set. (2) The radius and a center point of circle passing by three points were calculated. (3) Filtering step to check if the fitted circle was close to the properties of the podosome ring and sample. If the filtering criteria were met the circle center was saved (step 4) the steps 1–3 were repeated. Otherwise a new set of points was chosen (back to step 1). (5) The circle center positions were convolved with the Gaussian and saved as an image. (6) The podosome centers were found by identifying local maxima on the Gaussian image created in step 5.



**Fig. A.2.** The operation of a program calculating positions of proteins in the podosome ring. The list of podosome centers made by the podosome center localization program is used. For every podosome center the points which belong to this podosome are identified using masking image (step 1). Then for an easier calculation the coordinate system was changed from Cartesian to polar coordinates (step 2) – the podosome center becomes a center of the polar coordinate system and, each point belonging to this podosome is described by the distance to the center and its angular position. The relative distance between two proteins was calculated and saved (step 3).

## Appendix B. Podosome identification and localization microscopy images

### Appendix C. The minimal protein volume and size

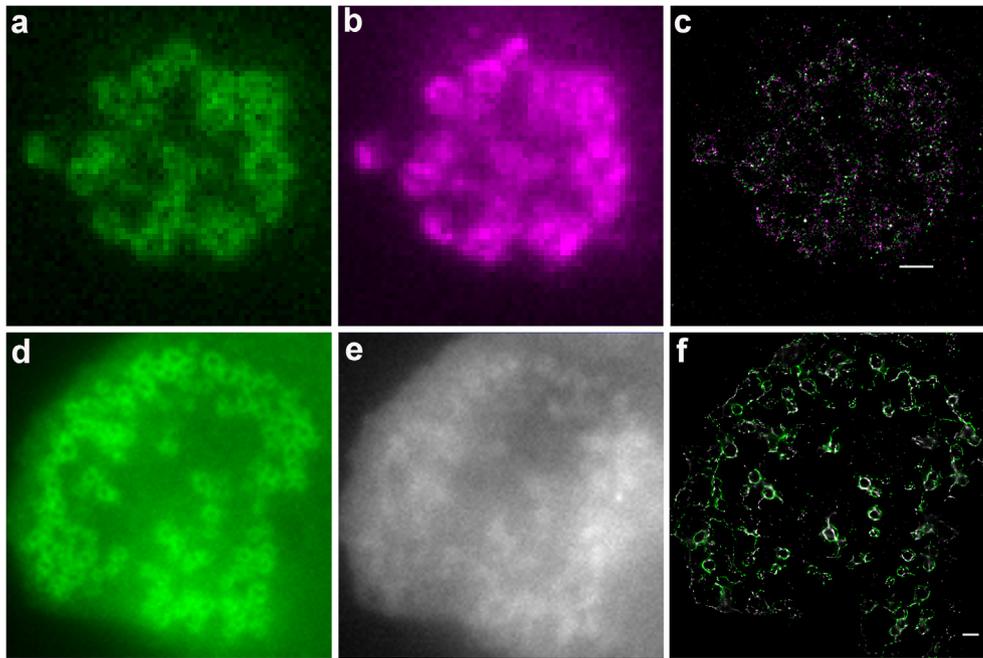
The minimal protein sizes were calculated according to the methodology presented in [17]. These calculations provide an estimation of the minimal size of the molecule and assume that it has a globular shape, however this is not true for a number of the proteins for example vinculin in its active state.

### Appendix D. The measured ring radius

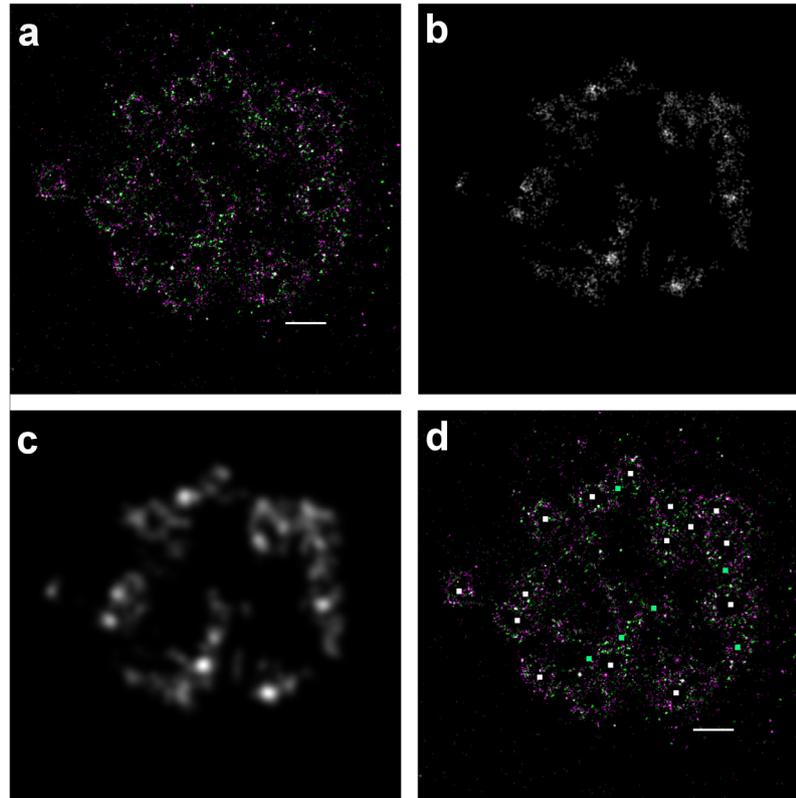
If the protein is evenly distributed through the podosome ring then the average position calculated will be slightly biased to the outer edge of the ring due to a higher number of molecules. However, if the fluorophores are equally likely to be anywhere on the ring then the outer side of the ring will have lower density of localizations than the inner side. Here, we have investigated how the size of this effect would vary with the ring thickness.

A number of datasets were simulated to account for the effect of sampling. Two sizes of inner rings with different thickness were simulated. These two types of rings correspond to the ratio of an average radius and thickness of the rings, observed using samples with two proteins labeled (here vinculin-paxillin) and one protein transfected and the second labeled (vinculin-talin). This ratio was estimated to be 0.6 (for vinculin-paxillin) and 0.1 (for vinculin-talin). The points/single molecule localizations in the rings were simulated with an equal density (see Fig. D.1). The averaged measured and expected radius are presented in Figure D.2.

The simulation indicated that the localization density (sampling) can have an influence on the measured value of the ring radius. For thicker rings the difference between the measured and expected values are bigger than for a thinner radius. However, since in our measurements we make comparisons between



**Fig. B.1.** Super-resolution imaging of podosome samples. Top row: (a-b) Wide-field and super-resolution reconstructed images of the podosome rings labeled with tandem dyes. Vinculin (green) was labeled with Cy2-Alexa Fluor 647 and paxillin (magenta) with Cy3-Alexa Fluor 647. Bottom row: (d-e) Wide-field and (f) 3B reconstructed super-resolution image of the podosome rings. Vinculin (green) was labeled with Alexa Fluor 488 and talin (grey) was transfected with mCherry-talin construct. Scale bar: 1  $\mu\text{m}$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

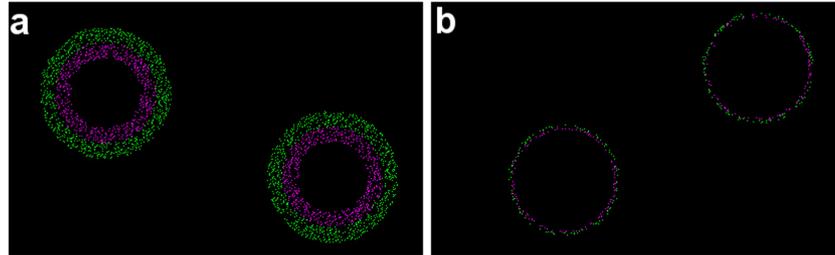


**Fig. B.2.** Podosome center identification for localization microscopy images. (a) Image of a cell displaying podosomes, vinculin (green) stained with Cy2-Alexa Fluor 647 and paxillin (red) with Cy3-Alexa Fluor 647. (b) Possible podosome centers found using the algorithm after applying filtration step (the image was blurred for improved visibility). (c) The center points were convolved with the Gaussian function. The actual centers of podosomes were found by identifying the local intensity maxima. (d) Identified podosome centers (marked with squares) displayed over the localization microscopy image of the podosome rings. White squares mark correct podosome identifications, green squares false positives, and white arrows false negatives (not identified podosomes). Scale bar 1  $\mu\text{m}$ . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

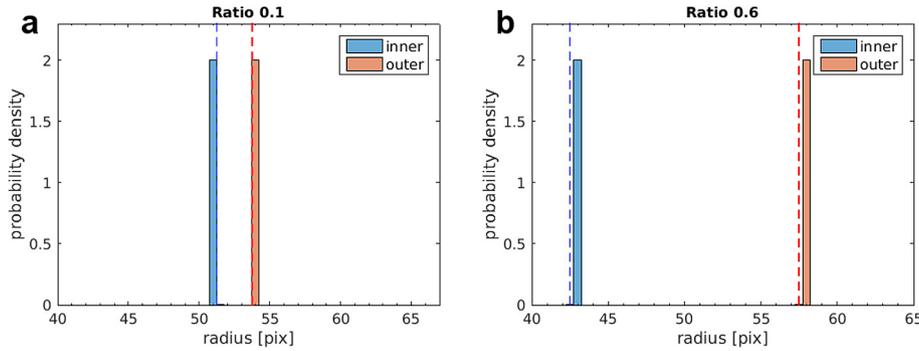
**Table C.1**

Minimal sizes of podosome ring proteins calculated using methodology presented in [17]. The minimal volume of space occupied by a protein with a certain mass can be calculated using equation  $R_{min}[nm] = 0.066M^{\frac{1}{3}}$ , where M is the mass of the the protein, measured in Daltons [17].

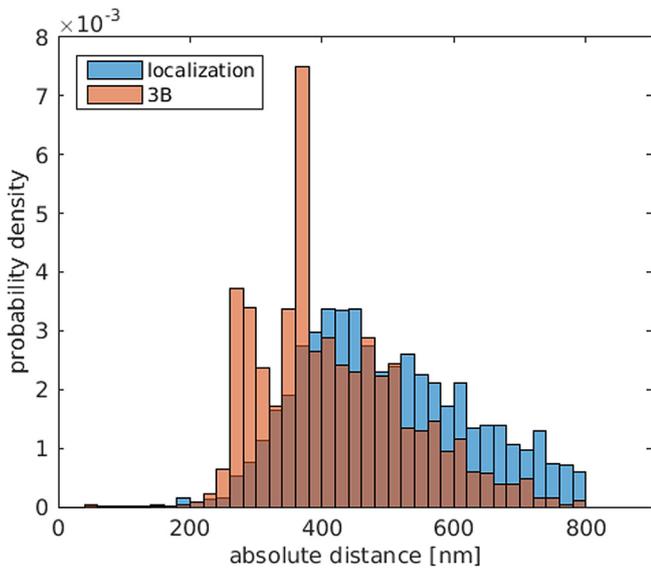
Protein	Mass [kDa]	$R_{min}$ [nm]	Notes
Vinculin	117	3.23	Circular only in inactive state
Paxillin	69	2.71	
Talin	270	4.27	



**Fig. D.1.** An example of simulated podosome rings. The rings were simulated as inner and outer concentric rings to correspond to the sample conditions. (a) The ration of thickness to radius is 0.6 corresponding to typical podosome ring observed with labeled samples with localization microscopy. (b) Thickness to radius ratio 0.1 corresponding to datasets analyzed with 3B method.



**Fig. D.2.** An average ring radius measured for different ratio of thickness and radius of simulated podosomes. Two values of ratio were simulated to correspond to observed average values observed in (a) 3B analyzed data sets and (b) labeled and QuickPALM analyzed localization microscopy. Podosome rings were simulated as two concentric rings corresponding to two proteins imaged for this study. Dashed lines correspond to an expected value of radius (calculated as a middle value between the ring radius). An average ring radius measured for different ratio of thickness and radius of simulated podosomes. (a) Measured average ring radius for a whole ring. (b) Inner and (c) outer concentric ring simulated to correspond to two protein imaged with localization microscopy. Dashed lines correspond to an expected value of radius (calculated as a middle value between the ring radius). Marked with blue for ratiom 0.6 and red for ratio 0.1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. D.3.** Comparison of the absolute positions of vinculin measured for localization microscopy and 3B datasets. The distributions are similar, however the distribution for localization microscopy data (blue) is shifted towards larger values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measurements with similar labeling and imaging techniques which exhibit very similar thicknesses, we do not believe that this will significantly bias our relative measurement.

There is, however a significant difference observed between the thickness of the rings observed using standard localization microscopy (thicker) and 3B (thinner). When directly comparing these two measurements (see Fig. D.3 and Table D.1), it appears that the 3B measurements are biased towards smaller values as might be expected if the density on the ring is a factor. This demonstrates the need to only compare results from similar techniques for the relative measurements.

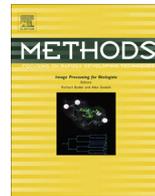
**Table D.1**

Statistical properties of the absolute positions of vinculin imaged with localization microscopy and 3B.

Protein	Mean distance [nm]	Median distance [nm]	Standard deviation	Quartiles
Localization	512	489	142	(420, 580)
3B	421	394	120	(360, 500)

## References

- [1] E. Betzig, G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, H.F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution, *Science* 313 (5793) (2006) 1642–1645.
- [2] M.J. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), *Nat. Methods* 3 (10) (2006) 793–795.
- [3] A. Löschberger, S. van de Linde, M.-C. Dabauvalle, B. Rieger, M. Heilemann, G. Krohne, M. Sauer, Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution, *J. Cell Sci.* 125 (9) (2011) 570–575.
- [4] A. Szymborska, A. de Marco, N. Daigle, V.C. Cordes, J.A.G. Briggs, J. Ellenberg, Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging, *Science* 341 (6146) (2013) 655–658.
- [5] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* 24 (6) (1981) 381–395.
- [6] C.L.L. Hendriks, M. van Ginkel, P.W. Verbeek, L.J. van Vliet, The generalized radon transform: sampling, accuracy and memory considerations, *Pattern Recogn.* 38 (12) (2005) 2494–2505.
- [7] V. Veillat, P. Spuul, T. Daubon, I. Egana, I. Kramer, E. Génot, Podosomes: multipurpose organelles?, *Int. J. Biochem. Cell Biol.* 65 (1) (2015) 52–60.
- [8] Y. Calle, S. Burns, A.J. Thrasher, G.E. Jones, The leukocyte podosome, *Eur. J. Cell Biol.* 85 (3–4) (2006) 151–157.
- [9] J. Monypenny, H.-C. Chou, I. Banón-Rodríguez, A.J. Thrasher, G.E.J. Inés, M. Antónb, Y. Calle, Role of wasp in cell polarity and podosome dynamics of myeloid cells, *Eur. J. Cell Biol.* 90 (2–3) (2011) 198–204.
- [10] M.B. Meddens, B. Rieger, C.G. Figdora, A. Cambi, K. van den Dries, Automated podosome identification and characterization in fluorescence microscopy, *Microsc. Microanal.* 19 (1) (2013) 180–189.
- [11] S. Cox, E. Rosten, J. Monypenny, T. Jovanovic-Talman, D.T. Burnette, J. Lippincott-Schwartz, G.E. Jones, R. Heintzmann, Bayesian localization microscopy reveals nanoscale podosome dynamics, *Nat. Methods* 9 (2) (2012) 195–200.
- [12] M.B. Meddens, K. van den Dries, A. Cambi, Podosomes revealed by advanced bioimaging: what did we learn?, *Eur. J. Cell Biol.* 93 (10) (2014) 380–387.
- [13] V. Vijayakumar, J. Monypenny, X.J. Chen, L. Machesky, S. Lilla, A.J. Thrasher, I. M. Antn, Y. Calle, G.E. Jones, Tyrosine phosphorylation of WIP releases bound WASP and impairs podosome assembly in macrophages, *J. Cell Sci.* (2014), <http://dx.doi.org/10.1242/jcs.154880>. <<http://jcs.biologists.org/content/early/2014/11/13/jcs.154880>>.
- [14] Nikon, Super Resolution Microscope N-STORM. STORM Protocol-Sample Preparation, [Online; retrieved on 10th of November 2015], 2015. <<https://nic.med.harvard.edu/sites/nic.med.harvard.edu/files/userfiles/N-STORM%20SAMPLE%20PREP.pdf>>.
- [15] R. Dave, D.S. Terry, J.B. Munro, S.C. Blanchard, Mitigating unwanted photophysical processes for improved single-molecule fluorescence imaging, *Biophys. J.* 96 (6) (2009) 2371–2381.
- [16] R. Henriques, M. Lelek, E.F. Fornasiero, F. Valtorta, C. Zimmer, M.M. Mhlanga, QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ, *Nat. Methods* 7 (5) (2010) 339–340.
- [17] H.P. Erickson, Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy, *Biol. Proced. Online* 11 (1) (2009) 32–51.
- [18] K.A.K. Tanaka, K.G.N. Suzuki, Y.M. Shirai, S.T. Shibutani, M.S.H. Miyahara, H. Tsuboi, M. Yahara, A. Yoshimura, S. Mayor, T.K. Fujiwara, A. Kusumi, Membrane molecules mobile even after chemical fixation, *Nat. Methods* 7 (11) (2010) 865–866.
- [19] A. Shivanandan, H. Deschout, M. Scarselli, A. Radenovic, Challenges in quantitative single molecule localization microscopy, *FEBS Lett.* 588 (19) (2014) 3595–3602.
- [20] R.E. Thompson, D.R. Larson, W.W. Webb, Precise nanometer localization analysis for individual fluorescent probes, *Biophys. J.* 82 (5) (2002) 2775–2783.
- [21] R.J. Ober, S. Ram, E.S. Ward, Localization accuracy in single-molecule microscopy, *Biophys. J.* 86 (2) (2004) 1185–1200.



# Mathematical imaging methods for mitosis analysis in live-cell phase contrast microscopy



Joana Sarah Grah<sup>a,\*</sup>, Jennifer Alison Harrington<sup>b</sup>, Siang Boon Koh<sup>b,1</sup>, Jeremy Andrew Pike<sup>b</sup>, Alexander Schreiner<sup>b,2</sup>, Martin Burger<sup>c</sup>, Carola-Bibiane Schönlieb<sup>a</sup>, Stefanie Reichelt<sup>b</sup>

<sup>a</sup> University of Cambridge, Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

<sup>b</sup> University of Cambridge, Cancer Research UK Cambridge Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, United Kingdom

<sup>c</sup> Westfälische Wilhelms-Universität Münster, Institute for Computational and Applied Mathematics, Einsteinstrasse 62, 48149 Münster, Germany

## ARTICLE INFO

### Article history:

Received 7 September 2016

Received in revised form 4 February 2017

Accepted 6 February 2017

Available online 9 February 2017

### Keywords:

Phase contrast microscopy

Mitosis analysis

Circular Hough transform

Cell tracking

Variational methods

Level-set methods

## ABSTRACT

In this paper we propose a workflow to detect and track mitotic cells in time-lapse microscopy image sequences. In order to avoid the requirement for cell lines expressing fluorescent markers and the associated phototoxicity, phase contrast microscopy is often preferred over fluorescence microscopy in live-cell imaging. However, common specific image characteristics complicate image processing and impede use of standard methods. Nevertheless, automated analysis is desirable due to manual analysis being subjective, biased and extremely time-consuming for large data sets. Here, we present the following workflow based on mathematical imaging methods. In the first step, mitosis detection is performed by means of the circular Hough transform. The obtained circular contour subsequently serves as an initialisation for the tracking algorithm based on variational methods. It is sub-divided into two parts: in order to determine the beginning of the whole mitosis cycle, a backwards tracking procedure is performed. After that, the cell is tracked forwards in time until the end of mitosis. As a result, the average of mitosis duration and ratios of different cell fates (cell death, no division, division into two or more daughter cells) can be measured and statistics on cell morphologies can be obtained. All of the tools are featured in the user-friendly MATLAB® Graphical User Interface *MitosisAnalyser*.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Mathematical image analysis techniques have recently become enormously important in biomedical research, which increasingly needs to rely on information obtained from images. Applications range from sparse sampling methods to enhance image acquisition through structure-preserving image reconstruction to automated analysis for objective interpretation of the data [1]. In cancer research, observation of cell cultures in live-cell imaging experiments by means of sophisticated light microscopy is a key technique for quality assessment of anti-cancer drugs [2,3]. In this context, analysis of the mitotic phase plays a crucial role. The balance between mitosis and apoptosis is normally carefully regulated, but many types of cancerous cells have evolved to allow

uncontrolled cell division. Hence drugs targeting mitosis are used extensively during cancer chemotherapy. In order to evaluate the effects of a given drug on mitosis, it is desirable to measure average mitosis durations and distribution of possible outcomes such as regular division into two daughter cells, apoptosis, division into an abnormal number of daughter cells (one or  $\geq 3$ ) and no division at all [4,5].

Since performance of technical equipment such as microscopes and associated hardware is constantly improving and large amounts of data can be acquired in very short periods of time, automated image processing tools are frequently favoured over manual analysis, which is expensive and prone to error and bias. Generally, experiments might last several days and images are taken in a magnitude of minutes and from different positions. This leads to a sampling frequency of hundreds of images per sequence with an approximate size of  $1000^2$  pixels.

### 1.1. Image characteristics in phase contrast microscopy

In live-cell imaging experiments for anti-cancer drug assessment, the imaging modality plays a key role. Observation of cell

\* Corresponding author.

E-mail address: [jg704@cam.ac.uk](mailto:jg704@cam.ac.uk) (J.S. Grah).

<sup>1</sup> Present address: Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, MA 02114, USA.

<sup>2</sup> Present address: PerkinElmer, Schnackenburgallee 114, 22525 Hamburg, Germany.

cultures originating from specific cell lines under the microscope requires a particular setting ensuring that the cells do not die during image acquisition and that they behave as naturally as possible [6]. Here, phase contrast is often preferred to fluorescence microscopy because the latter requires labelling or transgenic expression of fluorescent markers, both causing phototoxicity and possibly changes of cell behaviour [7–9]. As opposed to this, cells do not need to be stained for phase contrast microscopy. Moreover, phase shifts facilitate visualisation of even transparent specimens as opposed to highlighting of individual specific cellular components in fluorescence microscopy. We believe that one main advantage of our proposed framework is that it can be applied to data acquired with any standard phase-contrast microscope, which are prevalent in many laboratories and more widespread than for instance recently established quantitative phase imaging devices (e.g. Q-Phase by Tescan).

There are two common image characteristics occurring in phase contrast imaging (cf. Fig. 1). Both visual effects highly impede image processing and standard algorithms are not applicable in a straightforward manner. The shade-off effect leads to similar intensities inside the cells and in the background. As a result, edges are only weakly pronounced and imaging methods such as segmentation relying on intensity gradient information (cf. Section 2.2.2) often fail. Moreover, region-based methods assuming that average intensities of object and background differ from one another (cf. Section 2.2.3) are not applicable either. Secondly, the halo effect is characterised by areas of high intensity surrounding cell membranes. The brightness levels increase significantly immediately before cells enter mitosis due to the fact that they round up, form a nearly spherically-shaped volume and therefore the amount of diffracted light increases. In addition, both effects prohibit application of basic image pre-processing tools like for example thresholding or histogram equalisation (cf. [10]).

### 1.2. Brief literature review

Over the past few years a lot of cell tracking frameworks have been established (cf. [11]) and some publications also feature mitosis detection. In [12], a two-step cell tracking algorithm for phase contrast images is presented, where the second step involves a level-set-based variational method. However, analysis of the mitotic phase is not included in this framework. Another tracking method based on extended mean-shift processes [13] is able to incorporate cell divisions, but does not provide cell membrane segmentation. In [14] an automated mitosis detection algorithm based on a probabilistic model is presented, but it is not linked to cell tracking. A combined mitosis detection and tracking framework is established in [15], although cell outline segmentation is not included. Li et al. [16] provide a comprehensive framework facilitating both tracking and lineage reconstruction of cells in phase contrast image sequences. Moreover, they are able to distinguish between mitotic and apoptotic events.

In addition, a number of commercial software packages for semi- or fully automated analysis of microscopy images exist, for example *Volocity*, *Columbus* (both PerkinElmer), *Imaris* (Bitplane), *ImageJ/Fiji* [17] and *Icy* [18] (also cf. [19]). The last two are open source platforms and the latter supports graphical protocols while the former incorporates a macro language, allowing for individualisation and extension of integrated tools. However, the majority of plugins and software packages are limited to analysis of fluorescence data.

A framework, which significantly influenced development of our methods and served as a basis for our tracking algorithm, was published in 2014 by Möller et al. [20]. It incorporates a MATLAB® Graphical User Interface that enables semi-automated tracking of cells in phase contrast microscopy time-series. The user

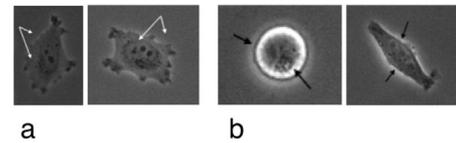


Fig. 1. Common image characteristics in phase contrast microscopy: shade-off effect (a) and halo effect (b) (HeLa DMSO control cells).

has to manually segment the cells of interest in the first frame of the image sequence and can subsequently execute an automatic tracking procedure consisting of two rough and refined segmentation steps. In the following section, the required theoretical foundations of mathematical imaging methods are discussed, starting with the concept of the circular Hough transform and continuing with a review of segmentation and tracking methods leading to a more detailed description of the above-mentioned framework. For a more detailed discussion, we refer the interested reader to [10] and the references therein.

## 2. Mathematical background

### 2.1. The circular Hough transform

The Hough transform is a method for automated straight line recognition in images patented by Paul Hough in 1962 [21]. It was further developed and generalised by Duda and Hart in 1972 [22]. More specifically, they extended the Hough transform to different types of parametrised curves and in particular, they applied it to circle detection.

The common strategy is to transform points lying on straight line segments or curves in the underlying image into a parameter space. Its dimension depends on the number of variables required in order to parametrise the sought-after curve. For the parametric representation of a circle, which can be written as

$$r^2 = (x - c_1)^2 + (y - c_2)^2, \quad (1)$$

the radius  $r$  as well as two centre coordinates  $(c_1, c_2)$  are required. Hence, the corresponding parameter space is three-dimensional. Each point  $(x, y)$  in the original image satisfying the above equation for fixed  $r, c_1$  and  $c_2$  coincides with a cone in the parameter space. Then, edge points of circular objects in the original image correspond to intersecting cones and from detecting those intersections in the parameter space one can again gather circles in the image space.

For simplification, we fix the radius and consider the two-dimensional case in Fig. 2. On the left, we have the image space, i.e. the  $x$ - $y$ -plane, and a circle in light blue with five arbitrary points located on its edge highlighted in dark blue. All points fulfil Eq. (1) for fixed centre coordinates  $(c_1, c_2)$ . On the other hand, fixing those specific values for  $c_1$  and  $c_2$  in the parameter space, i.e.  $c_1$ - $c_2$ -plane, on the right, and keeping  $x$  and  $y$  in (1) arbitrary, leads to the dashed orange circles, where the corresponding edge points are drawn in grey for orientation. All of the orange circles intersect in one point, which exactly corresponds to the circle centre in the original image. Hence, from intersections in the parameter space one can reference back to circular objects in the image space.

A discussion on how the circular Hough transform is embedded and implemented in *MitosisAnalyser* can be found in Section 3.1.

### 2.2. Image segmentation and tracking

In the following, we would like to introduce variational methods (cf. e.g. [23,24]) for imaging problems. The main aim is minimi-

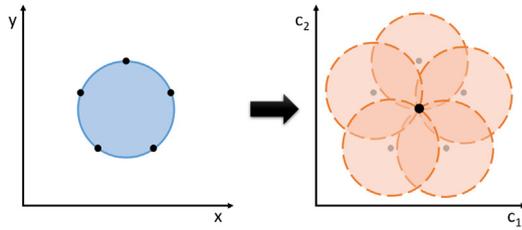


Fig. 2. The circular Hough transform.

sation of an energy functional modelling certain assumptions on the given data and being defined as

$$E(\phi) = D(\phi, \psi) + \alpha R(\phi). \tag{2}$$

It is dependent on the solution  $\phi$ , which represents the processed image to be obtained, and shall be minimised with respect to  $\phi$ . The given image to be processed is denoted by  $\psi$ . The functions  $\phi$  and  $\psi$  map from the rectangular image domain  $\Omega \subset \mathbb{R}^2$  to  $\Sigma \subset \mathbb{R}^d$  containing colour ( $d = 3$ ) or greyscale ( $d = 1$ ) intensity values. In the case of 8-bit phase contrast microscopy images,  $d = 1$  and  $\Sigma = \{0, \dots, 255\}$ , where 0 and 255 correspond to black and white, respectively.

The first part  $D$  on the right-hand side of (2) ensures data fidelity between  $\phi$  and  $\psi$ , i.e. the solution  $\phi$  should be reasonably close to the original input data  $\psi$ . This can be obtained by minimising a norm measuring the distance between  $\psi$  and  $\phi$ , where the choice of norm naturally depends on the given problem. The regulariser  $R$  in (2) incorporates a priori knowledge about the function  $\phi$ . For example,  $\phi$  could be constrained to be sufficiently smooth in a particular sense. The parameter  $\alpha$  is weighting the two different terms and thereby defines which one is considered to be more important. Energy functionals can also consist of multiple data terms and regularisers. Eventually, a solution that minimises the energy functional (2) attains a small value of  $D$  assuring high fidelity to the original data, of course depending on the weighting. Similarly, a solution which has a small value of  $R$  can be interpreted as having a high coincidence with the incorporated prior assumptions.

Here, we focus on image segmentation. The goal is to divide a given image into associated parts, e.g. object(s) and background. This can be done by finding either the objects themselves or the corresponding edges, which is then respectively called region-based and edge-based segmentation. However, those two tasks are very closely related and even coincide in the majority of cases. Tracking can be viewed as an extension of image segmentation because it describes the process of segmenting a sequence of images or video. The goal of object or edge identification remains the same, but the time-dependence is an additional challenge.

Below, we briefly discuss the level-set method and afterwards present two well-established segmentation models incorporating the former. Furthermore, we recap the methods in [20] building upon the above and laying the foundations for our proposed tracking framework.

### 2.2.1. The level-set method

In 1988 the level-set method was introduced by Osher and Sethian [25]. The key idea is to describe motion of a front by means of a time-dependent partial differential equation. In variational segmentation methods, energy minimisation corresponds to propagation of such a front towards object boundaries. In two dimensions, a segmentation curve  $c$  is modelled as the zero-level of a three-dimensional level-set function  $\phi$ . Two benefits are straightforward numerical implementation without need of parametrisation and implicit modelling of topological changes of the curve. The level-set evolution equation can be written as

$$\frac{\partial \phi}{\partial t} = F \cdot |\nabla \phi|$$

with curvature-dependent speed of movement  $F$  and suitable initial and boundary conditions.

For implementation, the level-set function  $\phi$  is assigned negative values inside and positive values outside of the curve  $c$ ,

$$\phi(t, x) \begin{cases} < 0, & \text{if } x \text{ is inside of } c, \\ = 0, & \text{if } x \text{ lies on } c, \\ > 0, & \text{if } x \text{ is outside of } c, \end{cases} \tag{3}$$

commonly chosen to be the signed Euclidean distances (cf. Fig. 3).

### 2.2.2. Geodesic active contours

Active contours or “snakes” have been developed and extended for decades [26–30] and belong to the class of edge-based segmentation methods. As the name suggests, the goal is to move segmentation contours towards image edges and stop at boundaries of objects to be segmented (e.g. by using the level-set method described above). Geodesic active contours constitute a specific type of active contours methods and have been introduced by Caselles, Kimmel and Sapiro in 1997 [31]. The level-set formulation reads

$$\frac{\partial \phi}{\partial t} = \nabla \cdot \underbrace{\left( g \frac{\nabla \phi}{|\nabla \phi|} \right)}_F \cdot |\nabla \phi| \tag{4}$$

with appropriate initial and boundary conditions and  $g$  is an edge-detector function typically depending on the gradient magnitude of a smoothed version of a given image  $\psi$ . A frequently used function is

$$g = \frac{1}{1 + |\nabla(G_\sigma * \psi(x))|^2} \tag{5}$$

with  $G_\sigma$  being a Gaussian kernel with standard deviation  $\sigma$ . The function  $g$  is close to zero at edges, where the gradient magnitude is high, and close or equal to one in homogeneous image regions, where the gradient magnitude is nearly or equal to zero. Hence, the segmentation curve, i.e. the zero-level of  $\phi$ , propagates towards edges defined by  $g$  and once the edges are reached, evolution is stopped. In the specific case of  $g = 1$ , (4) coincides with mean curvature motion.

Geodesic active contours are a well-suited method of choice for segmentation if image edges are strongly pronounced or can otherwise be appropriately identified by a suitable function  $g$ .

### 2.2.3. Active contours without edges

As the name suggests, the renowned model developed by Chan and Vese [32] is a region-based segmentation method and in contrast to the model presented in 2.2.2, edge information is not taken into account. It is rather based on the assumption that the underlying image can be partitioned into two regions of approximately piecewise-constant intensities. In the level-set formulation the variational energy functional reads

$$E(\phi, c_1, c_2) = \lambda_1 \int_{\Omega} (\psi(x) - c_1)^2 (1 - H(\phi(x))) dx + \lambda_2 \int_{\Omega} (\psi(x) - c_2)^2 H(\phi(x)) dx + \mu \int_{\Omega} |\nabla H(\phi(x))| dx + \nu \int_{\Omega} (1 - H(\phi(x))) dx, \tag{6}$$

which is to be minimised with respect to  $\phi$  as well as  $c_1$  and  $c_2$ . Recalling (3), we define the Heaviside function  $H$  as

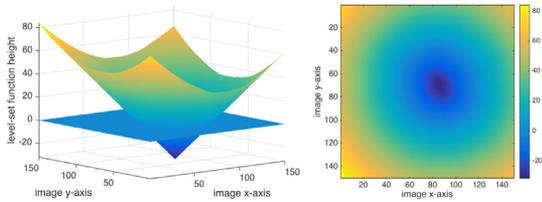


Fig. 3. Level-set function.

$$H(\phi) = \begin{cases} 0, & \text{if } \phi \leq 0, \\ 1, & \text{if } \phi > 0, \end{cases} \quad (7)$$

indicating the sign of the level-set function and therefore the position relative to the segmentation curve.

In (6) the structure in (2) is resembled. The first two data terms enforce a partition into two regions with intensities  $c_1$  inside and  $c_2$  outside of the segmentation contour described by the zero-level-set. The third and fourth terms are contour length and area regularisers, respectively.

The optimal  $c_1$  and  $c_2$  can be directly calculated while keeping  $\phi$  fixed:

$$c_1 = \frac{\int_{\Omega} \psi(x)(1 - H(\phi(x))) dx}{\int_{\Omega} (1 - H(\phi(x))) dx}, \quad c_2 = \frac{\int_{\Omega} \psi(x)H(\phi(x)) dx}{\int_{\Omega} H(\phi(x)) dx}.$$

In order to find the optimal  $\phi$  and hence the sought-after segmentation contour, the Euler-Lagrange equation defined as  $\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi} = 0$  needs to be calculated, which leads to the evolution equation

$$\frac{\partial \phi}{\partial t} = \delta_{\varepsilon}(\phi) \left( \lambda_1 (\psi - c_1)^2 - \lambda_2 (\psi - c_2)^2 + \mu \nabla \cdot \left( \frac{\nabla \phi}{|\nabla \phi|} \right) + v \right), \quad (8)$$

where  $\delta_{\varepsilon}$  is the following regularised version of the Dirac delta function:

$$\delta_{\varepsilon}(\phi) = \frac{\varepsilon}{\pi} (\varepsilon^2 + \phi^2).$$

Eq. (8) can be numerically solved with a gradient descent method.

This model is very advantageous for segmenting noisy images with weakly pronounced or blurry edges as well as objects and clustering structures of different intensities in comparison to the background.

#### 2.2.4. Tracking framework by Möller et al.

The cell tracking framework developed in [20] is sub-divided into two steps. First, a rough segmentation based on the model in Section 2.2.3 is performed. The associated energy functional reads

$$E(\phi, c_1, c_2) = \lambda_1 \int_{\Omega} (|v| - c_1)^2 (1 - H(\phi(x))) dx + \lambda_2 \int_{\Omega} (|v| - c_2)^2 H(\phi(x)) dx + \mu \int_{\Omega} |\nabla H(\phi(x))| dx + v \left( \int_{\Omega} (1 - H(\phi(x))) dx - V_{\text{old}} \right)^2. \quad (9)$$

In contrast to (6), the area or volume regularisation term weighted by  $v$  is altered such that the current volume shall be close to the previous volume  $V_{\text{old}}$ . Moreover, the data terms weighted by  $\lambda_1$  and  $\lambda_2$  incorporate the normal velocity image  $|v|$  instead of the image intensity function  $\psi$ :

$$|v| = \frac{|\frac{\partial \psi}{\partial t}|}{|\nabla \psi|_{\varepsilon}}, \quad (10)$$

where the expression in the denominator is a regularisation of the gradient magnitude defined as  $|\nabla \psi|_{\varepsilon} = \sqrt{(\partial_{x_1} \psi)^2 + (\partial_{x_2} \psi)^2 + \varepsilon^2}$  for small  $\varepsilon$ . The novelty here is that in contrast to only considering the image intensity both spatial and temporal information is used in order to perform the region-based segmentation. Indeed, cells are expected to move between subsequent frames. In addition, the gradient magnitude shall be increased in comparison to background regions. Therefore the incorporation of both temporal and spatial derivative provides a better indicator of cellular interiors.

In a second step, a refinement is performed using the geodesic active contours Eq. (4). The edge-detector function is customised and mainly uses information obtained by the Laplacian of Gaussian of the underlying image. In addition, topology is preserved throughout the segmentation by using the simple points scheme [33–35] and in order to reduce computational costs this is combined with a narrow band method [36], which we inherit in our framework as well.

### 3. MitosisAnalyser framework

In the following we present our proposed workflow designed in order to facilitate mitosis analysis in live-cell phase contrast imaging experiments. We specifically focused on applicability and usability while providing a comprehensive tool that needs minimal user interaction and parameter tuning. The MATLAB® Graphical User Interface *MitosisAnalyser* (The corresponding code is available at [github.com/JoanaGrah/MitosisAnalyser](https://github.com/JoanaGrah/MitosisAnalyser).) provides a user-friendly application, which involves sets of pre-determined parameters for different cell lines and has been designed for non-experts in mathematical imaging.

In Fig. 4 the main application window is displayed on the top left. The entire image sequence at hand can be inspected and after analysis, contours are overlaid for immediate visualisation. Moreover, images can be examined and pre-processed by means of a few basic tools (centre), although the latter did not turn out to be necessary for our types of data. Parameters for both mitosis detection and tracking can be reviewed, adapted and permanently saved for different cell lines in another separate window (bottom left). Mitosis detection can be run separately and produces intermediate results, where all detected cells can be reviewed and parameters can be adjusted as required. Consecutively, running the cell tracking algorithm results in an estimate of average mitosis duration and provides the possibility to survey further statistics (right).

Fig. 5 summarises the entire workflow from image acquisition to evaluation of results. First, live-cell imaging experiments are conducted using light microscopy resulting in 2D greyscale image sequences. Next, mitosis detection is performed. For each detected cell, steps 3–5 are repeated. Starting at the point in time where the cell is most circular, the circle-shaped contour serves as an initialisation for the segmentation. The tracking is then performed backwards in time, using slightly extended contours from previous frames as initialisations. As soon as cell morphology changes, i.e. area increases and circularity decreases below a predetermined threshold, the algorithm stops and marks the point in time at hand as start of mitosis. Subsequently, again starting from the detected mitotic cell, tracking is identically performed forwards in time until the cell fate can be determined. As already mentioned in Section 1, different cases need to be distinguished from one another: regular, abnormal and no division as well as apoptosis. The final step comprises derivation of statistics on mitosis duration and cell fate distribution as well as evaluation and interpretation thereof.

The double arrow connecting steps 1 and 5 indicates what is intended to be subject of future research. Ideally, image analysis

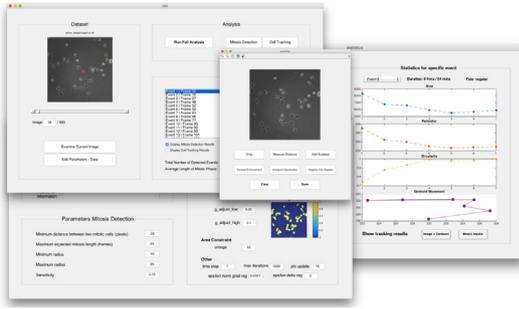


Fig. 4. MitosisAnalyser MATLAB®GUI.



Fig. 5. Summary of MitosisAnalyser framework.

shall be performed in on-line time during image acquisition and intermediate results shall be passed on to inform and influence microscopy software. Consequently, this may in turn lead to enhancement of image processing. Recently established concepts of bilevel optimisation and parameter learning for variational imaging models (cf. [37,38]) might supplement our framework.

### 3.1. Mitosis detection

In order to implement the circular Hough transform (CHT) described in Section 2.1, both image and parameter space need to be discretised. The former is naturally already represented as a pixel grid or matrix of grey values. The latter needs to be artificially discretised by binning values for  $r$ ,  $c_1$  and  $c_2$  and the resulting representation is called accumulator array. Once the CHT is performed for all image pixels, the goal is to find peaks in the accumulator array referring to circular objects.

There are several options in order to speed up the algorithm, but we will only briefly discuss two of them. First, it is common to perform edge detection on the image before applying the CHT, since pixels lying on a circle very likely correspond to edge pixels. An edge map can for instance be calculated by thresholding the gradient magnitude image in order to obtain a binary image. Then, only edge pixels are considered in the following steps. Furthermore, it is possible to reduce the accumulator array to two dimensions using the so-called phase-coding method. The idea is using complex values in the accumulator array with the radius information encoded in the phase of the array entries. Both enhancements are included in the built-in MATLAB®function `imfindcircles`.

The mitosis detection algorithm implemented into *MitosisAnalyser* uses this function in order to perform the CHT and search for circular objects in the given image sequences. Fig. 6 visualises the different steps from calculation of the gradient image, to identification of edge pixels, to computation of the accumulator matrix and transformation thereof by filtering and thresholding, to detection of maxima.

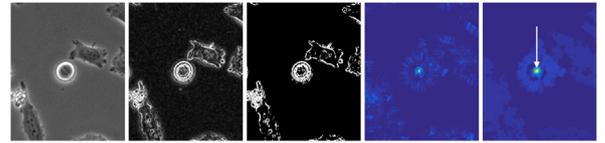


Fig. 6. Finding circles by means of the CHT. From left to right: Original greyscale image, gradient image, edge pixels, accumulator matrix, transformed matrix.

This method turned out to be very robust and two main advantages are that circles of different sizes can be found and even not perfectly circularly shaped or overlapping objects can be detected. At the beginning of analysis, the CHT is applied in every image of the given image sequence in order to detect nearly circularly shaped mitotic cells. Afterwards, the circles are sorted by significance, which is related to the value of the detected peak in the corresponding accumulator array. The most significant ones are picked while simultaneously ensuring that identical cells are neither detected multiple times in the same frame nor in consecutive frames. The complete procedure is outlined in [Supplementary Algorithm 1](#).

### 3.2. Cell tracking

We have already introduced variational segmentation methods in general as well as three models our framework is based on in more detail in Section 2.2. Here, we would like to state the cell tracking model we developed starting from the one presented in Section 2.2.4. The energy functional reads:

$$\begin{aligned}
 E(\phi, c_1, c_2) = & \lambda_1 \int_{\Omega} (|v| - c_1)^2 (1 - H(\phi(x))) dx + \lambda_2 \\
 & \times \int_{\Omega} (|v| - c_2)^2 H(\phi(x)) dx \\
 & + \mu \int_{\Omega} |\nabla H(\phi(x))| dx \\
 & + \nu \int_{\Omega} g(\psi(x)) |\nabla H(\phi(x))| dx - \omega \frac{1}{2} \\
 & \times \max \left\{ \int_{\Omega} (1 - H(\phi(x))) dx - t_{\text{area}}, 0 \right\}^2, \quad (11)
 \end{aligned}$$

with  $|v|$  and  $H$  defined as in (10) and (7), respectively.

The two terms weighted by  $\lambda_1$  and  $\lambda_2$  are identical to the ones in (9). Instead of having two separate segmentation steps as in [20], we integrate the edge-based term weighted by  $\nu$  into our energy functional. However, using a common edge-detector function based on the image gradient like the one in (5) was not suitable for our purposes. We noticed that the gradient magnitude image contains rather weakly pronounced image edges, which motivated us to search for a better indicator of the cells' interiors. We realised that the cells are very inhomogeneous in contrast to the background and consequently, we decided to base the edge-detector function on the local standard deviation of grey values in a  $3 \times 3$ -neighbourhood around each pixel. Additionally smoothing the underlying image with a standard Gaussian filter and rescaling intensity values leads to an edge-detector function, which is able to indicate main edges and attract the segmentation contour towards them.

Furthermore, we add a standard length regularisation term weighted by  $\mu$ . We complement our energy functional with an area regularisation term that incorporates a priori information about the approximate cell area and prevents contours from becoming too small or too large. This penalty method facilitates incorporation of a constraint in the energy functional and in this case the area shall not fall below the threshold  $t_{\text{area}}$ .

Optimal parameters  $c_1$  and  $c_2$  can be calculated directly. We numerically minimise (11) with respect to the level-set function  $\phi$  by using a gradient descent method (cf. 2.2.3). The third term weighted by  $\mu$  is discretised using a combination of forwards, backwards and central finite differences as proposed in [32]. We obtain the most stable numerical results by applying central finite differences to all operators contained in the fourth term weighted by  $\nu$ . In Fig. 7 we visualise level-set evolution throughout the optimisation procedure.

In order to give an overview of the backwards and forwards tracking algorithms incorporated in the mitosis analysis framework, we state the procedures in Supplementary Algorithm 2 and 3. Together with the mitosis detection step they form the foundation of the routines included in *MitosisAnalyser*.

#### 4. Material and methods

The *MitosisAnalyser* framework is tested in three experimental settings with MIA PaCa-2 cells, HeLa Aur A cells and T24 cells. Below, a description of cell lines and chemicals is followed by details on image acquisition and standard pre-processing.

##### 4.1. Cell lines and chemicals

The FUCCI (Fluorescent Ubiquitination-based Cell Cycle Indicator [39])-expressing MIA PaCa-2 cell line was generated using the FastFUCCI reporter system and has previously been characterised and described [40,41]. Cells were cultured in phenol red-free Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum (FBS).

T24 cells were acquired from CLS. The T24 cells were cultured in DMEM/F12 (1:1) medium supplemented with 5% FBS.

HeLa Aur A cells, HeLa cells modified to over-express aurora kinase A, were generated by Dr. Jennifer Harrington with Dr. David Perera at the Medical Research Council Cancer Unit, Cambridge, using the Flp-In T-Rex system from Invitrogen as described before [42]. The parental HeLa LacZeo/TO line, and pOG44 and pcDNA5/FRT/TO plasmids were kindly provided by Professor Stephen Taylor, University of Manchester. The parental line grows under selection with 50  $\mu\text{g}/\text{ml}$  Zeocin<sup>TM</sup> (InvivoGen) and 4  $\mu\text{g}/\text{ml}$  Blasticidin (Invitrogen). HeLa Aur A cells were cultured in DMEM supplemented with 10% FBS and 4  $\mu\text{g}/\text{ml}$  blasticidin (Invitrogen) and 200  $\mu\text{g}/\text{ml}$  hygromycin (Sigma Aldrich). Transgene expression was achieved by treatment with 1  $\mu\text{g}/\text{ml}$  doxycycline (Sigma Aldrich).

In all experiments, all cells were grown at 37 °C and 5% CO<sub>2</sub> up to a maximum of 20 passages and for fewer than 6 months following resuscitation. They were also verified to be mycoplasma-free using the Mycoprobe<sup>®</sup> Mycoplasma Detection Kit (R&D Systems). Paclitaxel (Tocris Bioscience), MLN8237 (Stratex Scientific) and Docetaxel (Sigma Aldrich) were dissolved in dimethylsulphoxide (DMSO, Sigma) in aliquots of 30 mM, kept at -20 °C and used within 3 months. Final DMSO concentrations were kept constant in each experiment ( $\leq 0.2\%$ ).

##### 4.2. Acquisition and processing of live-cell time-lapse sequences

Cells were seeded in  $\mu$ -Slide glass bottom dish (ibidi) and were kept in a humidified chamber under cell culture conditions (37 °C, 5% CO<sub>2</sub>). For experiments with T24 and HeLa Aur A cells they were cultured for 24 h before being treated with drugs or DMSO control. They were then imaged for up to 72 h. Images were taken from three to five fields of view per condition, every 5 min, using a Nikon Eclipse TE2000-E microscope with a 20X (NA 0.45) long-working distance air objective, equipped with a sCMOS Andor Neo camera

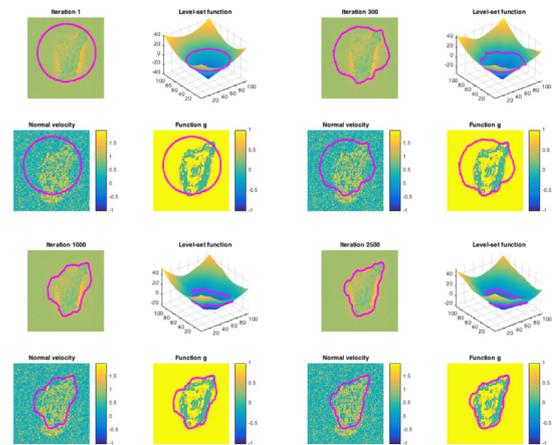


Fig. 7. Level-set evolution from initialisation to final iteration.

acquiring 2048 × 2048 images, which have been binned by a factor of two. Red and green fluorescence of the FUCCI-expressing cells were captured using a pE-300white COOLED source of light filtered by Nikon FITC B-2E/C and TRITC G-2E/C filter cubes, respectively. For processing, an equalisation of intensities over time was applied to each channel, followed by a shading correction and a background subtraction, using the NIS-Elements software (Nikon).

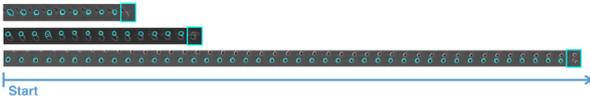
#### 5. Results and discussion

In this section we present and discuss results obtained by applying *MitosisAnalyser* to the aforementioned experimental live-cell imaging data. A list of parameters we chose can be found in Supplementary Table 1. For each cell line, we established a unique set of parameters. Nevertheless, the individual values are in reasonable ranges and do not differ significantly from one another. We did not follow a specific parameter choice rule, but rather tested various combinations and manually picked the best performing ones.

##### 5.1. MIA PaCa-2 cells

In a multi-modal experiment with FUCCI-expressing MIA PaCa-2 cells, both phase contrast images and fluorescence data were acquired. The latter consist of two channels with red and green intensities corresponding to CDT1 and Geminin signals, respectively. In this case we do use fluorescence microscopy imaging data as well, but we would like to stress that this analysis would not have been possible without the mitosis detection and tracking performed on the phase contrast data. As before, mitotic cells are detected using the circular Hough transform applied to the phase contrast images. Cell tracking is performed on the phase contrast images as well, but in addition, information provided by the green fluorescent data channel is used. More specifically, stopping criteria for both backwards and forwards tracking are based on green fluorescent intensity distributions indicating different stages of the cell cycle, which can be observed and is described in more detail in Supplementary Fig. 1.

The whole data set consists of nine imaging positions, where three at a time correspond to DMSO control, treatment with 3nM paclitaxel and treatment with 30nM paclitaxel. Fig. 8 visualises exemplary courses of the mitotic phase, which could be measured by means of our proposed workflow. Table 1 presents estimated average mitosis durations for the three different classes of data. Indeed, the average duration of 51 min for the control is consistent with that obtained from manual scoring (cf. [41], Figure S3D). Moreover, we can observe a dose-dependent increase in mitotic



**Fig. 8.** Three examples of mitotic events detected for FUCCI MIA PaCa-2 “DMSO control”, “treatment with 3 nM paclitaxel” and “treatment with 30 nM paclitaxel” data (from top to bottom).

duration for the two treatments, which was anticipated, since paclitaxel leads to mitotic arrest.

5.2. HeLa cells

In the following we discuss results achieved by applying *MitosisAnalyser* to sequences of phase contrast microscopy images showing HeLa Aur A cells. In addition to DMSO control data, cells have been treated with 25 nM MLN8237 (MLN), 0.75 nM paclitaxel (P), 30 nM paclitaxel (P) and with a combination of 25 nM MLN8237 and 0.75 nM paclitaxel (combined).

Fig. 9 shows exemplary results for detected and tracked mitotic events, where DMSO control cells divide regularly into two daughter cells. Particular treatments are expected to enhance multipolar mitosis and indeed our framework was able to depict the three daughter cells in each of the three examples (bottom rows) presented. In addition, mitosis duration is extended, as anticipated, for treated cells and specifically for the combined treatment. The segmentation of the cell membranes seems to work well by visual inspection, even in the case of touching neighbouring cells.

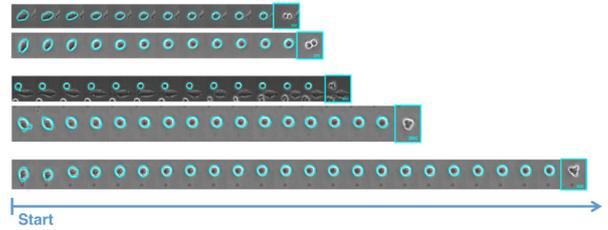
Table 2 summarises average mitosis durations that have been estimated for the different treatments. Again, the results are according to our expectations, i.e. mitosis durations for treated cells are extended in comparison to DMSO control.

5.3. T24 cells

For this data set we wanted to focus on cell fate determination and in order to distinguish between different fates in the T24 cell data set we combine the *MitosisAnalyser* framework with basic classification techniques. In particular, we manually segmented three different classes of cells: mitotic and apoptotic ones as well as cells in their normal state outside of the mitotic cell cycle phase (see Fig. 10).

In Fig. 11 we show boxplots of nine features based on morphology as well as intensity values we use for classification. Those include area, perimeter and circularity. Furthermore, we calculate both mean and standard deviation of the histogram. In addition, we consider the maximum of the gradient magnitude, the mean as well as the total variation of the local standard deviation and the total variation of the grey values. One can clearly observe that cells in mitosis have much higher circularity than in any other state. Flat cells differ significantly from the other two classes with respect to features based on intensity values.

In order to train a classifier solely based on those few features we used the MATLAB®Machine Learning Toolbox and its accompanying Classification Learner App. We chose a nearest-neighbour classifier with the number of neighbours set to 1 using Euclidean



**Fig. 9.** Five examples of mitotic events detected for HeLa Aur A “DMSO control” (one each in row one and two), “treatment with 25 nM MLN8237” (one each in row three and four), and “combined treatment with 25 nM MLN8237 and 0.75 nM paclitaxel” (bottom row) data.

distances and equal distance weights, which yielded a classification accuracy of 93.3% (cf. Supplementary Fig. 2).

Pie charts for T24 cell fate distributions for different drug treatments as preliminary results can be found in Supplementary Fig. 3, although integration of classification techniques will be subject of more extensive future research.

5.4. Validation

In order to validate performance of the segmentation, we compare results obtained with *MitosisAnalyser* with blind manual segmentation. For that purpose, we choose two different error measures: The Jaccard Similarity Coefficient (JSC) [43] and the Modified Hausdorff Distance (MHD) [44], which we are going to define in the following.

Let A and M be the sets of pixels included in the automated and manual segmentation mask, respectively. The JSC is defined as

$$JSC(A, M) = \frac{|A \cap M|}{|A \cup M|},$$

where  $A \cap M$  denotes the intersection of sets A and M, which contains pixels that are elements of both A and M. The union of sets A and M, denoted by  $A \cup M$ , contains pixels that are elements of A or M, i.e. elements either only of A or only of M or of  $A \cap M$ . The MHD is a generalisation of the Hausdorff distance, which is commonly used to measure distance between shapes. It is defined as

$$MHD(A, M) = \max \left\{ \frac{1}{|A|} \sum_{a \in A} d(a, M), \frac{1}{|M|} \sum_{m \in M} d(m, A) \right\},$$

where  $d(a, M) = \min_{m \in M} \|a - m\|$  with Euclidean distance  $\|\cdot\|$ .

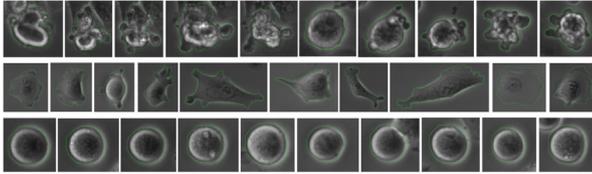
The JSC assumes values between 0 and 1 and the closer it is to 1 the better is the segmentation quality. The MHD on the other hand is equal to 0 if two shapes coincide and the larger the number, the farther they differ from each other. In Fig. 12 and Supplementary Table 2 we can observe that on average, *MitosisAnalyser* performs better than the standard Chan-Vese method (cf. Section 2.2.3) and Geodesic Active Contours based on the gradient magnitude (cf. Section 2.2.2) (both performed using the MATLAB *imageSegmenter* application) compared to manual segmentation of ten apoptotic T24 cell images (cf. Fig. 10., top row). Moreover, Fig. 13 shows successful segmentation of flat T24 cells affected by the shade-off effect in phase contrast microscopy images using *Mito-*

**Table 1**  
Average Mitosis Durations (AMD) for MIA PaCa-2 cell line in minutes.

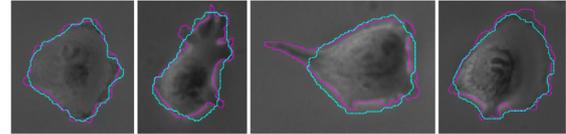
	DMSO control			3nM paclitaxel			30nM paclitaxel		
	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5	Pos 6	Pos 7	Pos 8	Pos 9
Events	14	11	13	12	8	19	10	13	35
AMD	51	41	60	52	88	94	146	104	112
Total AMD		51			78			121	

**Table 2**  
Average Mitosis Durations (AMD) for HeLa cell line in minutes.

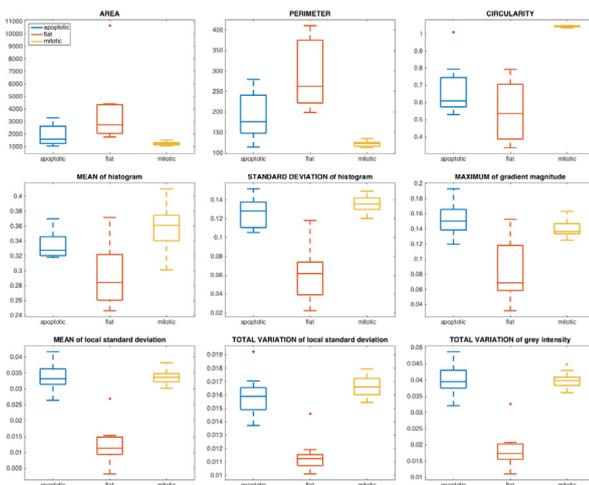
	DMSO	25 nM MLN	0.75 nM P	30 nM P	Combined
Events	44	75	10	35	43
AMD	58	73	68	116	105



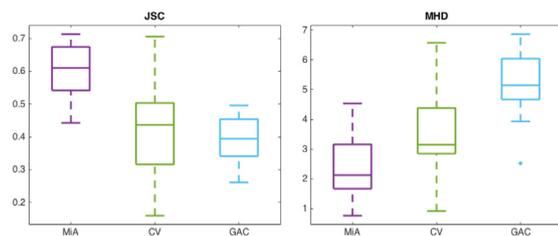
**Fig. 10.** Three manually segmented classes of T24 cells: apoptotic (top row), flat/normal (middle row) and mitotic (bottom row).



**Fig. 13.** Exemplary segmentations for flat cells in phase contrast images: Manual segmentation (magenta) is compared to performance of *MitosisAnalyser* (cyan). The average JSC and MHD values for the four images are 0.8377 and 0.3648, respectively.



**Fig. 11.** Key features for cell type classification.



**Fig. 12.** Boxplots showing JSC (left) and MHD (right) measures for segmentation of apoptotic cell images by *MitosisAnalyser* (MiA), the model by Chan and Vese (CV) and geodesic active contours (GAC) in comparison with manual segmentation.

*sisAnalyser*, where both the method by Chan and Vese and geodesic active contours failed.

### 5.5. Conclusions

We have used concepts of mathematical imaging including the circular Hough transform and variational tracking methods in order to develop a framework that aims at detecting mitotic events and segmenting cells in phase contrast microscopy images, whilst overcoming the difficulties associated with those images. Originating from the models presented in Section 2, we developed a customised workflow for mitosis analysis in live-cell imaging experiments performed in cancer research and discussed results we obtained by applying our methods to different cell line data.

### Acknowledgements

JSG acknowledges support by the NIHR Cambridge Biomedical Research Centre and would like to thank Hendrik Dirks, Fjedor Gaede [45] and Jonas Geiping [46] for fruitful discussions in the context of a practical course at WWU Münster in 2014 and significant speed-up and GPU implementation of earlier versions of the code. JSG and MB would like to thank Michael Möller for providing the basic tracking code and acknowledge support by ERC via Grant EU FP 7 – ERC Consolidator Grant 615216 LifeInverse. MB acknowledges further support by the German Science Foundation DFG via Cells-in-Motion Cluster of Excellence. CBS acknowledges support from the EPSRC grant Nr. EP/M00483X/1, from the Leverhulme grant “Breaking the non-convexity barrier”, from the EPSRC Centre for Mathematical And Statistical Analysis Of Multimodal Clinical Imaging grant Nr. EP/N014588/1, and the Cantab Capital Institute for the Mathematics of Information. JAH, SBK, JAP, AS and SR were funded by Cancer Research UK, The University of Cambridge and Hutchison Whampoa Ltd. SBK also received funding from Pancreatic Cancer UK.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ymeth.2017.02.001>.

### References

- [1] J. Rittscher, Characterization of biological processes through automated image analysis, *Ann. Rev. Biomed. Eng.* 12 (2010) 315–344.
- [2] C.H. Topham, S.S. Taylor, Mitosis and apoptosis: how is the balance set?, *Curr Opin. Cell Biol.* 25 (6) (2013) 780–785.
- [3] K.E. Gascoigne, S.S. Taylor, Cancer cells display profound intra-and interline variation following prolonged exposure to antimetabolic drugs, *Cancer cell* 14 (2) (2008) 111–122.
- [4] C.L. Rieder, H. Maiato, Stuck in division or passing through: what happens when cells cannot satisfy the spindle assembly checkpoint, *Developmental cell* 7 (5) (2004) 637–651.
- [5] B.A. Weaver, D.W. Cleveland, Decoding the links between mitosis, cancer, and chemotherapy: the mitotic checkpoint, adaptation, and cell death, *Cancer cell* 8 (1) (2005) 7–12.
- [6] D.J. Stephens, V.J. Allan, Light microscopy techniques for live cell imaging, *Science* 300 (5616) (2003) 82–86.
- [7] R. Dixit, R. Cyr, Cell damage and reactive oxygen species production induced by fluorescence microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy, *Plant J.* 36 (2) (2003) 280–290.
- [8] J.W. Dobrucki, D. Feret, A. Noatynska, Scattering of exciting light by live cells in fluorescence confocal imaging: phototoxic effects and relevance for frap studies, *Biophys. J.* 93 (5) (2007) 1778–1786.
- [9] R.M. Lasarow, R.R. Isseroff, E.C. Gomez, Quantitative in vitro assessment of phototoxicity by a fibroblast-neutral red assay, *J. Invest. Dermatol.* 98 (5) (1992) 725–729.

- [10] J.S. Grah, Methods for automatic mitosis detection and tracking in phase contrast images, Master's thesis, WWU – University of Münster, 2014.
- [11] E. Meijering, O. Dzyubachyk, I. Smal, W.A. van Cappellen, Tracking in cell and developmental biology, *Seminars in Cell & Developmental Biology*, vol. 20, Elsevier, 2009, pp. 894–902.
- [12] M.E. Ambühl, C. Brepant, J.-J. Meister, A.B. Verkhovskiy, I.F. Sbalzarini, High-resolution cell outline segmentation and tracking from phase-contrast microscopy images, *J. Microsc.* 245 (2) (2012) 161–170.
- [13] O. Debeir, P. Van Ham, R. Kiss, C. Decaestecker, Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes, *IEEE Trans. Med. Imaging* 24 (6) (2005) 697–711.
- [14] S. Huh, D.F.E. Ker, R. Bise, M. Chen, T. Kanade, Automated mitosis detection of stem cell populations in phase-contrast microscopy images, *IEEE Trans. Med. Imaging* 30 (3) (2011) 586–596.
- [15] K. Thirusittampalam, J. Hossain, P.F. Whelan, A novel framework for cellular tracking and mitosis detection in dense phase contrast microscopy images, *IEEE Trans. Biomed. Eng.* 17 (3) (2013) 642–653.
- [16] K. Li, E.D. Miller, M. Chen, T. Kanade, L.E. Weiss, P.G. Campbell, Cell population tracking and lineage construction with spatiotemporal context, *Med. Image Anal.* 12 (5) (2008) 546–566.
- [17] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (7) (2012) 676–682.
- [18] F. De Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, et al., Icy: an open bioimage informatics platform for extended reproducible research, *Nat. Methods* 9 (7) (2012) 690–696.
- [19] K.W. Eliceiri, M.R. Berthold, I.G. Goldberg, L. Ibáñez, B.S. Manjunath, M.E. Martone, R.F. Murphy, H. Peng, A.L. Plant, B. Roysam, et al., Biological imaging software tools, *Nat. Methods* 9 (7) (2012) 697–710.
- [20] M. Möller, M. Burger, P. Dieterich, A. Schwab, A framework for automated cell tracking in phase contrast microscopic videos based on normal velocities, *J. Visual Commun. Image Represent.* 25 (2) (2014) 396–409.
- [21] P.V.C. Hough, Method and means for recognizing complex patterns, *US Patent 3,069,654*, Dec. 18 1962.
- [22] R.O. Duda, P.E. Hart, Use of the hough transformation to detect lines and curves in pictures, *Commun. ACM* 15 (1) (1972) 11–15.
- [23] G. Aubert, P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, Springer, 2006.
- [24] T.F. Chan, J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*, *Siam* (2005).
- [25] S. Osher, J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations, *J. Comput. Phys.* 79 (1) (1988) 12–49.
- [26] V. Caselles, F. Catté, T. Coll, F. Dibos, A geometric model for active contours in image processing, *Numerische mathematik* 66 (1) (1993) 1–31.
- [27] L.D. Cohen, On active contour models and balloons, *CVGIP: Image Understanding* 53 (2) (1991) 211–218.
- [28] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, A. Yezzi, Gradient flows and geometric active contour models, *Proceedings Fifth International Conference on Computer Vision*, IEEE, 1995, pp. 810–815.
- [29] N. Paragios, R. Deriche, Geodesic active contours and level sets for the detection and tracking of moving objects, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (3) (2000) 266–280.
- [30] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, *Int. J. Comput. Vision* 1 (4) (1988) 321–331.
- [31] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, *Int. J. Comput. Vision* 22 (1) (1997) 61–79.
- [32] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Trans. Image Process.* 10 (2) (2001) 266–277.
- [33] G. Bertrand, Simple points, topological numbers and geodesic neighborhoods in cubic grids, *Pattern Recognit. Lett.* 15 (10) (1994) 1003–1011.
- [34] X. Han, C. Xu, J.L. Prince, A topology preserving level set method for geometric deformable models, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (6) (2003) 755–768.
- [35] T.Y. Kong, A. Rosenfeld, Digital topology: introduction and survey, *Comput. Vision Graphics Image Process.* 48 (3) (1989) 357–393.
- [36] D. Adalsteinsson, J.A. Sethian, A fast level set method for propagating interfaces, *J. Comput. Phys.* 118 (2) (1995) 269–277.
- [37] L. Calatroni, C. Chung, J.C.D.L. Reyes, C.-B. Schönlieb, T. Valkonen, Bilevel approaches for learning of variational imaging models, *arXiv preprint arXiv:1505.02120*.
- [38] K. Kunisch, T. Pock, A bilevel optimization approach for parameter learning in variational models, *SIAM J. Imaging Sci.* 6 (2) (2013) 938–983.
- [39] A. Sakaue-Sawano, H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, et al., Visualizing spatiotemporal dynamics of multicellular cell-cycle progression, *Cell* 132 (3) (2008) 487–498.
- [40] S.-B. Koh, P. Mascalchi, E. Rodriguez, Y. Lin, D.I. Jodrell, F.M. Richards, S.K. Lyons, A quantitative fastfucci assay defines cell cycle dynamics at a single-cell level, *J. Cell Sci.* 130 (2) (2017) 512–520.
- [41] S.-B. Koh, A. Courtin, R.J. Boyce, R.G. Boyle, F.M. Richards, D.I. Jodrell, Chk1 inhibition synergizes with gemcitabine initially by destabilizing the dna replication apparatus, *Cancer Res.* 75 (17) (2015) 3583–3595.
- [42] A. Tighe, V.L. Johnson, S.S. Taylor, Truncating apc mutations have dominant effects on proliferation, spindle checkpoint control, survival and chromosome stability, *J. Cell Sci.* 117 (26) (2004) 6339–6353.
- [43] P. Jaccard, *La distribution de la flore dans la zone alpine*, 1907.
- [44] M.-P. Dubuisson, A.K. Jain, A modified hausdorff distance for object matching, *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, 1994. Vol. 1-Conference A: Computer Vision & Image Processing, vol. 1, IEEE, 1994, pp. 566–568.
- [45] F. Gaede, Segmentation and tracking of cells in complete image sequences, WWU – University of Münster, 2015. Bachelor's thesis.
- [46] J. Geiping, Comparison of topology-preserving segmentation methods and application to mitotic cell tracking, WWU – University of Münster, 2014. Bachelor's thesis.



# Multiple objects detection in biological images using a marked point process framework



Xavier Descombes

Université Côte d'Azur, INRIA CRI-SAM, I3S, iBV, France

## ARTICLE INFO

### Article history:

Received 20 July 2016

Received in revised form 14 September 2016

Accepted 19 September 2016

Available online 21 September 2016

### Keywords:

Marked point process

Object detection

Biological imagery

## ABSTRACT

The marked point process framework has been successfully developed in the field of image analysis to detect a configuration of predefined objects. The goal of this paper is to show how it can be particularly applied to biological imagery. We present a simple model that shows how some of the challenges specific to biological data are well addressed by the methodology. We further describe an extension to this first model to address other challenges due, for example, to the shape variability in biological material. We finally show results that illustrate the MPP framework using the “simcep” algorithm for simulating populations of cells.

© 2016 Elsevier Inc. All rights reserved.

## Contents

1. Introduction	2
2. Method	3
2.1. Marked Point Process	3
2.2. Optimization	4
2.3. Results	4
3. Discussion	5
3.1. More general shapes	6
3.2. Energy function	7
3.3. Computational considerations	7
3.4. Further MPP applications to biology imagery	8
4. Conclusion	8
Acknowledgments	8
References	8

## 1. Introduction

Detecting multiple instances of a given object from images is a major issue in computer vision as it often represents the first step towards image understanding and interpretation. For example, in remote sensing, the description of land cover (especially when dealing with high resolution images) relies on a previous detection of objects in the scene such as buildings, trees or roads. In computational biology this problem also appears frequently in order to

evaluate, characterize or classify a population of biological objects such as cells, vesicles within cells or RNA/protein complexes [1,2]. A particular case can be the initialization of a tracking algorithm to study, for example, vesicles trajectories [3]. In addressing biological applications some specific issues have to be considered due to the variability of biological material within and between different classes of objects. For example, objects representing other biological material may be mixed with the actually targeted ones, thus the image cannot be simply modeled as a collection of objects of interest in a background. Besides, the size of these targeted objects is sometimes close to the voxel size, making the differentiation between objects and noise particularly complicated. In this paper,

E-mail address: [xavier.descombes@inria.fr](mailto:xavier.descombes@inria.fr)

we present a methodological framework that provides tools to solve the different issues raised by multiple biological objects detection from microscopic images. We will particularly develop the following:

- Issue 1: How to address the intensity heterogeneity that prevents from considering a global threshold on the intensity in order to separate to objects from background ?
- Issue 2: How to deal with nuisance objects that do not belong to the targeted class of objects but cannot be considered as background neither ?
- Issue 3: How to deal with a high density of objects that generates clusters of possibly overlapping objects ?
- Issue 4: How to handle the shape variability between objects ?
- Issue 5: How to detect objects that consist of a few pixels ?
- Issue 6: How to deal with both 2D and 3D datasets ?

Throughout the literature that addresses this problem, we distinguish both global as well as local methods. Global methods usually consider a threshold to separate the background from pixels belonging to objects. Each n-connected group of pixels tagged as object is then analyzed. A watershed segmentation is then performed on the distance map inside each component to split it into individual objects. Each individual object is finally selected or rejected depending on its size and shape, considering for example a circularity parameter. This classical approach is usually the one proposed by common image analysis software such as Matlab, the particle analyzer of Fiji or Cell Profiler [4,5]. Nevertheless, issue 1 is not addressed within this approach. In consequence, in order to remove background variation, a high pass filter has to be previously applied. Issues 2,3 and 4 are partially solved if the objects of interest have more or less a circular shape and can be bounded by particular minimum and maximum sizes that discriminate them from nuisance objects. The shape of the detected object is arbitrarily defined by the watershed algorithm, so issue 4 is not addressed. Finally, issue 5 is not addressed in case of noisy data. In local approaches, a first step usually consists of seeds detection. A growing process then extends each seed to define an object using, for example an active contour or marker controlled watershed. This process allows the object shape recovery only if they are initially properly localized by the seeds. Therefore, the seeds detection is crucial. Some strategies to obtain these seeds include local maxima after a global threshold or a template matching process [6]. Issue 1 can be partially solved by considering a low threshold when seeds are defined by local maxima. Issue 2 is not addressed whereas clusters are split arbitrarily when two growing objects intersect.

In this paper we present the marked point process modeling (MPP) as a framework to solve the different issues described above. These models derived from the application of point processes to spatial statistics. They have proven their efficiency and robustness in various fields of computer vision in order to evaluate populations of, for example, trees, buildings, roads, people in a crowd or flamingos. A survey of marked point processes applied to image analysis can be found in [7]. Herein we focus on biological images and show how to derive specific models to accurately address the different issues mentioned above.

## 2. Method

### 2.1. Marked Point Process

Let us consider an object space  $\mathcal{O} \subset \mathbb{R}^m$  that contains the geometrical description of the object of interest. For example if we consider the set of disks with radius bounded by  $r_{min}$  and  $r_{max}$ , then  $\mathcal{O} = [r_{min}, r_{max}] \subset \mathbb{R}$ .

We consider the configuration set  $\Omega$  as the union of all the possible finite sets of objects lying in a subspace  $\mathcal{S}$  of  $\mathbb{R}^m$  defined by the support of the image:

$$\Omega = \bigcup_{i=0}^{\infty} \Omega_i, \quad (1)$$

where

$$\Omega_i = \{\omega_1, \dots, \omega_i\} \in (\mathcal{S} \times \mathcal{O})^i \quad (2)$$

is the set of configurations containing exactly  $i$  objects,  $\omega_i = (p_i, m_i)$ ,  $p_i \in \mathcal{S}$  is the center of the object and  $m_i \in \mathcal{O}$  are the marks. We define a marked point process [8] by the Gibbs density as follows:

$$\forall \omega \in \Omega, d\pi(\omega) = \frac{1}{Z} \exp[-U(\omega)] d\pi_0(\omega), \quad (3)$$

where  $\pi_0$  is the measure of the Poisson process and  $U(\omega)$  is the energy function that evaluates each configuration of objects. The lower the energy function value the more probable is the particular object configuration. In the context of image analysis, the energy function embeds a data term,  $U_D(\Omega|I)$ , that evaluates the consistency of any object with respect to the data  $I$  as well as a prior,  $U_P(\Omega)$ , that reflects constraints on the objects geometry and repartition in the image plane.

Let us consider a first example, shown in Fig. 1, where the image  $\{I(s), s \in L\}$  on the lattice  $L$  consists of circular cells on a dark background. We first define a data term that measures the contrast between a candidate object and its neighborhood as follows:

$$P(I|\Omega = \{\omega_1, \dots, \omega_i, \dots, \omega_n\}) = \exp -U_D(\Omega|I) \text{ with} \quad (4)$$

$$U_D(\Omega|I) = \sum_{i=1}^n u_d(\omega_i),$$

where  $u_d(\omega_i)$  is a contrast term we defined as:

$$u_d(\omega_i) = \begin{cases} 1 - \frac{d(\omega_i)}{d_0} & \text{if } d(\omega_i) < d_0 \\ \exp\left(\frac{d_0 - d(\omega_i)}{3d_0}\right) - 1 & \text{otherwise.} \end{cases} \quad (5)$$

In Eq. (5),  $d(\omega_i)$  is a distance between pixels in the object  $\omega_i$  and pixels in the external boundary  $\partial\omega_i$  (see Fig. 2). For example the Bhattacharyya distance is defined by:

$$d(\omega) = \frac{1}{4} \frac{(\mu_o - \mu_b)^2}{\sigma_o^2 + \sigma_b^2} + \frac{1}{2} \log \left[ \frac{\sigma_o^2 + \sigma_b^2}{2\sigma_o\sigma_b} \right], \quad (6)$$

where  $\mu_o$  (resp.  $\mu_b$ ) and  $\sigma_o^2$  (resp.  $\sigma_b^2$ ) are the mean and variance of pixels in  $\omega$  (resp.  $\partial\omega$ ).

In order to prevent object overlap as much as possible, we add the following prior:

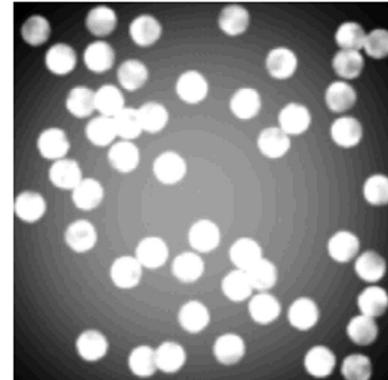


Fig. 1. Example of an image containing a collection of objects on a background.

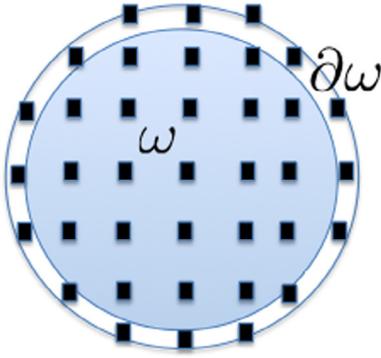


Fig. 2. Discretization of a disk  $\omega$  and its neighborhood  $\partial\omega$ .

$$U_p(\Omega) = \sum_{i,j:\omega_i \cap \omega_j \neq \emptyset} f(\omega_i, \omega_j) \quad \text{with} \quad (7)$$

$$f(\omega_i, \omega_j) = \begin{cases} \infty & \text{if } \frac{|\omega_i \cap \omega_j|}{\min(|\omega_i|, |\omega_j|)} > o \\ 0 & \text{otherwise} \end{cases}$$

where  $|\omega_i|$  refers to the size of object  $i$  and  $o$  is the maximum overlap ratio permitted. The solution is then defined as the minimizer of the global energy:

$$U(\Omega) = U_p(\Omega) + U_D(\Omega|I). \quad (8)$$

## 2.2. Optimization

Two practical issues arise when analyzing this problem. On the first place, the energy  $U(\Omega)$  we want to minimize is not convex. Secondly, the normalizing constant (or partition function)  $Z$  defined in Eq. (3) is analytically and numerically intractable. Therefore, the optimization is classically performed using a simulated annealing coupled with a sampling algorithm. This last that can be, for example, jump and diffusion processes or MCMC approaches, particularly Reversible Jump MCMC (RJMCMC), as the number of objects is unknown. More recently, the multiple births and deaths (MBD) algorithm based on a discretization scheme of a stochastic differential equation has been proposed [9]. As an advantage, this algorithm permits to address a whole set of objects in the same iteration. Besides, there is no rejection in the birth step that allows any new objects introduction at every stage of the simulated annealing. Some faster suboptimal algorithms have also been proposed such as the multiple births and

cut (MBC). In this paper we consider the MBD algorithm as the MBC algorithm is restricted to particular energy functions.

The MBC algorithm alternates births and deaths steps consisting in adding new objects and removing some of them with a certain probability that depends on the specific value of the energy function:

### Algorithm 1 Multiple Births and Deaths

1. Initialize the objects configuration with the empty set  $\Omega_0 = \emptyset$ , set  $T = T_0$ ,  $i = 0$  and  $\delta = \delta_0$
2. *Births step*: Set  $i = i + 1$ , Generate randomly a set of objects  $B_i = \{\omega_i^j\}$  and compute the data term for each object  $u_d(\omega_i^j)$ . The location and the marks of objects are drawn from a uniform distribution and the number of objects is drawn from a Poisson law of parameter  $\delta \times |L|$ ,  $|L|$  being the number of pixels. Set  $\Omega_i^b = \Omega_{i-1} \cup B_i$ .
3. *Sorting step*: Sort the objects in  $\Omega_i^b$  by descending order of the data energy (from the “worst” to the “best”).
4. *Death step*: For each object  $\{\omega_j \in \Omega_i^b\}$  taken sequentially in the ordered list, remove  $\omega_j$  from  $\Omega_i^b$  with probability  $p = \frac{\delta a(\omega_j, \Omega_i^b)}{1 + \delta a(\omega_j, \Omega_i^b)}$  such that

$$a(\omega_j, \Omega_i^b) = \exp -\frac{1}{T} (U(\Omega_i^b / \{\omega_j\}) - U(\Omega_i^b))$$

5. If not converged decrease  $T$  (resp.  $\delta$ ) by a factor  $\alpha_T$  (resp.  $\alpha_\delta$ ) and go back to step 2.

The convergence has been theoretically proved in [17], and is empirically obtained either after a fix number of iterations or when the configuration does not change during a couple of iterations.

## 2.3. Results

We validate this model on several synthetic images of cells simulated with the “Simcep” algorithm [10]. To do so, we compare the performance of the proposed MPP approach with the classical approach proposed by software such as Fiji or Matlab. This last one consists in binarizing the image and splitting the clusters using the watershed algorithm on the distance map. Resulting connected components are selected as objects depending on their size and a circularity coefficient. We first consider a noise free image of cells given on Fig. 3, and then we add noise on Fig. 4. Finally we increase the background heterogeneity due to the light source (see Figs. 5 and 6).

With this first model we partially address the different challenges (issues 1, 2 and 3). The main advantage of this approach is

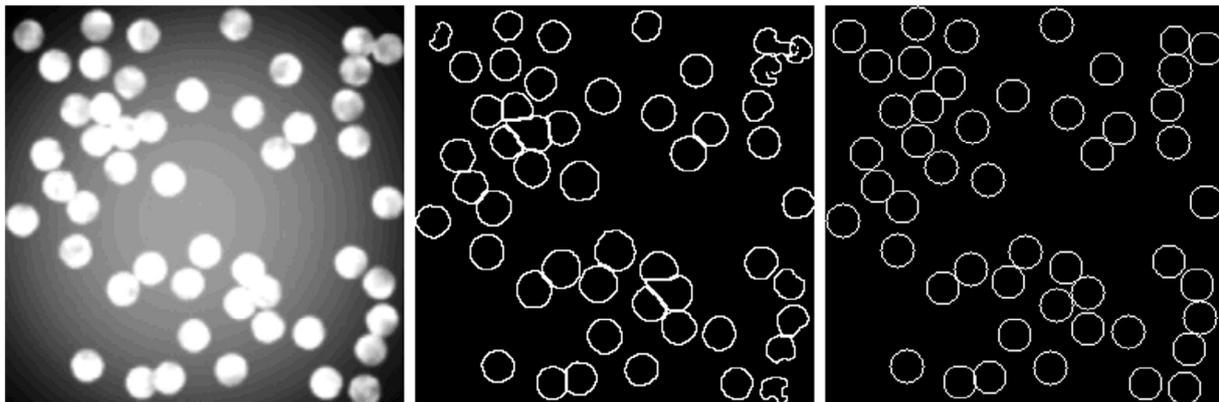
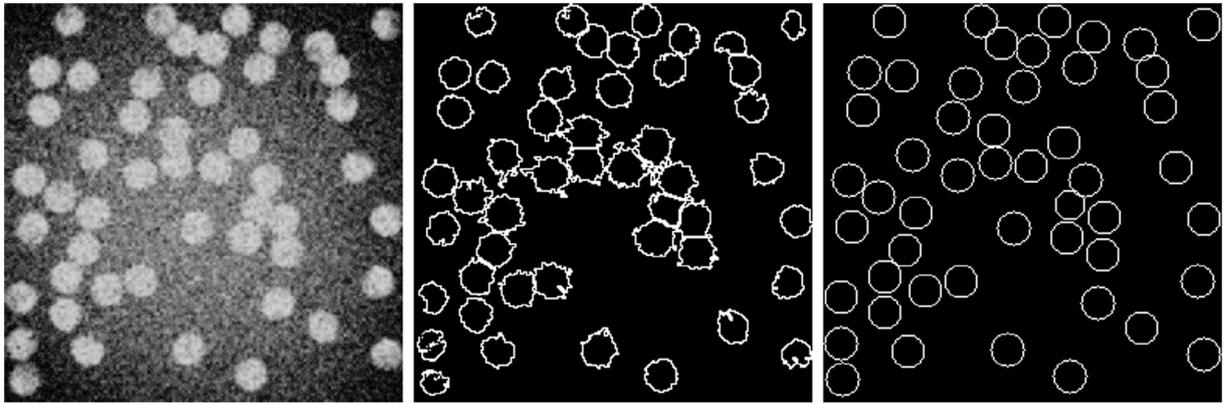
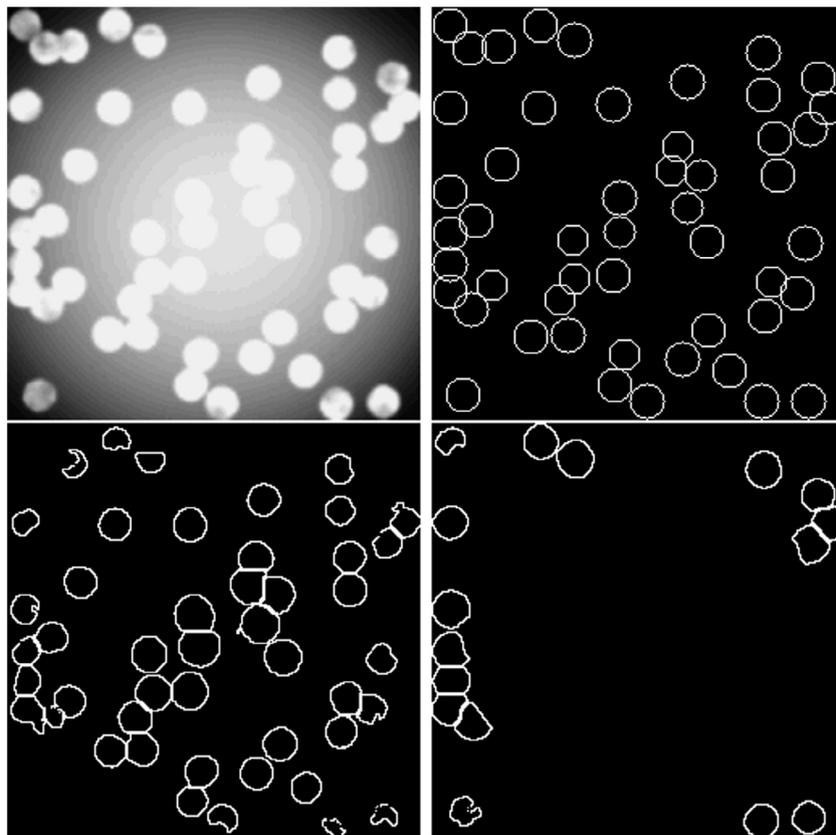


Fig. 3. A noise free example of image cells (left) and the detection obtained using the fiji particle analyser (middle) and the MPP approach (right).



**Fig. 4.** A noisy example of image cells (left) and the detection obtained using the fiji particle analyser (middle) and the MPP approach (right).



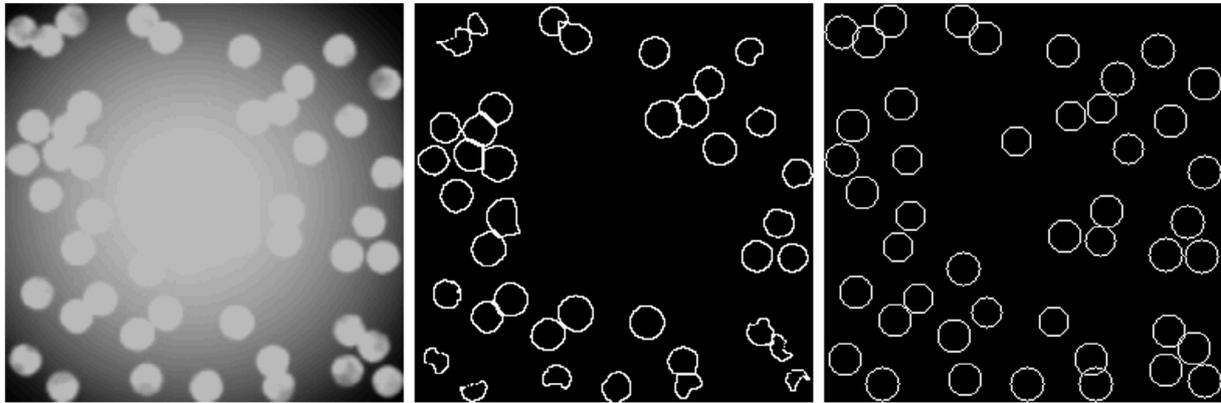
**Fig. 5.** A first example of an heterogeneous background (top left), results with the MPP approach (top right) and with two parameter settings for the fiji particle analyser (bottom left and right).

that the data is taken into account at the object level. In the data term, statistics of pixels contained in the whole object are considered providing a high robustness with respect to noise (see Fig. 4). Besides, we consider a local contrast term between the object and the surrounding pixels that gives robustness with respect to the background heterogeneity (see Figs. 5 and 6) compared to the classic approach. Finally, the shape model facilitates the discrimination between objects of interest and nuisance objects. Notice however that in the case of disk shaped objects the watershed algorithm (classic approach) performs quite well on the task of splitting clusters (see Fig. 3). However, this performance decreases rapidly with noise or background heterogeneity while the MPP approach remains robust. Finally, with this simple model, only circular cells

are addressed. Several extensions have to be considered in order to generalize the approach.

### 3. Discussion

We have shown in Section 2.3 that the MPP approach gives an answer to issues 1, 2 and, partially, 3. By considering objects instead of pixels as unknown variables we obtain a robust detection with respect to noise and to non homogeneous background. However, at this stage issues 4, 5 and 6 have not been yet addressed. In this section we discuss several extensions to the previous model that take into consideration the shape and size variability of the objects within the same MPP framework.

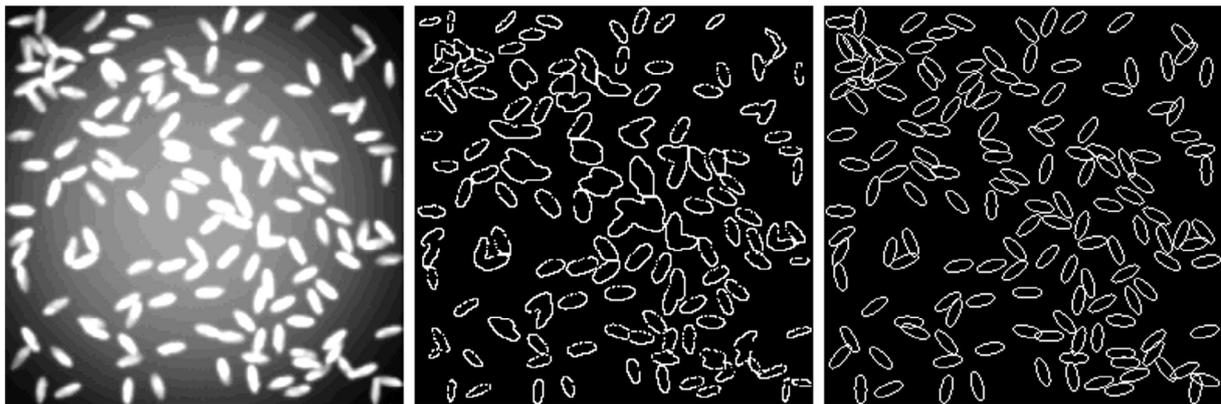


**Fig. 6.** A stongly heterogeneous background (left), results obtained the fiji particle analyser (middle) and with the MPP approach (right).

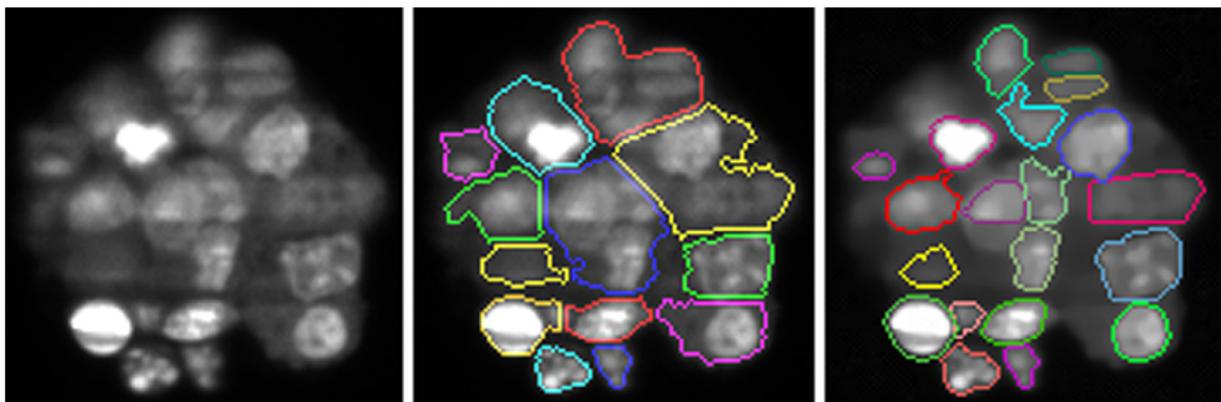
### 3.1. More general shapes

In Section 2, we have considered an object space consisting of disks. However, it is straightforward to take into account more complex parametric spaces in order to define other geometries. For example, if we consider the image given on Fig. 7, it is clear that a model based on disks would fail to detect the cells. The particle analyzer based on the distance transform (Fiji) is also inadequate as shown on Fig. 7 (middle). This problem can be solved within the marked point process framework by considering ellipses as objects, such that the object space is defined by  $\mathcal{O} = [a_{min}, a_{max}] \times [b_{min}, b_{max}] \times [0, \pi]$ . We can see on the MPP result

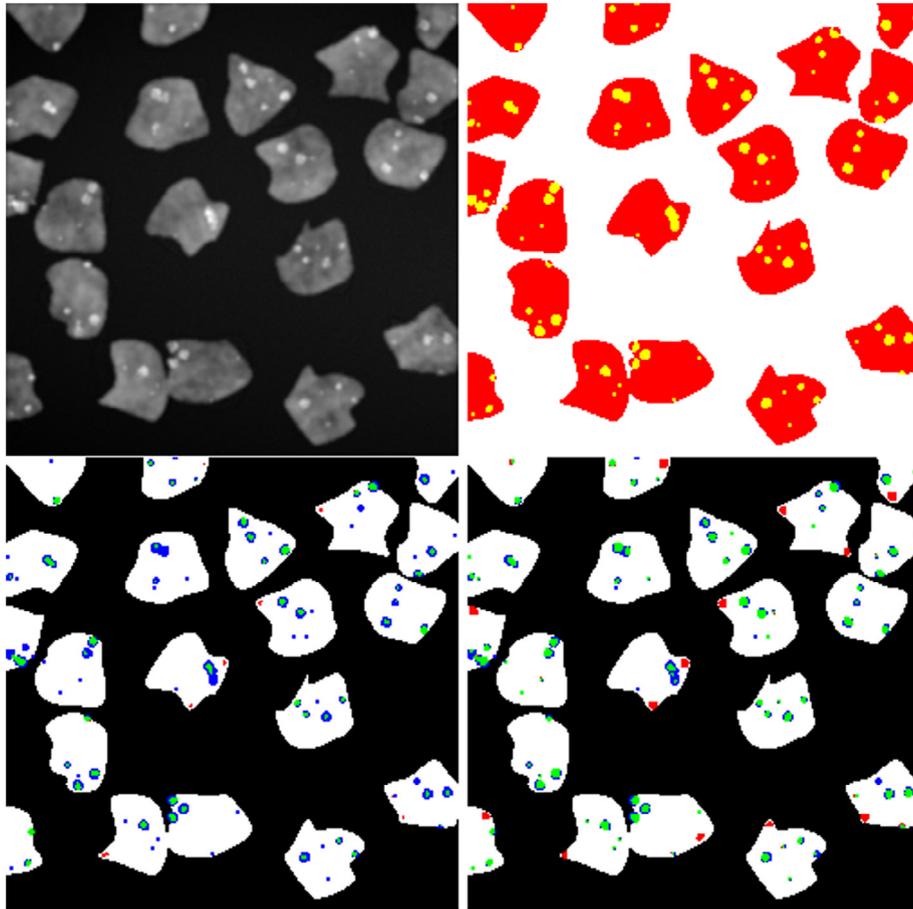
on Fig. 7 (right) that the objects have been correctly detected and delineated, contrary to the particle analyzer approach, that cannot correctly split clusters of ellipses. Several parametric shapes have been proposed in the literature including rectangles, segments or superquadrics. The mix of two shape spaces such as disks and ellipses is also possible. However, the parametric space that defines the shapes should have a low dimension, typically lower than 5, to avoid computational burden. This can be limiting in case of complex shapes such as the cells on Fig. 8, or when the objects are composed of a few pixels, causing the discretization to lead to a poor approximation of the parametric shapes. To overcome this limit, it has been proposed to define the object space as a



**Fig. 7.** Example of elliptic cells (left), results obtained by the Fiji particle analyser (middle) and the MPP approach (right).



**Fig. 8.** Spheroid containing cells of various shapes (left), results obtained with the Fiji particle analyser (middle) and the MPP approach based on a dictionary (right).



**Fig. 9.** An example of small particles within cells (top left), ground truth (top right), results obtained by the wavelet approach proposed by Icy (bottom left) and by the MPP approach (bottom right)– Green: true positive – Blue: false negative – Red: false positive. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

dictionary of precomputed shapes. Such a dictionary can be obtained from previous segmentation maps as in [11] (see the result on Fig. 8) or by constructing an exhaustive description of shapes included in a small bounding box. On Fig. 8 we can see that the use of a dictionary combining shapes obtained by the particle analyzer algorithm and by an active contour approach allows to select the most relevant ones from each method and for different parameters, thus improving the global result. On Fig. 9, the dictionary is defined by the whole set of convex shapes bounded by  $5 \times 5$  pixels square [12]. This last approach has been proven to overcome state of the art detection techniques, such as the one included in Icy which is based on wavelets (see Fig. 9 middle and right).

### 3.2. Energy function

The energy function is composed of a data term that fits the objects onto the image and a prior that favors or imposes properties on the whole configuration. In Section 2 we have considered a contrast term based on the Bhattacharyya distance between the pixel intensity inside the object and in a crown surrounding it. Several formulations to define the data term can be found that compare the mean, the median or the maximum value between the object and its neighborhood. These terms are based on the contrast norm of the object with respect to the surrounding pixels that are supposed to belong to the background. Another class of data terms is based on the scalar product of the normalized intensity gradient on the object frontier and the normal of the shape along this frontier. This notably leads to models that are invariant with respect to

the image contrast [13]. In this setting the data term is entirely based on the object geometry (in the data) independently on the contrast. This is therefore fully adapted to cases when the contrast is not constant within the image due to heterogeneous illumination in the background or variability in the object intensity. The second part of the energy function consists of prior information. In this paper we have considered a repulsive term that prevents objects overlap. Some attractive properties can also be defined, for example, to favor clusters of objects or to align them by sharing similar angles.

### 3.3. Computational considerations

The Multiple Birth and Death algorithm has been proposed in [9] as an alternative to the classical RJMCMC scheme. As for the RJMCMC, the convergence to the configuration that minimizes the energy has been proven in a simulated annealing context [17]. Regarding the MBD algorithm, its main advantage lies in the birth step, where several objects are added simultaneously to the configuration independently of the temperature. Therefore, even at a low temperature, the system can investigate new objects that locally increase the energy. To improve the convergence speed, one can introduce a birth map to favor introducing new objects with a higher probability in relevant locations within the image. For example this birth map can be based on a precomputation of the data terms for each possible location of the objects. A tradeoff needs then to be found between the complexity of the birth map computation and the gain in terms of number of iterations before convergence. Some suboptimal algorithms have been

proposed in order to speed up the convergence. For these algorithms there is no guarantee to reach the global optimizer, but they have proven to be efficient in practice. As for Markov Random Fields, the graph cut algorithm has been employed by replacing the death step in the MBD by a graph cut to select the most relevant objects. This has led to the Multiple Birth and Cut (MBC) algorithm [18]. The main advantage of the MBC over RJMCMC and MBD schemes is that it prevents from embedding the algorithm into a simulating annealing scheme thus avoiding the calibration of the cooling parameters (initial value and decreasing coefficient of the temperature). To speed up the convergence speed one can consider a deterministic version of the MBC algorithm that can be compared to the ICM for Markov Random Fields. It simply consists of the removal of an object - during the death step - in the case that this change in the configuration induces an energy decreasing. A quantitative comparison between stochastic samplers, in terms of accuracy and computational time, is given in [19] on a particular application. However, the efficiency of these algorithms highly depends on their design (kernel choice in case of RJMCMC, birth map in case of MBD), thus each case needs to be specially studied.

### 3.4. Further MPP applications to biology imagery

In this paper we focus on object detection. However, other problems can be addressed within the marked point process framework as it has been proposed for remote sensing applications. In this particular context, a hierarchical model has been proposed by [14] to model groups of vehicles (i.e.: each vehicle is a first order object and a group of them is a second order). Such a multi-level model can be applied to study populations of vesicles within cells. An extra dimension representing time can be added to MPP models in order to obtain an object tracking algorithm [15]. MPP can thus be employed to study vesicles trajectories. Finally, the transition between two states of a given object, for example a cell from alive to dead, can be addressed through change detection models [16].

## 4. Conclusion

The MPP approach, originally developed in the domain of spatial statistics for the modeling of populations, has been more recently successfully applied to solve image analysis problems and, more particularly, multiple object detection from images. In this paper we have shown that this framework is well suited to perform object detection in biological imagery. The different issues raised by these applications can be satisfactorily addressed with MPP modeling. This includes noise, shape variability and background heterogeneity. The non convexity of the functional to be minimized may lead to heavy computational time, especially when treating 3D datasets. However, some sub-optimal algorithms have been proposed that make the approach usable in practice. Some issues remain unsolved concerning MPP. Shapes are currently defined in a low dimensional parametric space or in a predefined dictionary. To consider general shapes defined in a shape space, as for example in [20], is a very challenging issue. Apart modeling issues we can also mention improvement in the optimization to speed up the convergence or to define parallel implementation [19]. Finally, estimating the parameters is still a largely open issue [21].

## Acknowledgments

The author would like to thank A. Razetti for careful proof reading.

## References

- [1] M.G. Forero, A. Hidalgo, Image processing methods for automatic cell counting in vivo or in situ using 3D confocal microscopy, in: Gaetano Gargiulo (Ed.), *Advanced Biomedical Engineering*, InTech, 2011, <http://dx.doi.org/10.5772/23147>.
- [2] I. Smal, M. Loog, W. Niessen, E. Meijering, Quantitative comparison of spot detection methods in fluorescence microscopy, *IEEE Trans. Med. Imaging* 29 (2) (2010) 282–301.
- [3] N. Chenouard, I. Smal, F. de Chaumont, M. Maska, I.F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, A.R. Cohen, W.J. Godinez, K. Rohr, Y. Kalaidzidis, L. Liang, J. Duncan, H. Shen, Y. Xu, K.E. Magnusson, J. Jaldén, H.M. Blau, P. Paul-Gilloteaux, P. Roudot, C. Kervrann, F. Waharte, J.Y. Tinevez, S.L. Shorte, J. Willemsse, K. Celler, G.P. van Wezel, H.W. Dan, Y.S. Tsai, C. Ortiz de Solórzano, J.C. Olivo-Marin, E. Meijering, Objective comparison of particle tracking methods, *Nat. Methods* 11 (3) (2014) 281–289. Epub 2014 Jan 19.
- [4] J. Schindelin, I. Arganda-Carreras, E. Frise, et al., Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (7) (2012) 676–682.
- [5] A.E. Carpenter, T.R. Jones, M.R. Lamprecht, C. Clarke, I.H. Kang, O. Friman, D.A. Guertin, J.H. Chang, R.A. Lindquist, J. Moffat, P. Golland, D.M. Sabatini, Cell Profiler: image analysis software for identifying and quantifying cell phenotypes, *Genome Biol.* 7 (2006) R100.
- [6] L. Miroslaw, A. Chorazyczewski, F. Buchholz, R. Kittler, Correlation-based method for automatic mitotic cell detection in phase contrast microscopy computer recognition systems, in: *Proceedings of 4th International Conference on Computer Recognition Systems CORES '05*, in: M. Kurzynski (Ed.), *Advances in Soft Computing*, Springer, Berlin, Heidelberg, 2005, pp. 627–634. Ch. 4.
- [7] *Stochastic geometry for image analysis* (2011). Xavier Descombes Ed. Wiley-ISTE, pp. 384, 2011, 978-1-84821-240-4.
- [8] M.N.M. van Lieshout, *Markov Point Processes and Their Applications* Imperial College Press, London, 2000.
- [9] X. Descombes, R. Minlos, E. Zhizhina, Object extraction using a stochastic birth-and-death dynamics in continuum, *J. Math. Imaging Vision* 33 (3) (2009) 347–359.
- [10] A. Lehmussola, P. Ruusuvaari, J. Selinummi, H. Huttunen, O. Yli-Harja, Computational framework for simulating fluorescence microscope images with cell populations, *IEEE Trans. Med. Imaging* 26 (7) (July 2007) 1010–1016.
- [11] E. Poulain, S. Prigent, E. Soubies, X. Descombes, Cells detection using segmentation competition, ISBI – International Symposium on Biomedical Imaging, IEEE, Brooklyn, United States, Apr 2015.
- [12] N. Cediñnik, E. Debreuve, X. Descombes, SPADE: Small Particle Detection, 2016, <<https://pypi.python.org/pypi/small-particle-detection/>>.
- [13] E. Soubies, P. Weiss, X. Descombes, Graph cut based segmentation of predefined shapes: applications to biological imaging, *Pattern Recognition Applications and Methods, Advances in Intelligent Systems and Computing*, 318, Springer-Verlag, 978-3-319-12610-4, 2015. pp. 153–170.
- [14] A. Börncs, C. Benedek, Extraction of vehicle groups in airborne lidar point clouds with two-level point processes, *IEEE Trans. Geosci. Remote Sens.* 53 (3) (2015) 1475–1489.
- [15] P. Craciun, M. Ortner, J. Zerubia, Joint detection and tracking of moving objects using spatio-temporal marked point processes, in: *IEEE Winter Conference on Applications of Computer Vision*, Hawaii, United States, Jan 2015.
- [16] C. Benedek, X. Descombes, J. Zerubia, Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (1) (2012) 33–50.
- [17] E. Zhizhina, X. Descombes, Double annealing regimes in the multiple birth-and-death stochastic algorithms, *Markov Processes Relat. Fields* 2012 (18) (2012) 441–456. Polymath.
- [18] A. Gamal Eldin, X. Descombes, G. Charpiat, J. Zerubia, Multiple birth and cut algorithm for multiple object detection, *J. Multimedia Process. Technol.* (2012).
- [19] Y. Verdié, F. Lafarge, Detecting parametric objects in large scenes by Monte Carlo sampling, *Int. J. Comput. Vision* 106 (1) (2014) 57–75.
- [20] A. Srivastava, S. Joshi, W. Mio, W. Liu, Statistical shape analysis: clustering, learning and testing, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (4) (April 2005) 590–602.
- [21] F. Chatelain, X. Descombes, J. Zerubia, Parameter estimation for marked point processes, Application to object extraction from remote sensing images, in: *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, Bonn, Germany, August 2009.



# Quantifying receptor trafficking and colocalization with confocal microscopy



Jeremy A. Pike<sup>a,b,c,\*</sup>, Iain B. Styles<sup>d</sup>, Joshua Z. Rappoport<sup>b,e</sup>, John K. Heath<sup>b</sup>

<sup>a</sup>PSIBS Doctoral Training Centre, School of Chemistry, University of Birmingham, Birmingham B15 2TT, UK

<sup>b</sup>School of Biosciences, University of Birmingham, Birmingham B15 2TT, UK

<sup>c</sup>Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK

<sup>d</sup>School of Computer Science, University of Birmingham, Birmingham B15 2TT, UK

<sup>e</sup>Center for Advanced Microscopy and the Nikon Imaging Center, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

## ARTICLE INFO

### Article history:

Received 14 July 2016

Received in revised form 20 January 2017

Accepted 23 January 2017

Available online 26 January 2017

### Keywords:

Colocalization

Trafficking

Endocytosis

Confocal

Receptor

EGFR

## ABSTRACT

Confocal microscopy is a powerful tool for the study of cellular receptor trafficking and endocytosis. Unbiased and robust image analysis workflows are required for the identification, and study, of aberrant trafficking. After a brief review of related strategies, identifying both good and bad practice, custom workflows for the analysis of live cell 3D time-lapse data are presented. Strategies for data pre-processing, including denoising and background subtraction are considered. We use a 3D level set protocol to accurately segment cells using only the signal from fluorescently labelled receptor. A protocol for the quantification of changes to subcellular receptor distribution over time is then presented. As an example, ligand stimulated trafficking of epidermal growth factor receptor (EGFR) is shown to be significantly reduced in both AG1478 and Dynasore treated cells. Protocols for the quantitative analysis of colocalization between receptor and endosomes are also introduced, including strategies for signal isolation and statistical testing. By calculating the Manders and Pearson coefficients, both co-occurrence and correlation can be assessed. A statistically significant decrease in the level of ligand induced co-occurrence between EGFR and rab5 positive endosomes is demonstrated for both the AG1478 and Dynasore treated cells relative to a control. Finally, a strategy for the visualisation of co-occurrence is presented, which provides an unbiased alternative to colour overlays.

© 2017 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The signalling and trafficking of cellular receptors are highly interlinked processes [1–3]. Ligand induced signalling regulates endocytosis and receptor trafficking within the endocytic network, which in turn attenuates receptor signalling. Moreover, the hypothesis of signalling endosomes, for which there is now extensive evidence, implies that the subcellular location of activated receptor triggers distinct signalling responses [3–8]. Homeostatic receptor trafficking is essential for organism development [9,10], and aberrant activity is implicated in numerous diseases [11,12].

Fluorescence microscopy is commonly used to study ligand induced changes to the quantity of receptor located at the plasma membrane [13], and also colocalization with subcellular structures, such as endosomes [14]. Developing a proper understanding of these experiments requires quantitative, unbiased, and repro-

ducible analysis protocols. In this paper, with these requirements in mind, we describe fully automated image analysis workflows for analysing live cell 3D time-lapse data. Confocal microscopy is used exclusively, but all protocols are equally applicable to deconvolved widefield images [15]. A HeLa cell line expressing fluorescent protein tagged constructs for both epidermal growth factor receptor (EGFR) and rab5 is used as a model system [16]. rab5 is an early endosome associated GTPase and key regulator of receptor trafficking [17]. Inhibitors for EGFR kinase (AG1478) [18] and dynamin (Dynasore) [19] are used to perturb the ligand (EGF) induced trafficking response. We demonstrate the effectiveness of the described workflows, and show that both drug treatments perturb EGFR trafficking and colocalization with rab5 positive endosomes.

The rest of the paper is structured as follows. In Section 2 a brief review of related approaches, identifying both good and bad practice, is presented. Section 3 describes, and shows the use of, the proposed protocols. Finally, Sections 4 and 5 provide a discussion and conclusion.

\* Corresponding author at: Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge CB2 0RE, UK.

E-mail address: [jeremy.pike@cruc.cam.ac.uk](mailto:jeremy.pike@cruc.cam.ac.uk) (J.A. Pike).

## 2. Related approaches

### 2.1. Workflows to quantify the subcellular distribution of receptor

Confocal microscopy data can be acquired in either two, or three, spatial dimensions. In a 3D approach multiple axial slices are acquired at different focal planes through the sample. A 2D approach is inherently superior to 2D acquisition as the entire cellular volume can be sampled [5]. In a 2D approach only a single plane through the cell is acquired, hence key information can be missed [13]. If the axial position is not set automatically, for example at a set distance from the coverslip, or at the widest nuclear plane, then user bias is introduced to the acquisition [20]. Note that for live samples light exposure should be kept low. Therefore for time-lapse imaging there is a practical trade-off between the number of axial slices, and the frame-rate.

Post-acquisition, data can be processed to isolate biologically relevant regions of interest (ROIs) such as the plasma membrane. Subsequently, the fluorescence intensity of a specific marker can be quantified within each ROI. Selection of ROIs can either be manual or automated. Manual selection should be avoided as it is prone to user bias and error, and time-consuming and difficult to implement in 3D. With time-lapse data the change in (normalised) intensity, within each ROI, over time can be calculated. For example Fortian and Sorkin (2014) acquired 3D time-lapse data with spinning disk confocal microscopy, and used an automated 3D edge based segmentation protocol to identify the cellular ROI [5]. The segmentation was eroded by a set number of pixels to identify ROIs for the intra-cellular region and plasma membrane. This was used to calculate the normalised percentage, of both EGF and Grb2, associated with the plasma membrane over time. This is an excellent example of an automated 3D strategy for the quantification of temporal changes to the subcellular distribution of a fluorescent construct. However, the pre-processing and segmentation protocols are not fully defined, only the software package and associated components are cited. As the specific image processing algorithms are not referenced, reproduction of this methodology has not been possible in an alternative software application.

ROI intensities, and colocalization measures, can be calculated using either the raw or pre-processed data. Raw data refers to the unprocessed data as acquired by the microscope. There is extensive literature on both image denoising and deconvolution [15,21,22]. These techniques respectively aim to remove corruption and out of focus contributions within image data. Although these approaches can be inaccessible for biological researchers, due to either lack of knowledge or user-friendly tools, working with raw data cannot be considered best practice. However, the use of unjustified or poorly specified methods is worse as results cannot be reproduced. Following the initial pre-processing steps, data can be further processed to enhance, or isolate, biologically meaningful components. Note there is no generalised workflow for image pre-processing and care should be taken to match the approaches used to both the data and the biological context. For example, Dunn et al. (2011) suggest that background subtraction, as calculated with a median filter, is appropriate for the quantification of signal within endosomes [14]. When the width of the filter is at least twice as large as the endosomal structures a reliable estimation of local background is produced.

### 2.2. Colocalization analysis

Colocalization analysis is typically used to determine if labelled proteins colocalize, or cluster, to the same subcellular structures. High quality analysis relies on high quality data, and particular care must be taken to avoid detector saturation and cross-talk

between channels [23]. The spatial sensitivity of colocalization analysis is limited by the resolution limit of the microscope, which is determined by the point spread function (PSF) [24]. According to the Nyquist criterion the pixel size, and the axial spacing, should be less than approximately half this limit to accurately represent the sample at this resolution [25]. However, sampling at this optimal rate may be practically infeasible for live experiments, or large scale screens. When using larger pixels, or axial spacing, artefacts can be introduced and it is the pixel size, not the resolution of the microscope, which limits the spatial sensitivity of the colocalization analysis. For example, consider the imaging of endosomes using a pixel size of 0.25  $\mu\text{m}$ , and axial spacing of 0.5  $\mu\text{m}$ . When using standard imaging wavelengths, and an objective with a numerical aperture of 1.4, this is larger than the size defined by the Nyquist criterion. Two endosomes, can only be distinguished if they are separated by more than approximately 0.5  $\mu\text{m}$  laterally, or 1  $\mu\text{m}$  axially. Therefore colocalization analysis, even using super-resolution techniques, is poorly suited to the identification of direct protein-protein interaction [26]. Techniques such as Förster Resonance Energy Transfer (FRET) are more appropriate for this purpose [27]. Conversely, when there is no direct interaction between the proteins but association within subcellular structures, such as endosomes, FRET cannot be used.

In studies of receptor trafficking, quantitative analysis is often neglected, and colour merges are used to provide qualitative evidence for colocalization [6,28–30]. This can leave interpretation and presentation of results open to user bias, either through the image display settings, or the choice of *representative* images. Visualisation of correlation is better performed using joint-histograms, not colour merges [14]. There are two distinct strategies for colocalization analysis. The first is based on the overlap, or correlation, between pixels [14,23,25]. The second detects objects within the data and uses the centre of mass for each object to determine clustering statistics such as Ripley's K-function [25,26,31]. Object based methods have shown promising results for localization and TIRF microscopy, where the data is well modelled by point, or spot like, objects [26]. However, for the application of receptor trafficking using confocal microscopy, the receptor is typically localised to either the plasma membrane or endosomal structures, the former of which is not well represented by a point distribution. Therefore we will focus on pixel based measures which can be split into two categories; co-occurrence and correlation [32]. Co-occurrence measures quantify how often, or how much, signal from each channel overlaps with the other channel based only on the presence, or absence, of signal. For example, 50% of channel 1 signal overlaps with channel 2 signal. Correlation measures assess the extent of a relationship between the signals from each fluorophore. For example, if there is high positive linear correlation a pixel with high intensity in channel 1 would typically also have high intensity in channel 2. For high negative linear correlation a pixel with high intensity in channel 1 would typically have low intensity in channel 2.

The Manders Coefficients (MCs) (M1 and M2) are well-established co-occurrence measures which simply calculate the percentage of total signal from one channel which overlaps with signal from the other, such that [33],

$$M1 = \sum_i \frac{C_{1,coloc}}{C1_i} \quad M2 = \sum_i \frac{C_{2,coloc}}{C2_i} \quad (2.2.1)$$

where  $C1_i$  and  $C2_i$  represent the intensities of individual pixels for channels 1 and 2 respectively.  $C_{1,coloc}$  and  $C_{2,coloc}$  represent the colocalizing pixels such that  $C_{1,coloc} = C1_i$  when  $C2_i > 0$  and  $C_{1,coloc} = 0$  otherwise. Similarly  $C_{2,coloc} = C2_i$  when  $C1_i > 0$  and  $C_{2,coloc} = 0$  otherwise. The Pearson coefficient (PC), R is a well-established measure of linear correlation, defined such that [34],

$$R = \frac{\sum_i (C1_i - C1_{av}) \times (C2_i - C2_{av})}{\sqrt{\sum_i (C1_i - C1_{av})^2 \times \sum_i (C2_i - C2_{av})^2}} \quad (2.2.2)$$

where  $C1_{av}$  and  $C2_{av}$  are the average intensities for channel 1 and channel 2 signal respectively. In this work we have avoided the use of combined correlation and co-occurrence measures such as the Manders overlap coefficient (distinct from the MCs), and the more recent measure introduced by Singan et al. (2011), as it is advantageous to maintain the ability to distinguish between co-occurrence and correlation [33,35]. The Manders overlap coefficient is also sensitive to small variations in background signal and detector offset [32]. A promising new colocalization measure was introduced by Humpert et al. (2015) which is robust at both low signal to noise ratios (SNR) and varying background levels [36]. It can also be used to simultaneously analyse more than two fluorescent channels. However the interpretation of this measure is not yet clear.

To calculate the MCs it is necessary to isolate the pixels containing biologically relevant signal from both channels. Although signal isolation is not an essential step for the calculation of the PC, Adler et al. (2010) argue that, to prevent artificial inflation, only the pixels containing isolated signal from both channels should be used to calculate the PC [32]. When calculated using only pixels containing isolated signal from both fluorophores, the PC is easily interpreted as a measure of linear correlation within pixels containing probes from both channels [14,26]. This will often have a clear biological interpretation, for example the correlation in endosomes positive for both rab5 and EGFR. Therefore, if both the PC and the MCs are calculated, the level of overlap between the two channels, and the correlation within that overlap can be simultaneously assessed. If all pixels in the ROI, for example a cell, are used to calculate the PC then the extent of linear correlation between the two signals across the whole cell is evaluated. In this case we consider the biological interpretation to be less clear. Signal isolation should be performed with an automated approach to avoid user bias and error. In the popular approach introduced by Costes et al. (2004) a linear fit is found for the joint histogram of the data [37]. The point on this line of best fit below which there is no correlation ( $R \leq 0$ ) is used to define the threshold values for both channels. A critical discussion of this approach is presented in Section 4.2.

Colocalization analysis is typically performed to test one of two hypotheses; that the level of colocalization is higher than that predicted for randomly distributed signal (within a ROI), or that there is a difference in the level of colocalization between test conditions. In the former case each ROI can be considered independently, but care must be made to avoid auto-correlation effects (see [26] for a recent review). However, for most biological studies it is more informative to consider if the distribution of measurements taken across biologically independent replicates is significant. To do this McDonald and Dunn (2013) showed, using simulated data, that a distribution of PC measurements can be compared to the expected value of  $R = 0$  using a  $t$ -test. Similarly, a distribution of either M1, or M2, measurements can be compared to the fractional volume of the ROI occupied by either channel 2, or channel 1, respectively [38]. The difference between the MCs and the expected value can be expressed as,

$$M1_{diff} = M1 - \frac{V_2}{V_{ROI}} \quad M2_{diff} = M2 - \frac{V_1}{V_{ROI}} \quad (2.1.1)$$

where  $V_{ROI}$  is the total volume of the ROI and  $V_1$ ,  $V_2$  are the volumes of the isolated signal from channel 1 and channel 2 respectively. Note that an accurate segmentation of the ROI containing the signal from both channels is essential for this approach. For the application of receptor trafficking, the ROI should contain the cytoplasm and plasma membrane, but not intracellular structures such as the nucleus. Moreover it is not clear if an expected value of  $R = 0$

is suitable for real data where auto-correlation, or imperfect ROI selection, can lead to inflation of the PC. McDonald and Dunn also showed that a two way  $t$ -test can be used to test the null-hypothesis that two distributions of either the PC or the MCs, have the same mean. This approach is robust, and also easy to implement, as the effects of ROI selection, signal isolation and auto-correlation will be reproduced for both conditions. Therefore, when possible, we consider it highly desirable to design experiments with a negative control and to test for changes in the level of colocalization.

### 3. Methods and results

In this section the proposed analysis workflows, designed specifically for live cell 3D time-lapse datasets, are presented (Fig. 1). The use of these workflows is demonstrated on confocal microscopy data of cells expressing both EGFR-EGFP, and rab5-mRFP constructs. The cells were treated with either AG1478, Dynasore or a DMSO control. The cell culture and microscopy methodology is described in Appendix A. Unless otherwise stated, the algorithms were implemented in 3D using Matlab (v2015a, The Math-Works, Inc., Natick, MA, USA).<sup>1</sup>

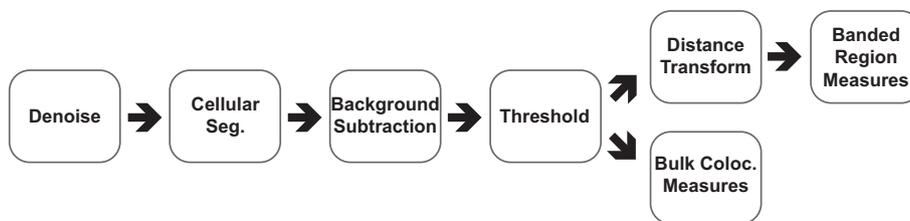
#### 3.1. Data pre-processing and cellular segmentation

Image pre-processing and segmentation are essential components of the analysis. The first step is the manual cropping of each time-series such that the cropped data contains a single cell. This quality control measure was the only manual component of the workflow. This was done blindly and efficiently using a custom-designed interface, where the only criterion for selection is that the cell be alive and non-mitotic. A maximal projection of the first time-point is used to define a region for the cropping of the entire time-lapse.

Each cropped time-lapse was then denoised using an ImageJ plugin implementing the PURE-LET scheme described by Luisier et al. (2010) [39]. ImageJ was run within a Matlab script using MIJ and the plugin was set to automatically estimate noise parameters (4 spin cycles, 3 multi-frame) [40,41]. The PURE-LET scheme is designed for the removal of Poisson noise, it is simple to use, relatively fast, and has been shown to have similar performance to other state of the art methods. Noise with a Poisson distribution is produced in the imaging process by the inherent uncertainty in arrival time of photons at the detector [42]. Briefly, the PURE-LET scheme estimates and minimises the error between the unknown noiseless image and the processed image based on an assumption of Poisson noise.

To identify the cellular volume (segmentation), the denoised EGFR-EGFP channel was processed using a 3D level set segmentation protocol. Specifically, we used a 3D implementation of the edge based distance regularized level set evolution (DRLSE) framework described by Li et al. (2010) (parameters listed in Table A.2) [43,44]. This powerful approach facilitated the segmentation of the cellular boundary using only the EGFR-EGFP signal, where conventional thresholding approaches would fail. The DRLSE term allows for a simple finite difference implementation without the need for re-initialising sub-routines. The computational costs of a level set framework are high so it is advantageous to implement a fast protocol to obtain an initial estimate of the segmentation, and to use this estimate as the starting point for the level set algorithm. To do this an algorithm based on K-means clustering was used (described in Appendix A, example shown in Fig. 2B). After both

<sup>1</sup> Matlab code available at <https://github.com/JeremyPike/receptor-trafficking-toolbox>



**Fig. 1.** Flowchart summary of image analysis workflows. The first step is data denoising which is performed to reduce the corruption introduced during image acquisition. This is followed by automated segmentation of the cellular boundary using the membrane bound receptor signal. This is done to produce an accurate region of interest (ROI) for the subsequent analysis protocols. Background subtraction is followed by automated thresholding to isolate signal in both channels. At this point, bulk colocalization statistics can be calculated for each cell, or the cellular ROI can be split into banded volumes using a distance transform. The percentage of receptor (without thresholding), or colocalization measures, can then be calculated for each band. Together these statistics and measures provide a thorough description of the subcellular receptor distribution and colocalization with endosome sub-populations.

the initial and level set based segmentation the largest connected component was selected, and any *holes* were filled.

Ligand treatment triggers receptor internalisation and a decrease in the SNR for the EGFR-EGFP channel at the plasma membrane. Therefore this is a complex segmentation problem. Validation of the segmentation protocol was performed using a reference produced by blind manual segmentation of EGFR-EGFP and transmission images. The SNR at the plasma membrane was sufficient at all time-points to perform manual validation, although a separate membrane stain could have been used if this was not the case. A key limitation of this approach is that manual segmentation will contain bias and not be perfect, but is currently a widely accepted approach to segmentation validation. The Jaccard similarity index,  $J$ , was used as a performance measure [45]. The Jaccard index is defined as the intersection divided by the union of the manual segmentation,  $M$ , and the automated segmentation,  $A$ , such that a perfect match would give  $J = 1$ ;

$$J(M,A) = \frac{|M \cap A|}{|M \cup A|} \quad (3.1.1)$$

The results of this comparison are shown in Fig. 2D. Importantly, a high mean performance ( $J = 0.87$ ) and a statistically significant improvement from the level set method over the K-means based estimate was demonstrated.

The final pre-processing step was background subtraction with a 3D rolling ball approach [46]. In this approach, a background volume was obtained by morphological opening (erosion followed by dilation) of the denoised data using a spherical structure element (1  $\mu\text{m}$  radius). This background was then subtracted from the denoised data to produce background subtracted data. When performing rolling ball background subtraction it is important to set the radius to be at least as large as the width of the largest biologically relevant structures, in this case, endosomes. Note, for the denoising, segmentation and background subtraction steps, each 3D time-point, and each channel, were processed independently.

### 3.2. Quantifying subcellular receptor distribution over time

In this subsection, a protocol for quantifying the subcellular distribution of receptor over time is presented. The data was first pre-processed, and the cellular ROI was segmented, as described in the previous subsection. Subsequently, the cellular ROI (for each time-point) was split into banded volumes of equal width (0.5  $\mu\text{m}$ ), based on Euclidean distance from the segmentation edge (Fig. 2E). This was done using the computationally fast 3D distance transform described by Mishchenko (2015) [47]. Importantly, the transform calibrates for differences in the lateral and axial spacing.

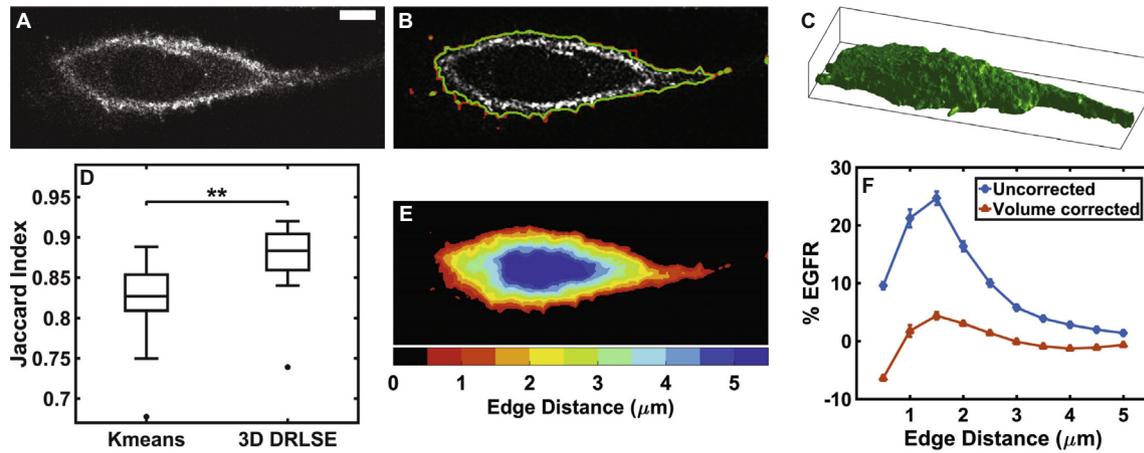
The percentage of the total receptor signal (after pre-processing) contained within each band was calculated to characterise the subcellular distribution at each time-point (Fig. 2F).

These measurements were then volume corrected by subtracting the fractional volume of the band. This was done to calibrate for differences in cell size and shape resulting in bands of varying volume (Section 4.1). Note, bands further than 5  $\mu\text{m}$  from the segmentation edge are not shown as only larger cells will exceed this depth. When analysing time-lapse data the change in receptor distribution, for each cell, can be calculated by subtracting the measurements from the first time-point. This is advantageous and justified as it isolates the ligand induced change in receptor distribution, and corrects for cellular variation in the receptor distribution before ligand treatment.

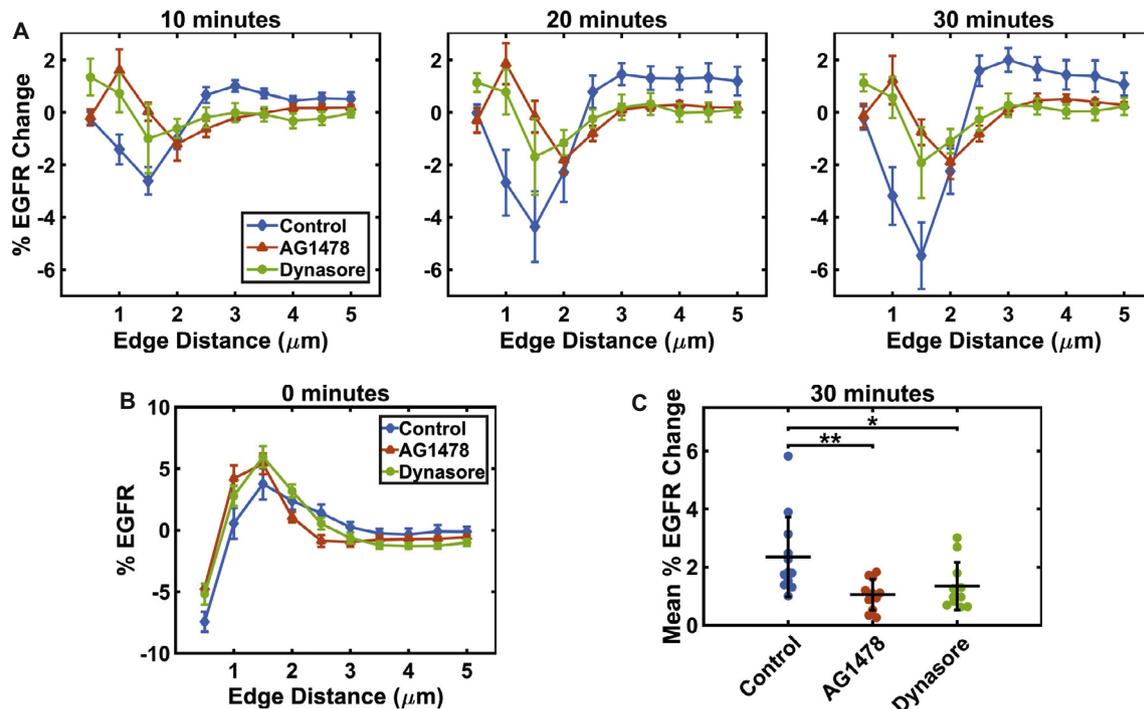
The plots in Fig. 3A show the mean change in percentage EGFR as a function of distance from the segmentation edge across all time-points and conditions. These plots are useful for identifying condition-dependent changes to (ligand induced) receptor trafficking. Fig. 3B shows the initial distribution of receptor at the start of the time-lapse. This can be used to check for any condition dependent variation of receptor distribution before the ligand induced trafficking response. It is informative to perform a statistical analysis to determine if the effects of a specific treatment are significant, relative to the control population. However, when considering distributions of single cell measurements the assumption of a normal distribution is unlikely to be valid, and there will be outliers. Non-parametric tests equivalent to the two-way ANOVA are not well established, so for simplicity we restricted the statistical testing to the change between the first and final time-points (Fig. 3C). For each cell, the mean (absolute) percentage EGFR change across all bands (up to 5  $\mu\text{m}$ ) was calculated. This characterises the total change in subcellular receptor distribution 30 min post ligand treatment. Using this measure, a non-parametric Kruskal-Wallis analysis of variance was performed. If significant ( $p < 0.05$ ), post hoc-testing using a Mann-Whitney  $U$  test, with Bonferroni correction for multiple hypothesis testing ( $n = 2$ ), was performed for each treatment relative to the control. With this approach, a statistically significant reduction in the magnitude of the trafficking response, for both the AG1478 and Dynasore treatments, was identified.

### 3.3. Signal isolation for colocalization analysis

In the previous subsection, the EGFR-EGFP signal was used to characterise changes in the subcellular distribution of receptor. In this section, we suggest and validate a strategy for thresholding both the EGFR and rab5 channels. It is necessary to threshold the data to calculate the Manders coefficients and good practice for the Pearson coefficient (Section 2.2) [32]. After data denoising and background subtraction, as described in Section 3.1, an automated global thresholding approach is typically sufficient. Specifically, we used the Otsu approach where it is assumed that the data histogram is bi-modal, consisting of background and signal peaks



**Fig. 2.** Pre-processing, cellular segmentation and band based analysis of 3D time-lapse data. (A) Representative raw EGFR-EGFP image slice from a 3D stack. Scale bar set at 5 μm and contrast enhanced such that the display range is between zero and half the maximum pixel intensity. (B) Pre-processed data after denoising and background subtraction. The results from the K-means based (red) and level set (green) segmentation protocols are shown. (C) Surface rendering of the level set segmentation result for a single time-point. (D) The segmentation performance of both the K-means estimation and level set algorithm (3D DRLSE) was quantified using the Jaccard index. This was done using 10 datasets where 14 evenly spaced slices per time-point were manually segmented. The mean Jaccard index for the K-means and level set protocols were 0.82 and 0.87 respectively. A sign test was used to determine that the level set protocol produced a significant increase in performance ( $p = 0.002$ ). Central mark on boxplot represents the median, and the edges of the box are the 25th and 75th percentiles. (E) The cellular ROI is split into banded volumes based on distance from the segmentation boundary. Each band has a width of 0.5 μm. (F) Uncorrected plot shows percentage of total cellular EGFR signal contained within each band. For the volume corrected plot, the fractional volume of each band has been subtracted. This was calculated using data from the control population for the first time-point (immediately after EGF treatment). Error bars given by the SEM ( $n = 12$ ). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Band based spatio-temporal analysis of subcellular receptor distribution. (A) Plots of the volume corrected change in percentage EGFR signal over time, for each banded volume (up to 5 μm), in response to EGF treatment. The data for the DMSO control and the two drug treatments are shown (AG1478 and Dynasore). Error bars given by the SEM and  $n > 10$  for all treatments. (B) Plot showing the volume corrected subcellular distribution of EGFR for the first time-point. (C) Plot of mean (absolute) change in percentage EGFR signal (volume corrected) between the first and final time-points (30 min). A Kruskal-Wallis one-way analysis of variance returned  $p = 0.005$ , indicating that there are statistically significant differences between treatments. Post-hoc testing of both drug treatments relative to the control by the Mann-Whitney  $U$  test (corrected by the Bonferroni method), returned  $p = 0.007$ ,  $p = 0.03$  for the AG1478 and Dynasore treatments respectively. The central band represents the mean, and the error bars are the standard deviation.

[48]. The threshold value is defined such that the intra-class variance between the two peaks is minimised. This value best separates the signal and background components of the data. Only pixels contained within the cellular ROIs were included in calcula-

tions, and by combining the data from all time-points, a single threshold value was found for each cell.

To validate this approach on real data, threshold values were manually set for 15 cells at a single time-point. This was done

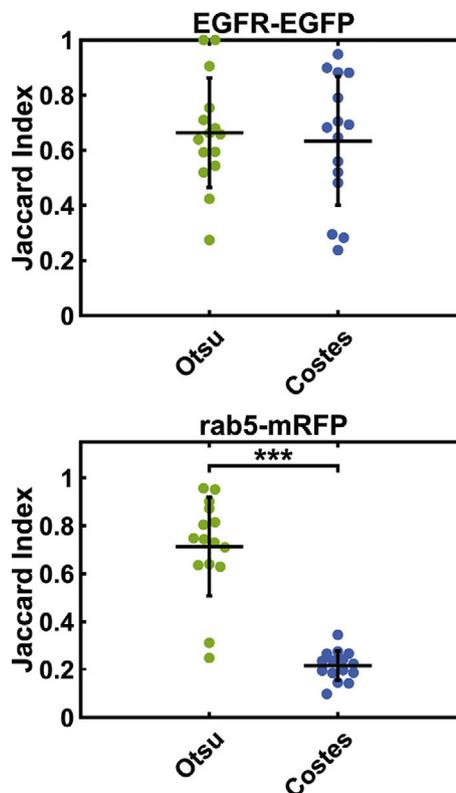
blindly, using a script which randomly selected cells, and time-points, from the entire data-set. Manual thresholds were subsequently set using Fiji [49]. The threshold values were chosen to isolate both the membrane and endosomal associated signal. The resulting binary data was compared to that generated by automated Otsu thresholding using the Jaccard index (Eq. (3.1.1)). The approach proposed by Costes et al. (2004) was also tested (Section 2.2) [37]. Both approaches demonstrated a comparable mean performance for the EGFR channel, but the performance of the Otsu approach was significantly higher for the rab5 channel (Fig. 4).

A comparison of the Otsu and Costes methods was also performed on synthetic data. 3D two channel image stacks of spots (with Gaussian profile to approximate point spread function) were generated in the open image analysis platform, Icy, using a mixed Poisson-Gaussian noise model (Fig. 5A) (detailed in Appendix A) [50,51]. Synthetic data was generated for low, medium and high levels of noise. The level of colocalization was varied from 100% colocalized (spot overlap) to 100% anti-colocalized (spot avoidance). Spots not specified as colocalized (or anti-colocalized) were distributed randomly. 3D Gaussian filtering (width = 1 pixel) was used as a simple pre-processing step. Threshold values were then calculated using either a Costes or Otsu approach on both the raw and pre-processed data. The MCs ( $M1$  and  $M2$ ) were calculated before subtracting the expected value to obtain  $M1_{diff}$  and  $M2_{diff}$  (Eq. (2.1.1)). Fig. 5B shows the rate of failure for the Costes approach in the low noise test data. Failure is defined as extreme over segmentation resulting in 100% signal overlap and a MC equal to one. The Costes approach has non-zero fail rate when there is either no colocalization, or anti-colocalization, indicating that it is not appropriate under these conditions. The Otsu approach does not fail for any tested condition. Fig. 5C shows  $M1_{diff}$  for all levels of colocalization and noise.  $M1_{diff}$  is shown only where the failure rate is zero. In all cases, pre-processing increases the performance. Otsu thresholding (with pre-processing) outperforms the Costes approach across all noise levels. The implications of this analysis are discussed in more detail in Section 4.2.

#### 3.4. Quantitative colocalization analysis for 3D time-lapse data

Section 3.1 describes the cellular segmentation and pre-processing steps of the proposed workflows. Section 3.3 introduces a strategy for signal isolation. Using only the isolated and pre-processed signal, the Manders ( $M1$  and  $M2$ ) and Pearson ( $R$ ) coefficients were calculated for all time-points. To assess the ligand induced change in colocalization over time, the change in all coefficients was calculated, for each cell, by subtracting the measurement from the first time-point. Fig. 6A–C show plots for each of the coefficients across all time-points, and for all conditions.

To identify statistically significant differences in the colocalization response for a specific condition relative to the control, the colocalization coefficients were processed as conventional measurement variables. As in Section 3.2 we restrict our statistical analysis to the change between the first and final time-points. This characterises the change in either co-occurrence, or correlation, for the Manders and Pearson coefficients. A Kruskal-Wallis analysis of variance with post hoc testing of each treatment relative to the control (Mann-Whitney  $U$  test with Bonferroni correction) was then used to identify statistically significant changes in the colocalization response. Fig. 6D–F show the results of this analysis for the Manders and Pearson coefficients. A statistically significant decrease in the level of EGF induced co-occurrence, as measured by the Manders coefficients, was shown for both the AG1478, and Dynasore treatments. No significant change in correlation, as measured by the Pearson coefficient, was established for either treatment.



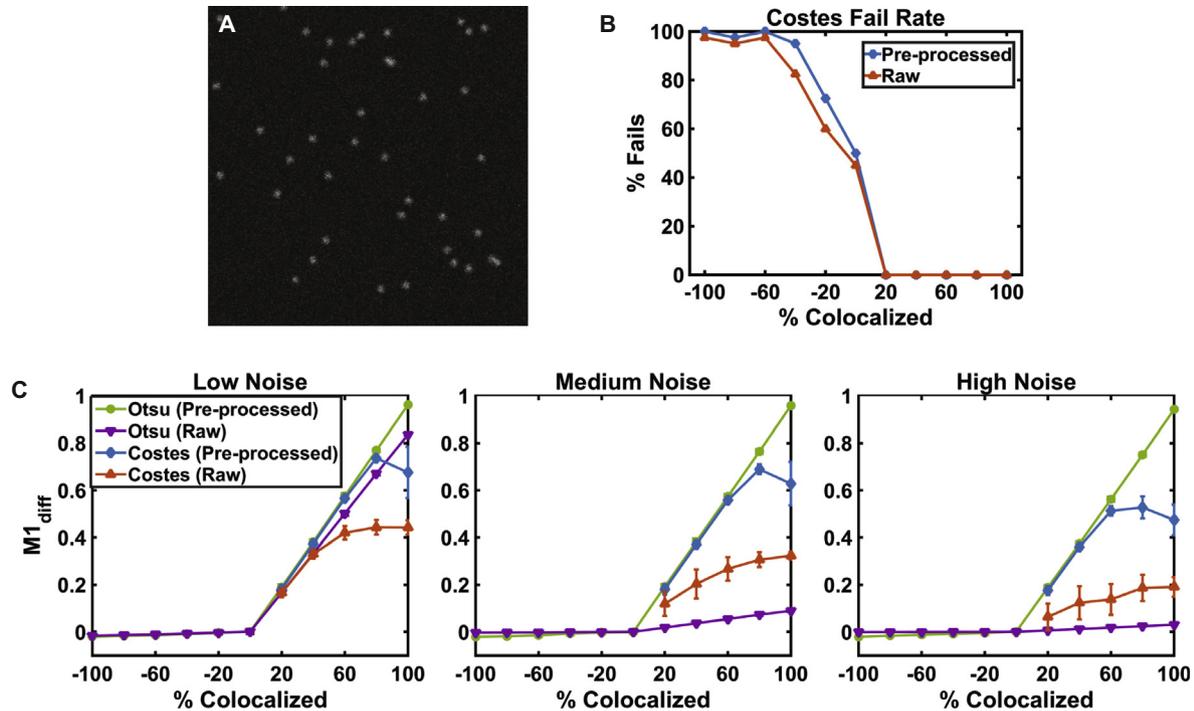
**Fig. 4.** Comparison of the Otsu and Costes thresholding approaches for signal isolation using real data. Before calculating the threshold values, denoising and background subtraction was performed. Both methods were compared to manually set threshold levels from 15 randomly selected cells and time-points. This was done by calculating the Jaccard index,  $J$ , using the resulting binary images. For the EGFR-EGFP channel, the mean performance was  $J = 0.66$  and  $J = 0.63$ , for the Otsu and Costes methods respectively. A sign test returned  $p = 0.6$  indicating that no significant difference in performance was detected. For the rab5-mRFP channel, the mean performance was  $J = 0.71$  and  $J = 0.22$  for the Otsu and Costes methods respectively. A sign test returned  $p = 9.8 \times 10^{-4}$  indicating that the Otsu method performed significantly better than the Costes approach.

#### 3.5. Visualising co-occurrence in 3D

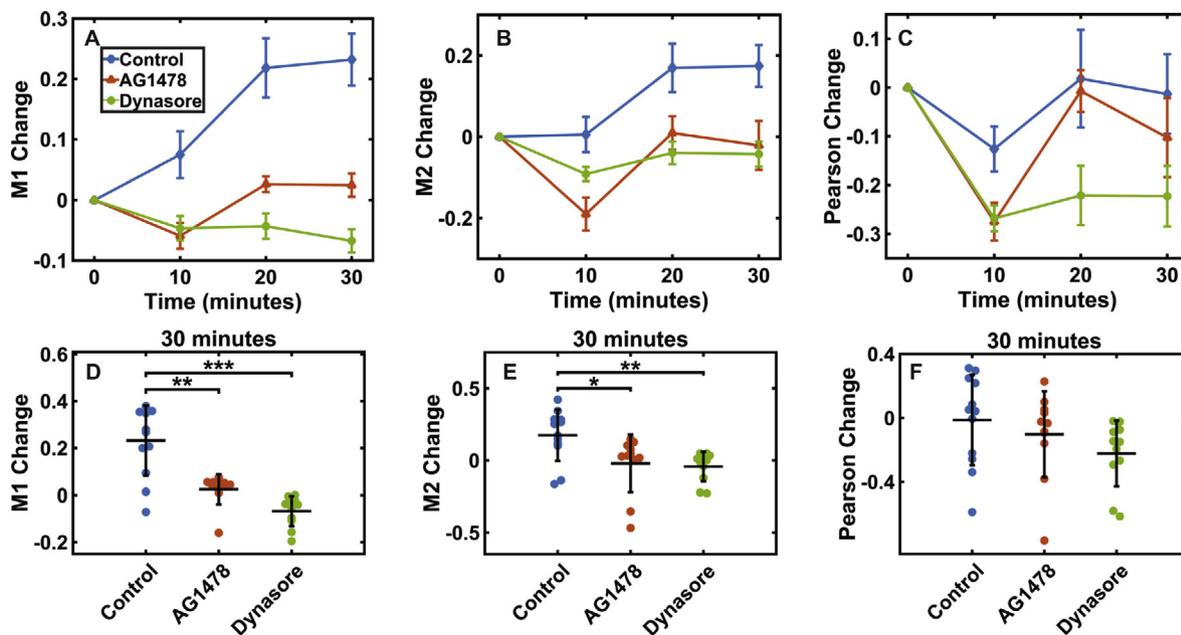
In this subsection an unbiased strategy for the visualisation of colocalization is presented. The proposed method produces a spatial map of co-occurrence. For a specific cell and time-point, the relative contribution of each pixel to one of the Manders coefficients (either  $M1$  or  $M2$ ) is used to determine the intensity of the co-occurrence map. For  $M1$  this is the isolated signal from channel 1, which overlaps with isolated signal in channel 2. This 3D map is then scaled between the maximum and minimum display values, and normalised by multiplying by  $M1$  (or  $M2$ ) for the given time-point. This normalisation allows for unbiased visual comparison between multiple cells.

The co-occurrence maps (and cellular ROIs) can be visualised interactively in 3D (Fig. 7A). To do this, the Matlab vol3d v2 script<sup>2</sup> was used, but any software capable of 3D data rendering could be used for this purpose. If 2D visualisation is required then either the maximal projection, or single slices, of the map can be shown. In Fig. 7B, maximal projections of  $M1$  co-occurrence maps are shown using a false colour lookup table (LUT) for representative cells from the control, AG1478 and Dynasore treatments (20 min post EGF treatment). Finally, recall that it is informative to include joint-

<sup>2</sup> available at <http://uk.mathworks.com/matlabcentral/fileexchange/22940-vol3d-v2>, accessed 26/06/16



**Fig. 5.** Comparison of the Otsu and Costes thresholding approaches for signal isolation using synthetic data. (A) 3D test data of spots was generated for varying levels of colocalization (and anti-colocalization), and for three levels of noise. 50 stacks were generated per condition. A single slice from the medium noise level is shown. (B) In some instances the Costes approach fails to find appropriate threshold values resulting in Manders coefficients equal to one. The percentage of stacks for which the Costes approach fails is shown for each level of colocalization (low noise). (C)  $M1_{diff}$  is calculated by subtracting the expected value from the Manders Coefficient M1. This is shown for all methods, levels of colocalization and noise levels. Note  $M1_{diff}$  is shown only when the failure rate is zero.



**Fig. 6.** Quantifying colocalization with 3D time-lapse data. (A–C) Plots of change in either the Manders (M1 and M2), or Pearson coefficients. Measurements for a DMSO control, AG1478 and Dynasore treatments are shown. These plots characterise the change in co-occurrence (Manders) or correlation (Pearson) between EGFR-EGFP and rab5-mRFP, in response to EGF treatment. Error bars are given by the SEM and  $n > 10$  for all treatments. (D, E) Plots of the change between the first and final time-points for the Manders and Pearson coefficients. Central band represents the mean, and the error bars are the standard deviation. This indicates that there are statistically significant differences between treatments for both M1 and M2. (D) For M1 post-hoc testing of both drug treatments, relative to the control, by the Mann-Whitney  $U$  test (corrected by the Bonferroni method) returned  $p = 4 \times 10^{-5}$ ,  $p = 0.002$ ,  $p = 0.08$  for M1, M2 and the Pearson coefficient respectively. This indicates that there are statistically significant differences between treatments for both M1 and M2. (D) For M1 post-hoc testing of both drug treatments, relative to the control, by the Mann-Whitney  $U$  test (corrected by the Bonferroni method) returned  $p = 0.008$ ,  $p = 6 \times 10^{-4}$  for the AG1478 and Dynasore treatments respectively. (E) For M2 Post-hoc testing returned  $p = 0.02$ ,  $p = 0.007$  for the AG1478 and Dynasore treatments respectively. (F) For the Pearson coefficient the analysis of variance was not significant so no post-hoc testing was performed.

histograms as a visualisation of correlation [14]. Fig. 7 C shows the corresponding joint-histograms where only pixels within the cellular ROI have been included.

### 3.6. Quantifying subcellular colocalization distribution over time

In Section 3.2 a method was proposed which split the cellular segmentation into banded regions using a 3D distance transform. In Sections 3.2–3.5 workflows for quantifying and visualising colocalization were presented. Here these two strategies are merged to quantify the subcellular distribution of colocalization. We propose two approaches. In the first we consider the signal contained within each band individually and calculate both the Manders coefficients and the Pearson coefficient. This quantifies the level of co-occurrence and correlation for each band. In this approach, if we compare the measurements from two different bands, we cannot say if there is more, or less, co-occurring signal, only that a greater or smaller percentage of that band's signal is co-occurring. In the second approach, the percentage of signal from the co-occurrence map (Section 3.5) contained within each band is calculated and volume corrected. This approach follows the same protocol as that described for the EGFR signal in Section 3.2 and quantifies where the co-occurring signal is located. When comparing two bands we can say that there is more co-occurrence in one band but we cannot state what percentage of the band's signal is co-occurring, or comment on correlation.

The results from the first approach are shown in Fig. 8. Fig. 8A shows the change in M1 with respect to the first time-point across all banded regions up to 5  $\mu\text{m}$  from the cellular segmentation. This characterises the ligand induced changes in co-occurrence on a subcellular level. Note, in the control population, the dominant increase in co-occurrence occurs less than 3  $\mu\text{m}$  from the edge. This is as expected since we are quantifying colocalization between receptor and early endosomes. Fig. 8B shows the same analysis for the change in the Pearson coefficient. Fig. 8C and D show M1 and the Pearson coefficient at the first time-point. Fig. 9 shows the results from the second approach. Note from Fig. 9A that there is no clear change in how the co-occurring signal is distributed in response to ligand. Fig. 9B shows the distribution of the co-occurrence map at the first time-point. Fig. 9C shows the mean (absolute) change in percentage signal between the first and final time-point across all bands up to 5  $\mu\text{m}$ . No statistically significant change was found.

## 4. Discussion

### 4.1. Band based analysis of subcellular receptor distribution

In the band-based approach introduced in Section 3.2, the cellular ROI is split into volumetric regions based on distance from the segmentation edge. The relative distribution of receptor signal across the banded volumes can then be calculated. This approach facilitates the quantification of both the surface (membrane associated) and intracellular signal. Importantly, the analysis is automated and performed in 3D, hence avoiding any user bias. The level set based cellular segmentation protocol has high performance when using the receptor signal as the input (Fig. 2D). This is advantageous for live cell microscopy as the use of a secondary marker for the plasma membrane, or cytoplasm, would increase sample exposure and experimental complexity. This increase in light exposure, and associated photo-toxicity, would reduce the sampling and/or frame-rate achievable.

The measurement for each band was volume corrected by subtracting the expected value assuming a homogenous distribution of signal, specifically the fractional volume of each band. This is

justified because cells have varying size and shape which in turn, change over time. By subtracting the expected value, we calibrate for varying band volume which facilitates direct comparison between cells over time. The bands were chosen to be of equal width (0.5  $\mu\text{m}$ ) for ease of interpretation. In this approach the number of bands per cell will vary and therefore we chose to restrict plots to a maximum of 5  $\mu\text{m}$  from the segmentation edge. An alternative approach would be to fix the number of bands but vary the band width between cells. This would allow for quantitative comparison of the entire cellular volume, but the interpretation would be less clear.

In the presented workflows, the nucleus is not excluded from the analysis ROI. Due to cellular size and shape variation, the proportion of nuclear region contained in a specific band will vary. Since the nuclear region contains less receptor, and no endosomes, this will introduce noise to bands in the cellular interior. Therefore using a nuclear stain to segment the nucleus would be advantageous. However, as with the addition of a membrane stain, the use and imaging of a nuclear marker would require additional light exposure. If a nuclear stain is used, the band based analysis could be repeated within the nuclear region, or alternatively extending outwards from the nucleus to the plasma membrane. Although this is not particularly relevant for the EGFR and rab5 data presented, this is an excellent example of how the proposed workflows could be adapted for a variety of applications.

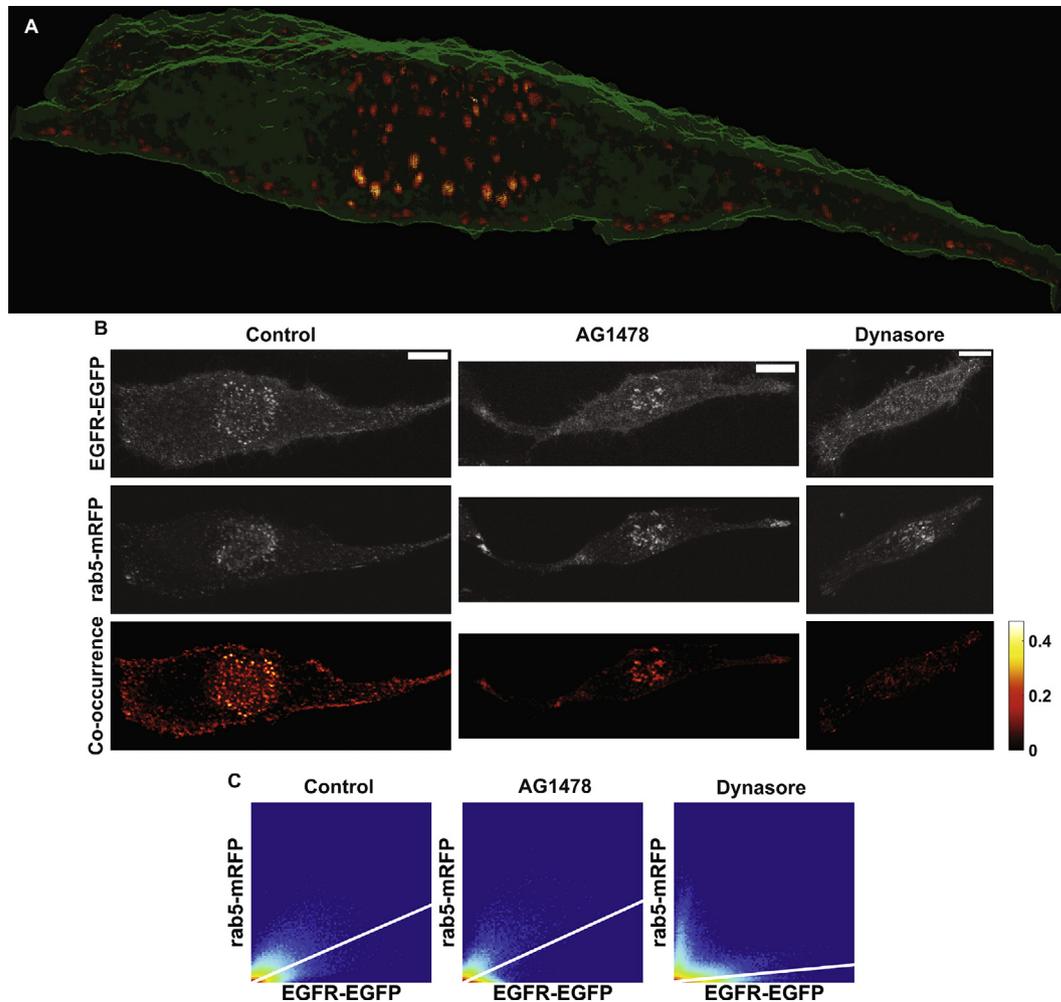
When using live-cell time-lapse data, the change in a given measurement can be isolated at a single cell level, by subtracting the measurement value at the first time-point. This is advantageous as it corrects for cellular variation in the receptor distribution prior to ligand stimulation. Therefore the ligand induced change in receptor distribution is isolated (Fig. 3A). This approach can also be applied to colocalization measurements as shown in Figs. 6A, 8A, B and 9A.

The mean absolute change in percentage signal (between the first and last time-points) across all bands provides a single measurement per cell. This quantifies the magnitude of the, ligand induced, trafficking response (Fig. 3C). This summary measurement is useful as non-parametric hypothesis testing can be employed to identify statistically significant changes to the trafficking response between two conditions (typically a control and treatment) (Fig. 3C). However, information is lost by the reduction of the analysis to a single measurement per cell.

In summary, the proposed band-based analysis provides researchers with a powerful tool to identify, and quantify, perturbations to receptor trafficking in live cells. Moreover this approach to quantifying the subcellular distribution of a fluorophore is adaptable and could be applied to applications outside this field.

### 4.2. A critical review of Costes thresholding

Costes thresholding is an automated strategy for signal isolation in colocalization analysis (Section 2.2) [37]. Briefly, the linear line of best fit for the joint-histogram of the data is found, and the point on this line below which the data has a Pearson coefficient of less than zero is used to define global threshold values. This strategy is cited in review articles as good practice [14,25], and implemented in many popular image processing applications including Fiji and Imaris (Bitplane AG, Zurich, Switzerland) [49]. Adler and Parmryd (2013) note that a Costes approach will fail if there is no correlation in the data, and Dunn et al. (2011) suggest the approach may fail if the SNR is too low, the labelling density is too high, or if there are too many structures in one channel [14,23]. We also note that with a Costes approach there is an assumption that the data is well represented by a single linear fit, and therefore has strong correlation. In Fig. 5 B we show that a Costes approach has non-zero fail rate when tested on simulated data when there is either no colocaliza-



**Fig. 7.** Unbiased visualisation of colocalization. (A) Snapshot of a M1 based map of co-occurrence which has been visualised by 3D rendering. The surface of the cellular ROI is shown in green. (B) Maximal projections of representative cells (20 min post EGF treatment) for the control, AG1478 and Dynasore treatments respectively. Raw unprocessed data for the EGFR-EGFP and rab5-mRFP channels are shown in the first two rows respectively. The third row shows the normalised M1 based map of co-occurrence. Each map has been scaled by the corresponding M1 calculation to facilitate visual comparison. Scale bar set at 5  $\mu\text{m}$ . (C) Joint-histograms for the corresponding cells and time-point (logarithmic scale used). Only pixels within the cellular ROI were included. The line of best fit is shown in white. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tion, or anti-colocalization. This analysis was performed on a very simple test data set, where the intensities of spots in both channels were non-varying and equal (Appendix A). The Costes approach fails under these conditions as the gradient of the line of best fit can be less than or equal to zero. Therefore we conclude that a Costes approach is appropriate only if it can be assumed before the study that well correlated colocalization is present. Since the Costes approach is used to calculate measures of colocalization this should clearly not be assumed in the majority of cases.

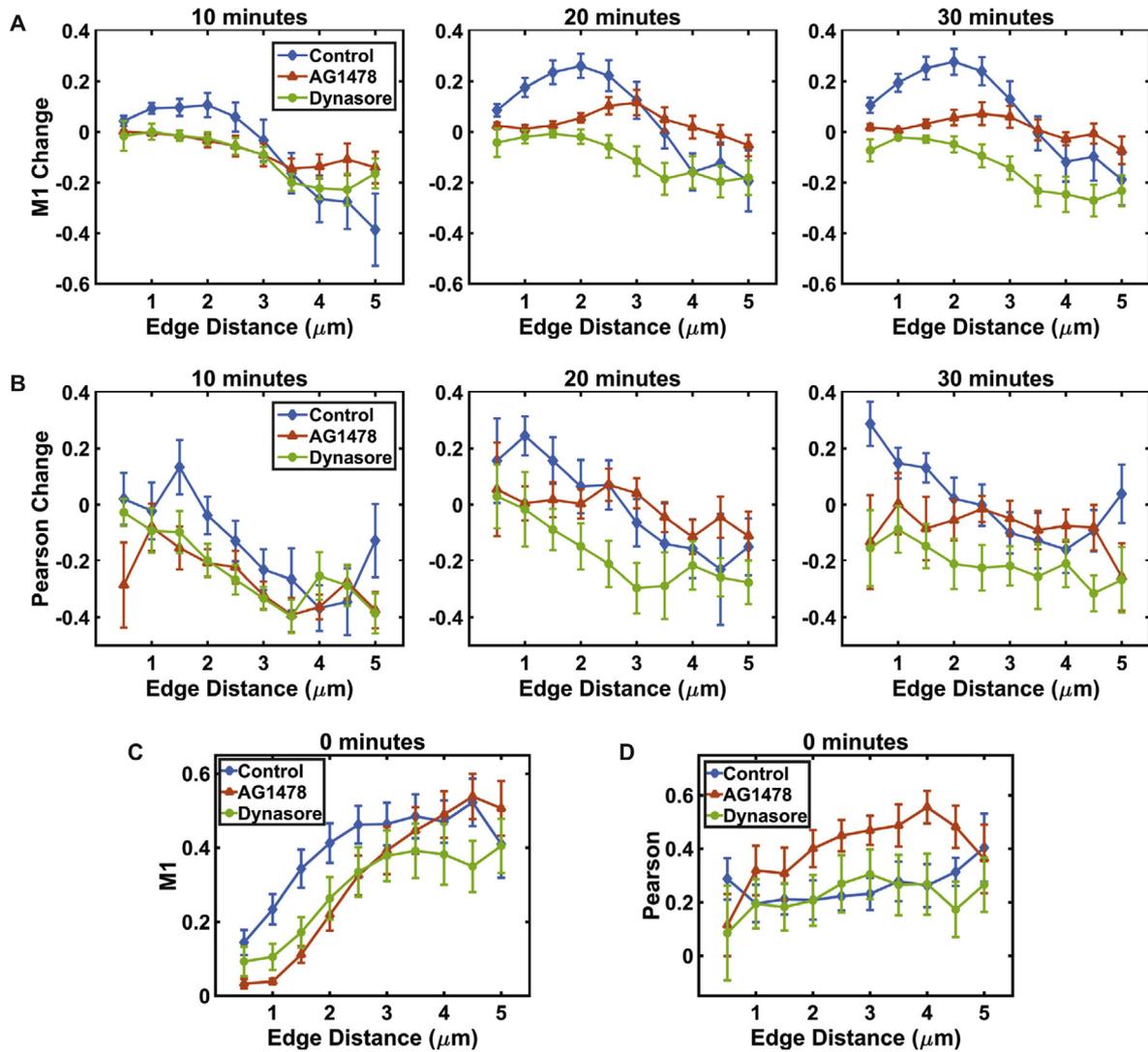
Consider the specific case of receptor colocalization with endosomal sub-populations. In such a case there are three key structures; membrane bound receptor, receptor positive endosomes and receptor negative endosomes. Each of the structures will have very different levels of correlation. Therefore the joint histogram representation will not be well represented by a single linear fit, and a Costes approach is not applicable (Fig. 10). Moreover it should not be assumed that there will be correlation between receptor and early endosomes in the drug treated populations or for all time-points. Note from Fig. 4 that the Costes approach fails to reliably isolate signal in the rab5 channel, when compared to manually set threshold values.

We consider well defined pre-processing of data, followed by global thresholding strategies, such as Otsu or minimum cross

entropy thresholding, to be a more appropriate strategy for signal isolation in colocalization studies [48,52]. Fig. 4 demonstrates that Otsu thresholding outperforms Costes thresholding on the real dataset. For the simulated data, Fig. 5C demonstrates that when calculated on data pre-processed with a Gaussian filter, the Otsu approach outperforms the Costes approach for all tested levels of colocalization and noise parameters. The Otsu approach performs substantially better at higher noise levels when performed on pre-processed data. This emphasises the importance of the pre-processing steps. Finally note, we are not proposing that the Otsu method will be appropriate for all applications and datasets. Our conclusion is that custom design and testing of pre-processing and thresholding steps is necessary to reliably segment biologically relevant signal for colocalization analysis. Importantly the developed strategy should be automated and applied consistently to all conditions to avoid user bias and variability.

#### 4.3. Quantitative colocalization analysis

In Section 3.3, a workflow was presented to quantify colocalization for 3D time-lapse data. Both the Manders (M1 and M2) and Pearson coefficients are calculated (Fig. 4A–C). Note the Pearson coefficient is calculated only across pixels containing isolated sig-



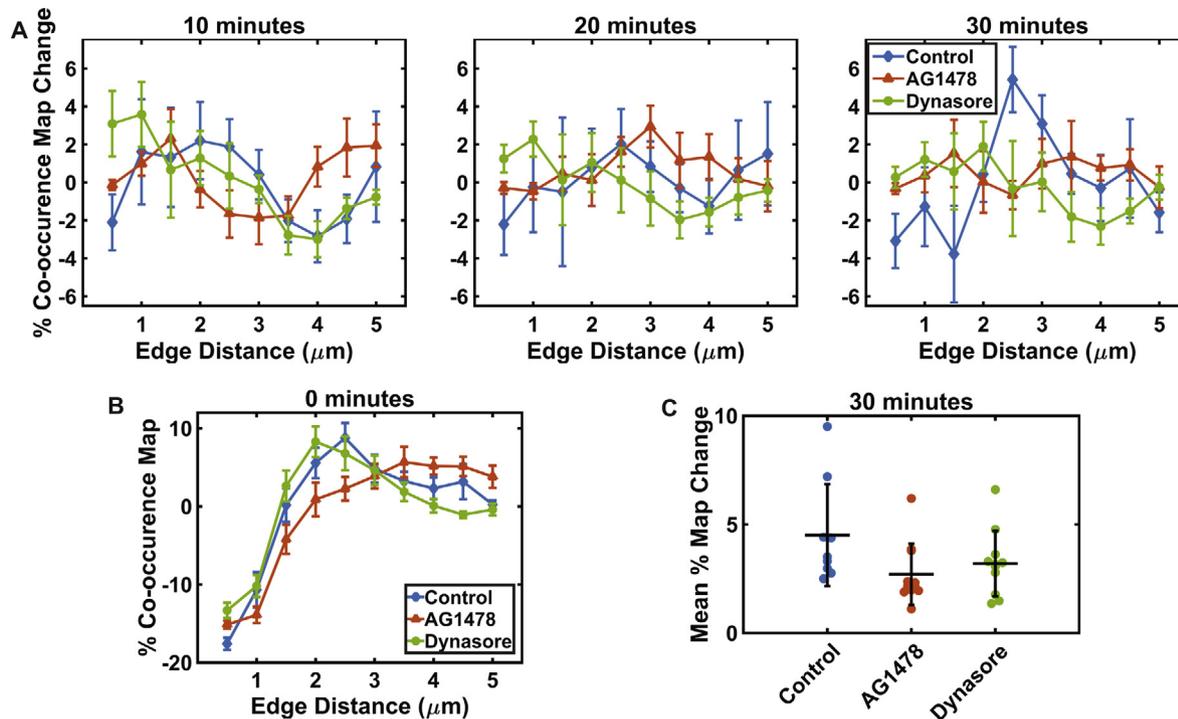
**Fig. 8.** Band based analysis using colocalization measures. (A) Plot of the change in the first Manders coefficient, M1, over time, for each banded volume (up to 5  $\mu\text{m}$ ), in response to EGF treatment. The data for the DMSO control and the two drug treatments are shown (AG1478 and Dynasore). Error bars given by the SEM and  $n > 10$  for all treatments. (B) Plot of the change in the first Pearson coefficient, R, over time, for each banded volume, in response to EGF treatment. (C, D) Plots of M1 and the Pearson coefficient for each banded volume at the first time-point.

nal in both channels. Recall this pairing of colocalization measures is very informative, and has a clear interpretation; the Manders coefficients quantify signal co-occurrence, and the Pearson coefficient quantifies the correlation within the co-occurring signal. For the test data used in this study, M1 represents the percentage of EGFR signal overlapping with rab5 positive endosomes, and is therefore the more relevant of the two Manders coefficients.

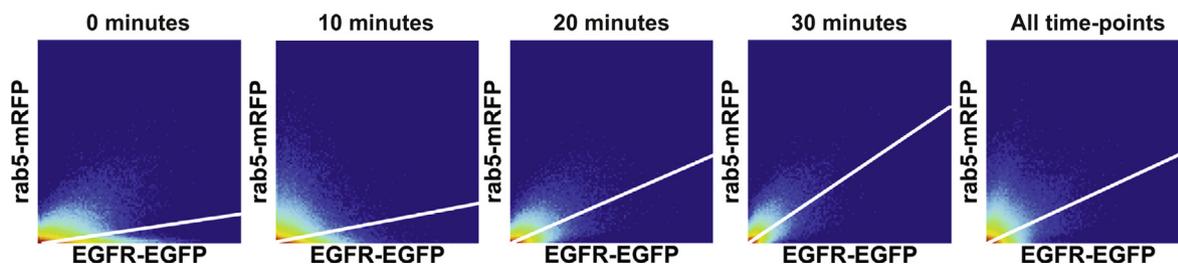
Non-parametric statistical testing can identify statistically significant differences in either co-occurrence, or correlation, between conditions (Fig. 4D–F). In this approach the statistical analysis is performed across multiple, as opposed to individual, ROIs which is advantageous as population based conclusions are essential. We consider the null hypothesis that the ligand induced change in co-occurrence, or correlation, is the same in the control and treated samples. The null hypothesis that the signal is randomly distributed within the cellular ROI is not considered. There are two reasons for this; firstly we consider rejecting the former null-hypothesis to be more informative and useful for the study of receptor trafficking. Secondly, the technical complications of auto-correlation and accurate ROI detection must be considered to reject the latter null-

hypothesis. For example the nuclear region should be isolated and removed from calculations. The proposed approach is therefore robust, unbiased and comparatively simple to implement.

In Section 3.6 the band based analysis from Section 3.2 is used to characterise the subcellular distribution of colocalization. Two distinct approaches were taken. Firstly, the level of co-occurrence and correlation was calculated for each band independently (Fig. 8). Secondly, the distribution of total cellular co-occurrence was characterised (Fig. 9). These analyses were included to demonstrate how the workflows presented in this paper can be adapted and combined for new applications. To clarify the difference between these two approaches consider a simple example where a cell is split into just two bands, with equal volume. The outer band contains 90% of the signal from channel 1 which has 50% overlap with channel 2 signal. The inner band only has 10% of channel 1 signal but 100% overlap with channel 2 signal. The first approach would return  $M1 = 0.5$  and  $M1 = 1$  for the outer and inner bands respectively, but this could be misleading as 82% of the co-occurring signal is contained in the outer band. This distribution of co-occurring signal is calculated by the second approach.



**Fig. 9.** Band based spatio-temporal analysis of subcellular co-occurrence distribution. (A) Plots of the change in (volume corrected) percentage co-occurrence map signal over time, for each banded volume (up to 5  $\mu\text{m}$ ), in response to EGF treatment. The data for the DMSO control and the two drug treatments are shown (AG1478 and Dynasore). Error bars given by the SEM and  $n > 10$  for all treatments. (B) Plot showing the (volume corrected) subcellular distribution of the co-occurrence map for the first time-point. (C) Plot of mean (absolute) change in (volume corrected) percentage co-occurrence map between the first and final time-points (30 min). A Kruskal-Wallis one-way analysis of variance returned  $p = 0.07$ , indicating that there are no statistically significant differences between treatments. Therefore no post-hoc testing was performed.



**Fig. 10.** Receptor and endosomal two channel data is not well represented by a single linear fit. Joint histograms (logarithmic scale) for a representative cell expressing EGFR-EGFP and rab5-mRFP. Immediately prior to imaging the cell was treated with EGF. Only pixels within the cellular ROI were included in the plots. The line of best fit for each plot is shown in white. Note the gradient of the linear fit changes over time as the receptor is internalised. Therefore, if a Costes thresholding approach is used the ratio between the two threshold values will also change. Also note that for the first two time-points, where the majority of the receptor is membrane localised, the data is weakly correlated. This is also the case if all time-points are combined to produce a single joint-histogram.

#### 4.4. Unbiased visualisation of colocalization

Colocalization is typically visualised using colour overlays (Section 2.2). This approach is open to bias through the choice of display parameters and contrast. Joint-histograms provide an effective visualisation of correlation, but do not preserve spatial information. In Section 3.4 we introduced a spatial map for the visualisation of co-occurrence. Importantly, this allows for unbiased visualisation of the colocalizing signal, where the spatial information is preserved. Using 3D rendering techniques, researchers can interactively view the co-occurrence in 3D (Fig. 5A). However 2D images are often required, and as an alternative to colour overlays, a maximal projection (or single slice) of the co-occurrence map can be used. The use of maximal projections can be misleading. If there is more than one pixel with co-occurring signal along the projection axis then only the maximal value will be displayed. Therefore visualisation in 3D is preferable. Note that these visualisation strategies should be performed in parallel with,

and not as an alternative to, quantitative analysis (Section 3.4). Finally note that, as with all visualisation strategies, the choice of the representative cell can bias the interpretation, emphasising the importance of quantitative analysis of a population.

## 5. Conclusion

All of the workflows introduced in this study are unbiased and described in sufficient detail as to be reproducible. These two key requirements should always be fulfilled by an image analysis protocol. Custom image analysis solutions are typically required for the analysis of subcellular signal distribution and colocalization. Therefore the aim of this work was not to provide step by step protocols, which researchers should follow exactly. Instead, illustrative examples were used to demonstrate the implementation of custom workflows which should be adapted by researchers for different datasets. The described protocols are specifically designed

for 3D live cell time-lapse data, and the statistical analysis is constructed to identify differences between treated and control samples. Such an approach is robust and ideal for confirmatory studies from larger screens. For example, the workflows presented in this work could be used to validate and further investigate hits from a RNAi screen for regulators of endocytosis and trafficking [53,54]. Finally, we note that the workflows could easily be adapted for applications other than receptor trafficking.

## Acknowledgments

This work was supported by the EPSRC through a studentship to JP from the PSIBS Doctoral Training Centre (EP/F50053X/1). The authors also acknowledge the Birmingham Advanced Light Microscopy (BALM) facility and Abdullah Khan. Data supporting this research is openly available from the University of Birmingham data archive at <http://findit.bham.ac.uk/>.

## Appendix A. Supplementary materials and methods

### A.1. Cell culture and transfection

The HeLa cell line was cultured in DMEM (with 4.5 g/l L-glutamine) (Invitrogen) supplemented with 10% FBS (Biosera) and 1% Penicillin-Streptomycin solution (Invitrogen) at 37 °C and 5% CO<sub>2</sub>. The EGFR-EGFP construct was a gift from Alexander Sorokin (University of Pittsburgh, USA) and the rab5-mRFP construct was a gift from Ari Helenius (Institute of Biochemistry, ETH Zurich). 24 h prior to transfection,  $2 \times 10^6$  cells/well were seeded into 35 mm glass bottomed dishes (MaTek) with antibiotic free media. Transient dual transfection of both constructs was performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions (10  $\mu$ l of Lipofectamine 2000 and 2  $\mu$ g of both constructs were used per well). 3 h after the addition of the constructs the media was changed (antibiotic free). Cells were imaged 24 h after expression.

### A.2. Drug treatment and confocal microscopy

Prior to imaging, cells were washed with 1 ml DPBS (Lonza) and serum starved for 30 min in serum/antibiotic free media (37 °C and 5% CO<sub>2</sub>). Media was then replaced with 2.5 ml cell imaging media (10 mM HEPES–Hank's balanced salt solution (Sigma), pH 7.4) and cells were treated with either 5  $\mu$ g/ml AG1478 (Sigma) (1  $\mu$ l/ml DMSO:MeOH 1:1 vehicle), 80  $\mu$ M Dynasore hydrate (Sellek Chemicals) (1  $\mu$ l/ml DMSO vehicle) or 1  $\mu$ l/ml DMSO (Sigma). Cells were then incubated for a further 30 min before transfer to the microscope where cells were treated with 2 ng/ml EGF (Bachem) immediately prior to initiating the time-lapse. Confocal laser scanning microscopy was performed with an inverted microscope (Eclipse Ti, Nikon A1R) at 37 °C and 5% CO<sub>2</sub> using a 1.49 NA 100X NA oil-immersion objective. Note a water immersion lens, which closer matches the refractive index of living cells, could be used to reduce artefacts and spherical aberration. An argon-ion 457–514 nm laser was used to excite the EGFR-EGFP construct and a green diode 561 nm laser was used for the rab5-mRFP construct. Images were acquired, or the data was scaled down, such that the pixel size was 0.25  $\mu$ m. Z-stacks with 0.3  $\mu$ m slice spacing were acquired every 10 min for a total duration of 30 min. Each stack took approximately 3 min to acquire.

### A.3. K-means based segmentation estimate

To produce the segmentation estimate for the initialisation of the level set function, the data from the EGFR-EGFP channel was

**Table A.1**

Parameters for generation of simulated colocalization data using the Colocalization Simulator plugin from the software application Icy. The colocalization percentage was varied and three noise levels were generated (low/medium/high).

Parameter	Value
Sequence width/height	256
Number of slices	50
Number of spots (for both channels)	100
Minimum spot intensity (for both channels)	20
Maximum spot intensity (for both channels)	20
Mean colocalization distance	0
Std colocalization distance	0
Maximum number of iterations	50
Poisson noise (low/medium/high)	0/2/4
Mean Gaussian noise (low/medium/high)	0/2/4
Std Gaussian noise	1

**Table A.2**

Parameters for the distance regularized level set evolution (DRLSE) segmentation protocol. The full details of the protocol can be found in Li et al. (2010) [43]. A stopping condition is employed such that the segmentation protocol is stopped if the volume contained within the zero level set changes less than 0.01% between two iterations. For the 3D extension the axial and spatial derivatives are scaled using the pixel size and axial spacing.

Parameter	Value
Edge term constant, $\lambda$	8
Constant force term, $\alpha$	0
DRLSE constant, $\mu$	0.15
Binary step size, $c_0$	2
Update step, $\Delta t$	1
Dirac delta width, $\epsilon$	1.5
Stopping condition	0.01%
Maximum number of iterations	50

clustered using a K-means approach (4 clusters) [55]. The 3 clusters with the highest means were combined and subsequently dilated, filled and eroded with a spherical structural element (with a 1.5  $\mu$ m radius). Holes in each axial slice were filled separately.

### A.4. Generation of synthetic data for colocalization analysis

To produce the simulated data with positive, or no, colocalization the Colocalization Simulator plugin (version 0.0.1.2), from the open image analysis package Icy, was used [50,51]. To generate the anti-colocalized data the Colocalization Simulator plugin was modified. First, the spots in both channel 1 and 2 were randomly placed. If a spot from channel 2 was less than a set distance (20 pixels) from a channel 1 spot, then it was randomly re-positioned. This re-positioning was repeated until all anti-colocalized spots from channel 2 were at least the minimum distance from all spots in channel 1. Finally, the remaining spots from channel 2 (neither colocalized or anti-colocalized) were randomly distributed. The plugin parameters for both the colocalized and anti-colocalized data were set as specified in Table A.1. 50 stacks were generated per colocalization percentage and noise level.

## References

- [1] S. Jones, J.Z. Rappoport, Interdependent epidermal growth factor receptor signalling and trafficking, *Int. J. Biochem.* 51 (2014) 23–28.
- [2] A. Sorokin, M. von Zastrow, Endocytosis and signalling: intertwining molecular networks, *Nat. Rev. Mol. Cell Biol.* 10 (2009) 609–622.
- [3] R. Irannejad, N.G. Tsvetanova, B.T. Lobingier, M. von Zastrow, Effects of endocytosis on receptor-mediated signaling, *Curr. Opin. Cell Biol.* 35 (2015) 137–143.
- [4] R. Villaseñor, H. Nonaka, P. Del Conte-Zerial, Y. Kalaidzidis, M. Zerial, Regulation of EGFR signal transduction by analogue-to-digital conversion in endosomes, *Elife* 4 (2015) e06156.

- [5] A. Fortian, A. Sorkin, Live-cell fluorescence imaging reveals high stoichiometry of Grb2 binding to the EGF receptor sustained during endocytosis, *J. Cell Sci.* 127 (2014) 432–444.
- [6] D. Teis, N. Taub, R. Kurzbauer, D. Hilber, M.E. de Araujo, M. Erlacher, M. Offterdinger, A. Villunger, S. Geley, G. Bohn, Others, p14-MP1-MEK1 signaling regulates endosomal traffic and cellular proliferation during tissue homeostasis, *J. Cell Biol.* 175 (2006) 861–868.
- [7] C.L. Howe, W.C. Mobley, Signaling endosome hypothesis: a cellular mechanism for long distance communication, *J. Neurobiol.* 58 (2004) 207–216.
- [8] G. Di Guglielmo, P. Baass, W. Ou, B. Posner, J. Bergeron, Compartmentalization of SHC, GRB2 and mSOS, and hyperphosphorylation of Raf-1 by EGF but not insulin in liver parenchyma, *EMBO J.* 13 (1994) 4269.
- [9] B.D. Grant, J.G. Donaldson, Pathways and mechanisms of endocytic recycling, *Nat. Rev. Mol. Cell Biol.* 10 (2009) 597–608.
- [10] A. Schenck, L. Goto-Silva, C. Collinet, M. Rhinn, A. Giner, B. Habermann, M. Brand, M. Zerial, The endosomal protein Appl1 mediates Akt substrate specificity and cell survival in vertebrate development, *Cell* 133 (2008) 486–497.
- [11] L. Munsie, A. Milnerwood, P. Seibler, D. Beccano-Kelly, I. Tatarnikov, J. Khinda, M. Volta, C. Kadgien, L. Cao, L. Tapia, et al., Retromer-dependent neurotransmitter receptor trafficking to synapses is altered by the Parkinson's disease VPS35 mutation p. D620N, *Hum. Mol. Genet.* 24 (2015) 1691–1703.
- [12] A. Tomas, C.E. Futter, E.R. Eden, EGF receptor trafficking: consequences for signaling and cancer, *Trends Cell Biol.* 24 (2014) 26–34.
- [13] G. Auciello, D.L. Cunningham, T. Tatar, J.K. Heath, J.Z. Rappoport, Regulation of fibroblast growth factor receptor signalling and trafficking by Src and Eps8, *J. Cell Sci.* 126 (2013) 613–624.
- [14] K.W. Dunn, M.M. Kamocka, J.H. McDonald, A practical guide to evaluating colocalization in biological microscopy, *Am. J. Physiol. Cell Physiol.* 300 (2011) C723–C742.
- [15] P. Sarder, A. Nehorai, Deconvolution methods for 3-D fluorescence microscopy images, *IEEE Signal Process. Mag.* 23 (2006) 32–45.
- [16] M.A. Lemmon, J. Schlessinger, K.M. Ferguson, The EGFR family: not so prototypical receptor tyrosine kinases, *CSH Perspect. Biol.* 6 (2014) a020768.
- [17] H. Stenmark, Rab GTPases as coordinators of vesicle traffic, *Nat. Rev. Mol. Cell Biol.* 10 (2009) 513–525.
- [18] A. Levitzki, A. Gazit, Tyrosine kinase inhibition: an approach to drug development, *Science* 267 (5205) (1995) 1782–1788.
- [19] E. Macia, M. Ehrlich, R. Massol, E. Boucrot, C. Brunner, T. Kirchhausen, Dynasore, a cell-permeable inhibitor of dynamin, *Dev. Cell* 10 (2006) 839–850.
- [20] J.-M. Hsu, C.-T. Chen, C.-K. Chou, H.-P. Kuo, L.-Y. Li, C.-Y. Lin, H.-J. Lee, Y.-N. Wang, M. Liu, H.-W. Liao, et al., Crosstalk between Arg 1175 methylation and Tyr 1173 phosphorylation negatively modulates EGFR-mediated ERK activation, *Nat. Cell Biol.* 13 (2011) 174–181.
- [21] A. Buades, B. Coll, J.-M. Morel, A review of image denoising algorithms, with a new one, *Multiscale Model. Simul.* 4 (2005) 490–530.
- [22] C. Kervrann, C.O.S. Sorzano, S.T. Acton, J.-C. Olivo-Marin, M. Unser, A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy, *IEEE J. Sel. Top. Signal Process.* 10 (2016) 6–30.
- [23] J. Adler, I. Parmryd, Colocalization analysis in fluorescence microscopy, *Methods Mol. Biol.* 931 (2013) 97–109.
- [24] E. Abbe, Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung, *Arch. Mikrosk. Anat.* 9 (1873) 413–418.
- [25] S. Bolte, F. Cordelieres, A guided tour into subcellular colocalization analysis in light microscopy, *J. Microsc.* 224 (2006) 213–232.
- [26] T. Lagache, N. Sauvonnnet, L. Danglot, J.-C. Olivo-Marin, Statistical analysis of molecule colocalization in bioimaging, *Cytometry A* 87 (6) (2015) 568–579.
- [27] R.B. Sekar, A. Periasamy, Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations, *J. Cell Biol.* 160 (2003) 629–633.
- [28] M. Miaczynska, S. Christoforidis, A. Giner, A. Shevchenko, S. Uttenweiller-Joseph, B. Habermann, M. Wilm, R.G. Parton, M. Zerial, APPL proteins link Rab5 to nuclear signal transduction via an endosomal compartment, *Cell* 116 (2004) 445–456.
- [29] M. Ward, A. Cunningham, Developmental expression of vascular endothelial growth factor receptor 3 and vascular endothelial growth factor C in forebrain, *Neuroscience* 303 (2015) 544–557.
- [30] A. Frederick, H. Yano, P. Triflief, H. Vishwasrao, D. Biezonski, J. Mészáros, E. Urizar, D. Sibley, C. Kellendonk, K. Sonntag, et al., Evidence against dopamine D1/D2 receptor heteromers, *Mol. Psychiatry* (2015).
- [31] B.D. Ripley, The second-order analysis of stationary point processes, *J. Appl. Probab.* (1976) 255–266.
- [32] J. Adler, I. Parmryd, Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient, *Cytometry A* 77 (2010) 733–742.
- [33] E. Manders, F. Verbeek, J. Aten, Measurement of co-localization of objects in dual-colour confocal images, *J. Microsc.* 169 (1993) 375–382.
- [34] E. Manders, J. Stap, G. Brakenhoff, R. Van Driel, J. Aten, Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy, *J. Cell Sci.* 103 (1992) 857–862.
- [35] V.R. Singan, T.R. Jones, K.M. Curran, J.C. Simpson, Dual channel rank-based intensity weighting for quantitative co-localization of microscopy images, *BMC Bioinformatics* 12 (2011) 407.
- [36] F. Humpert, I. Yahiatene, M. Lummer, M. Sauer, T. Huser, Quantifying molecular colocalization in live cell fluorescence microscopy, *J. Biophotonics* 8 (2015) 124–132.
- [37] S.V. Costes, D. Daelemans, E.H. Cho, Z. Dobbin, G. Pavlakis, S. Lockett, Automatic and quantitative measurement of protein-protein colocalization in live cells, *Biophys. J.* 86 (2004) 3993–4003.
- [38] J.H. McDonald, K.W. Dunn, Statistical tests for measures of colocalization in biological microscopy, *J. Microsc.* 252 (2013) 295–302.
- [39] F. Luisier, C. Vonesch, T. Blu, M. Unser, Fast interscale wavelet denoising of Poisson-corrupted images, *Signal Process.* 90 (2010) 415–427.
- [40] D. Sage, D. Prodanov, J.-Y. Tinevez, J. Schindelin, MIJ: making interoperability between ImageJ and Matlab possible, *ImageJ User & Developer Conference*, 2012, pp. 24–26.
- [41] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, *Nat. Methods* 9 (2012) 671–675.
- [42] J.B. Pawley, Sources of noise in three-dimensional microscopical data sets, *Three-Dimensional Confocal Microscopy: Volume Investigation of Biological Specimens*, 2012, pp. 47–94.
- [43] C. Li, C. Xu, C. Gui, M.D. Fox, Distance regularized level set evolution and its application to image segmentation, *IEEE Trans. Image Process.* 19 (2010) 3243–3254.
- [44] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, *Int. J. Comput. Vision* 22 (1997) 61–79.
- [45] M. Maška, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbiola, T. España, S. Venkatesan, D.M. Balak, et al., A benchmark for comparison of cell tracking algorithms, *Bioinformatics* 30 (2014) 1609–1617.
- [46] S.R. Sternberg, Biomedical image processing, *IEEE Comput.* 16 (1983) 22–34.
- [47] Y. Mishchenko, A fast algorithm for computation of discrete Euclidean distance transform in three or more dimensions on vector processing architectures, *Signal Image Video Process.* 9 (2015) 19–27.
- [48] N. Otsu, A threshold selection method from gray-level histograms, *Automatica* 11 (1975) 23–27.
- [49] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (2012) 676–682.
- [50] F. De Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, et al., Icy: an open bioimage informatics platform for extended reproducible research, *Nat. Methods* 9 (2012) 690–696.
- [51] N. Chenouard, Advances in probabilistic particle tracking for biological imaging (thesis), 2011.
- [52] C.H. Li, C. Lee, Minimum cross entropy thresholding, *Pattern Recognit.* 26 (1993) 617–625.
- [53] C. Collinet, M. Stöter, C.R. Bradshaw, N. Samusik, J.C. Rink, D. Kenski, B. Habermann, F. Buchholz, R. Henschel, M.S. Mueller, et al., Systems survey of endocytosis by multiparametric image analysis, *Nature* 464 (2010) 243–249.
- [54] P. Kozik, N.A. Hodson, D.A. Sahlender, N. Simecek, C. Soromani, J. Wu, L.M. Collinson, M.S. Robinson, A human genome-wide screen for regulators of clathrin-coated vesicle formation reveals an unexpected role for the V-ATPase, *Nat. Cell Biol.* 15 (2013) 50–60.
- [55] S.P. Lloyd, Least squares quantization in PCM, *IEEE Trans. Inf. Theory* 28 (1982) 129–137.



# Super-resolution for everybody: An image processing workflow to obtain high-resolution images with a standard confocal microscope



France Lam<sup>a</sup>, Damien Cladière<sup>b</sup>, Cyndélia Guillaume<sup>a</sup>, Katja Wassmann<sup>b</sup>, Susanne Bolte<sup>a,\*</sup>

<sup>a</sup> Sorbonne Universités, UPMC Univ Paris 06, CNRS, Core Facilities - Institut de Biologie Paris Seine (IBPS), 75005 Paris, France

<sup>b</sup> Sorbonne Universités, UPMC Univ Paris 06, CNRS, Laboratory of Developmental Biology - Institut de Biologie Paris Seine (IBPS), 75005 Paris, France

## ARTICLE INFO

### Article history:

Received 27 July 2016

Received in revised form 2 November 2016

Accepted 4 November 2016

Available online 5 November 2016

### Keywords:

Confocal microscopy

Deconvolution

Refractive index matching

3D-SIM-like resolution improvement

Meiotic spindle

## ABSTRACT

In the presented work we aimed at improving confocal imaging to obtain highest possible resolution in thick biological samples, such as the mouse oocyte. We therefore developed an image processing workflow that allows improving the lateral and axial resolution of a standard confocal microscope. Our workflow comprises refractive index matching, the optimization of microscope hardware parameters and image restoration by deconvolution. We compare two different deconvolution algorithms, evaluate the necessity of denoising and establish the optimal image restoration procedure. We validate our workflow by imaging sub resolution fluorescent beads and measuring the maximum lateral and axial resolution of the confocal system. Subsequently, we apply the parameters to the imaging and data restoration of fluorescently labelled meiotic spindles of mouse oocytes. We measure a resolution increase of approximately 2-fold in the lateral and 3-fold in the axial direction throughout a depth of 60  $\mu\text{m}$ . This demonstrates that with our optimized workflow we reach a resolution that is comparable to 3D-SIM-imaging, but with better depth penetration for confocal images of beads and the biological sample.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

During the past ten years one of the most fundamental challenges in 3D imaging was to break the diffraction barrier of light in optical microscopy to attain super-resolution images [1]. These new super-resolution techniques may be based on the precise localization of fluorochromes, such as PALM/STORM, which needs high power excitation and numerous acquisitions to collect enough photons. On the other hand, non-linear fluorophore responses such as Stimulated Emission Depletion (STED) help increasing the resolution. STED is based on the selective deactivation of fluorophores by a high power depletion laser, minimising the area of illumination at the focal point, and thus enhancing the achievable resolution for a given system. Sample preparation in PALM/STORM and STED-microscopy demands for dyes with specific characteristics and fluorochrome environment, especially if multicolour labelling is needed. Another major drawback is that 3D-information is difficult to achieve and the depth penetration of these techniques is limited to a few  $\mu\text{m}$  for PALM/STORM and below 20  $\mu\text{m}$  for STED.

3D-Structured Illumination Microscopy (3D-SIM), a technique based on illumination patterning has recently emerged as a

practicable approach to circumvent the diffraction limit of light, avoiding complex sample preparation protocols. 3D-SIM requires the acquisition of at least 15 images per plane and the image contrast has to be good enough to extract high frequencies. 3D-SIM allows obtaining a spatial resolution of 100–150 nm laterally and 250–350 nm axially, thus improving resolution by a factor of approximately two compared to confocal microscopy [1–3]. This technique has enabled high resolution imaging of the cellular components such as the microtubule (MT) network in cells and organisms [4,5]. However, standard 3D-SIM is limited to an imaging depth of 10–20  $\mu\text{m}$  [6] and therefore difficult to achieve in thicker samples. The reasons for this are the inherent out-of-focus background of this wide-field approach. Furthermore, bulky samples do not allow obtaining a good fringe contrast necessary for the reconstruction of the 3D-SIM-image. Using two-photon SIM, the depth penetration limit has been broken enabling super-resolution imaging in depth beyond 45  $\mu\text{m}$  [5,7]. Yet, 3D-SIM and 2-photon SIM are methods that call for specialised equipment and expertise, and are not accessible to every biologist. PALM/STORM, STED and 3D-SIM require image reconstruction/processing as a final step and can be applied in a restricted manner to live imaging [1]. However, even though these techniques have been democratized recently by the emergence of commercial solutions they ask for high technical expertise and specific hardware, they are time consuming and display inherently high photo-toxicity [1].

\* Corresponding author.

E-mail address: [susanne.bolte@upmc.fr](mailto:susanne.bolte@upmc.fr) (S. Bolte).

On the other hand, confocal microscopy, a valuable tool for optical sectioning of thick samples has become a standard technique accessible to every biologist. A conventional confocal microscope may reach resolutions of 200 nm laterally and 500 nm axially, if properly adjusted [8]. Furthermore, image restoration by deconvolution, a powerful computational process used to reduce out-of-focus light in 3D-fluorescent images, can be combined with confocal imaging. This combination has been proven useful to enhance image contrast, improve the signal-to-noise ratio and resolution [9–11]. However, deconvolution of confocal data is still scarcely used. This might be due to the fact that deconvolution requires expertise and a neat measurement of the impulse response of the optical system; the point spread function (PSF). When focussing deep into a biological specimen the PSF is distorted by spherical aberrations due to refractive index mismatches [12], and might then not be precise enough to perform high fidelity deconvolution. Besides, scattering is also an issue in thick and optically dense samples. To counteract the negative influence of the distorted PSF on the deconvolution results, one might use a PSF taken directly from the biological image, if the structures investigated have the appropriate shape [13]. It is also possible to compute theoretical PSF [14]. Alternatively, adaptive optics could be used to minimize the distortion of the wave front in depth of the sample [15]. Another possibility to avoid spherical aberration and thus distortion of the PSF would be refractive index matching of the sample [16]. We have recently shown that accurate refractive index matching of the sample and the mounting medium improves considerably the axial resolution and depth penetration in fixed brain tissue [17]. We furthermore showed that the fixed biological tissue has an overall refractive index that is close to that of the optical system (1.518), rendering the tissue very transparent. This minimizes scattering and the PSF measured after focussing up to 120  $\mu\text{m}$  into this sample was not impaired by spherical aberration and thus distortion-free.

In the present study, we aim at applying this knowledge on the mouse oocyte, a rather bulky sample with a diameter of nearly 100  $\mu\text{m}$ . Our scientific project aims at improving microtubule (MT) imaging at high resolution in the meiotic spindle. High resolution imaging of the meiotic spindle will be important to tackle outstanding questions regarding the influence of the structural organisation of the MTs network on meiotic division in the future. Microtubules of the meiotic spindle are hollow tubes of 25 nm in diameter, consisting of 13 protofilaments composed of alpha- and beta-tubulin heterodimers. These polymers are highly dynamic and stochastically switch between polymerisation and depolymerisation, a process known as dynamic instability [18]. MT dynamics and structural organisation is tightly regulated by a plethora of MT-associated proteins (MAPs) [19]. To understand how the precise organisation of the MT network contributes to essential cellular processes such as chromosome segregation during meiosis, a detailed reconstruction and analysis of the spatial distribution of MTs as well as their length, orientation and anchoring to kinetochores in oocytes is decisive. However, conventional fluorescence microscopy of these polymers and their associated proteins in mammalian oocytes is limited by light diffraction and by the thickness of the sample. Visualizing the attachment of individual microtubule fibres on each sister kinetochore is beyond the resolution limit of classical confocal imaging.

In the present paper, we combine refractive index matching, optimization of imaging parameters and deconvolution of confocal data. Several commercial solutions have been recently developed to improve lateral and axial resolution of confocal imaging. Leica's Hyvolution and Olympus' FV-OSR solution both use a standard confocal microscope, customized with optimized acquisition and detection system (pinhole closure and high sensitivity, low-noise

detector) coupled to image restoration. A different approach is used in the Zeiss' Airyscan configuration where all photons are collected from the volume excited with concentrically arranged hexagonal detector array. Its detection area consists of 32 single detector elements, acting as very small pinholes. The signals from these detector elements are reassigned to their correct position, producing an image with increased signal-to-noise ratio and resolution.

We propose a workflow similar to the solution from Leica and Olympus, with the difference, that we adapt the mounting medium by refractive index matching with the biological sample. This matching decreases the impact of the spherical aberration and improves thus axial resolution. Our workflow allows to improve the lateral and axial resolution of the confocal microscope, gaining resolutions comparable to 3D-SIM-imaging, but with a much better depth penetration. In a first step, we characterize the resolution of our optical system by acquiring sub resolution fluorescent beads using this workflow. We then demonstrate that this optimized workflow improves considerably the resolution of biological imaging when applied to the meiotic spindle of mouse oocytes.

## 2. Materials and methods

### 2.1. Oocyte culture and whole-mount immunofluorescence

Oocytes were obtained from ovaries of adult Swiss mice 10–16 weeks old (Janvier, France) and cultured in self-made M2 medium, as previously described [20]. Oocytes entering meiosis I, visible through Germinal Vesicle Breakdown (GVBD), within 60 min after harvesting were used for experiments. Before using oocytes for fixation, the zona pellucida was removed by incubation in Tyrode's solution [21].

Between 6 h and 6h30 after GVBD, zona pellucida-free oocytes were placed into self-made chambers in order to manipulate them easily. Oocytes were cytospinned in chambers coated with concanavalin A (at 0.2 mg/ml, from Sigma) to keep oocytes in place, for 13 min at 1400 RPM at 38 °C. Cold-stable spindles were obtained by incubating the individual chambers on an ice-water bath for 4 min before fixation. Oocytes were fixed with 2% Formaldehyde solution (Sigma-Aldrich; F1635) in BRB80 buffer with Triton TX-100 during 30 min at 38 °C, as described [22]. After PBS washing, oocytes were incubated overnight in a PBS-BSA 3% Triton TX-100 solution for permeabilisation and blocking unspecific antibody binding.

Antibody staining of fixed oocytes was done using the antibodies below at the indicated dilutions in PBS: Human CREST autoimmune antibody (Cellon SA, HCT-0100; 1/100), and mouse monoclonal  $\alpha$ -tubulin (DM1A) coupled to FITC (Sigma-Aldrich, F2168; 1/100). As secondary antibody CY3 anti rabbit (Jackson ImmunoResearch, 711-166-152; 1/200) was used. Antibody incubation times were at least one hour. Chromosomes were stained with 5  $\mu\text{g}/\text{ml}$  Hoechst 33342 during incubation with the secondary antibody. Oocytes were washed several times with PBS between antibody incubations and before mounting the chamber [22].

### 2.2. Mounting media

We used AF1 (Citifluor, UK), a commercially available mounting medium with a refractive index of 1.463 and AF1+, a modified AF1 solution harbouring a refractive index of 1.518. The refractive index increase of AF1+ solution was obtained by adding 83% (w/w) of Methyl-Phenylsulfoxid (Sigma-Aldrich, #261696) to AF1-solution. Refractive indices were verified at 21 °C using a refractometer (Mettler Toledo, Switzerland).

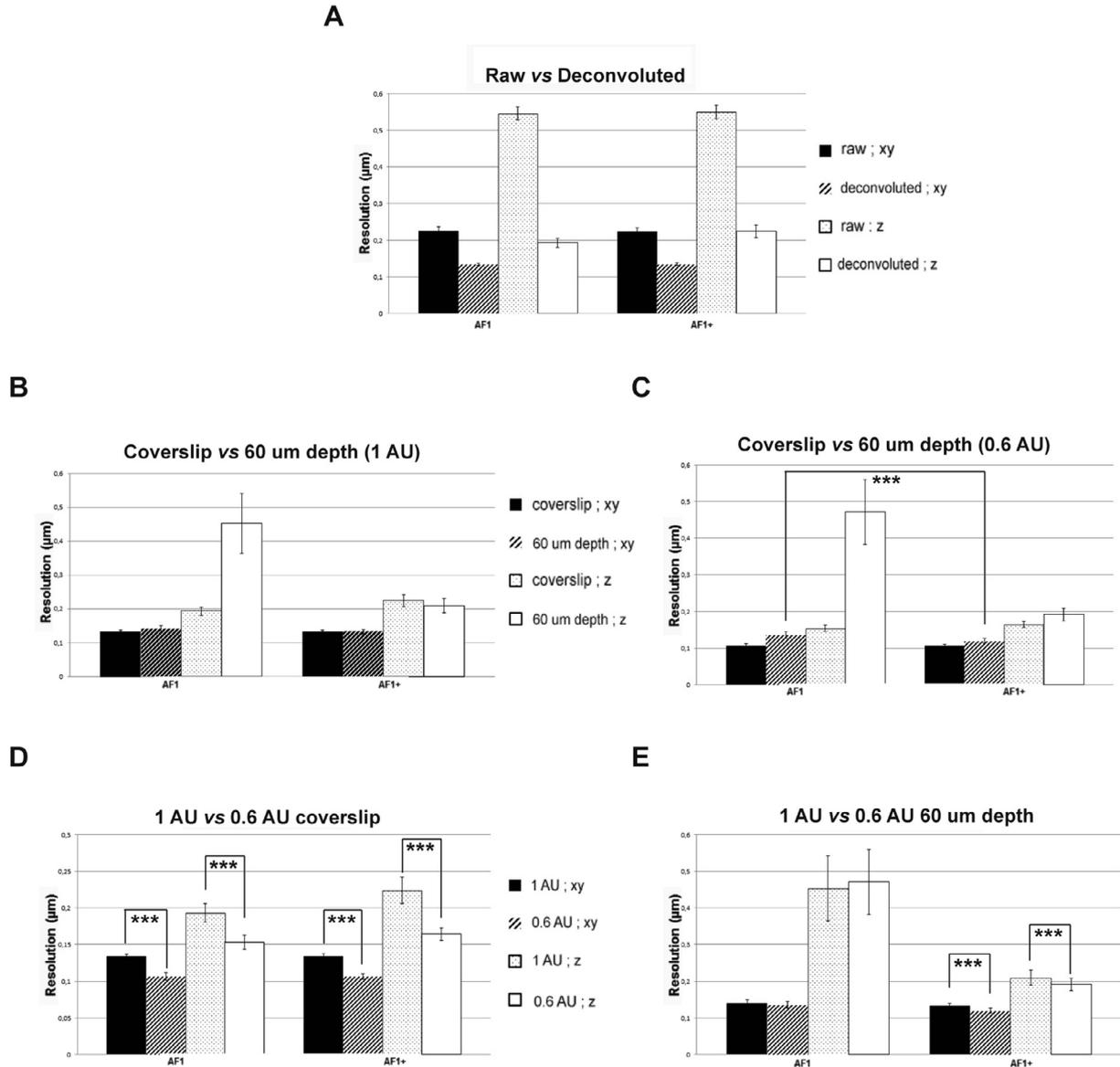
2.3. Preparation of Sub-resolution fluorescent beads

The experimental setup for evaluating lateral and axial resolution at the coverslip and in depth was designed as followed: fluorescent beads (PS-Speck, Lifetechnologies), of a diameter of 100 nm, and loaded with yellow-green fluorescent dye, were diluted in water (1/800 v/v). Drops of the water-diluted sample were put on the surface of the coverslip (Menzel Glaeser #1.5, Agar scientific) or slide and air-dried. The coverslip and slide were

separated by one layer of adhesive tape (Scotch®, 3 M) with a nominal thickness of 60 μm and the volume filled with a drop of the respective mounting medium.

2.4. Confocal laser scanning microscopy

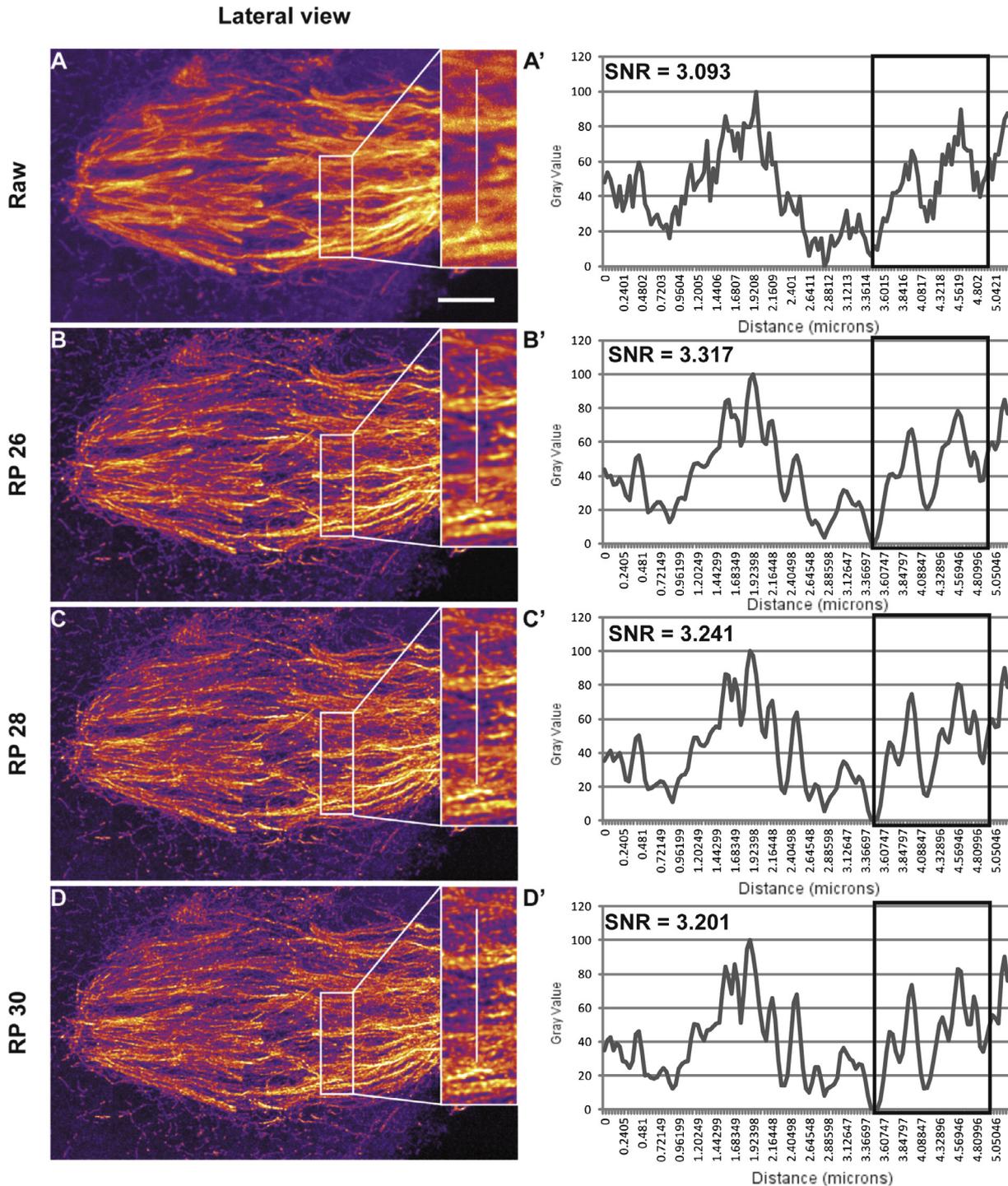
8-bit Images were collected using a Leica 63× oil immersion objective (HCX Plan APO CS, NA 1.4, working distance 0.14 mm) with an inverted Leica laser-scanning confocal microscope TCS



**Fig. 1.** Resolution measurements with sub-resolution microspheres: influence of deconvolution, refractive index and pinhole aperture. 100 nm green fluorescent beads were used to measure the lateral and axial resolution of the optical system at the coverslip and in 60 μm depth. For excitation, the 488 nm laser line was used and two series with the detection pinhole set to 1 Airy Unit (AU) and 0.6 AU were acquired. The Full Width at Half Maximum (FWHM) was measured using a line profile in x, y-direction and x, z-direction in the maximum intensity plane of the image stack, respectively and transferred to a graph (A–E). All data in brackets are given as mean ± SD and have the unit nm. (A) Raw bead images (black and dotted bars) and deconvoluted bead images (striped and white bars) taken at the coverslip were compared. Deconvolution improved significantly the lateral (raw: 225 ± 12; decon: 134 ± 3) and axial resolution (raw: 546 ± 18; decon: 193 ± 13) of beads in AF1. Deconvolution improved also significantly the lateral (raw: 223 ± 11; decon: 134 ± 3) and axial resolution (raw: 550 ± 19; decon: 224 ± 18) of beads in AF1+-medium. (B and C) Lateral and axial resolution of deconvoluted bead images at the coverslip (black and dotted bars) and in 60 μm depth (striped and white bars) at 1 AU (B) and 0.6 AU (C). Resolution measurements in depth show a significant axial resolution loss in AF1-medium at 1 AU (coverslip: 193 ± 13; depth: 453 ± 89) and at 0.6 AU (coverslip: 182 ± 10; depth: 471 ± 89). (D and E) Comparison of the resolution of beads taken at 1 Airy unit (black and dotted bars) and 0.6 Airy unit (striped and white bars) at the coverslip (D) and in 60 μm depth (E). Resolution measurements at the coverslip (D) show a significant lateral resolution improvement when closing the pinhole to 0.6 AU in AF1-medium (1 AU: 134 ± 3; 0.6 AU: 120 ± 5, Mann-Whitey test, p = 0.0001), and a substantial axial resolution increase (1 AU: 193 ± 13; 0.6 AU: 182 ± 10, Mann-Whitey test, p = 0.0001). A significant lateral resolution improvement at 0.6 AU is also observed in AF1+-medium (1 AU: 134 ± 3; 0.6 AU: 106 ± 4, Mann-Whitey test, p = 0.0001) and axial resolution improves substantially, too (1 AU: 210 ± 21; 0.6 AU: 192 ± 17, Mann-Whitey test, p = 0.0001). In depth (E), no resolution improvement is observed laterally (1 AU: 140 ± 10; 0.6 AU: 136 ± 9) or axially (1 AU: 453 ± 89; 0.6 AU: 471 ± 89) in AF1-medium. However, when using AF1+-medium, resolution improves significantly in the lateral direction (1 AU: 133 ± 6; 0.6 AU: 120 ± 7, Mann-Whitey test, p = 0.0001) and in the axial direction (1 AU: 210 ± 21; 0.6 AU: 192 ± 17, Mann-Whitey test, p = 0.0003).

SP5 II (Leica Microsystems, Heidelberg, Germany) equipped with a GaAsP hybrid detection system at a sampling rate of 40 nm in x,y- and 83 nm in z-direction unless otherwise stated. Fluorochromes were detected using laser lines 405 nm, 488 nm and 561 nm. Imaging was performed in a temperature-controlled room at 21 °C. The detection pinhole aperture was adjusted to 1 Airy unit or 0.6 Airy units. Slow scanning speeds may improve the signal-to-noise ratio and the resolution of the image since more photons may be

collected by pixel. The cutoff value has to be determined for every biological sample, since slower scan speed results in higher photo bleaching. We tested scan speeds of 400, 200, 100 and 50 Hz. For the given biological sample the optimal scan speed was 400 Hz. We also tested Immersion oils with refractive indexes from 1.510 to 1.520 to minimize spherical aberration (supp. Fig. 2). We used an immersion oil harbouring a refractive index of 1.518 for all image acquisitions.



**Fig. 2.** Determination of optimal deconvolution parameters, R-Parameter. Mouse oocytes stained with FITC-tubulin and imaged at 488 nm with the detection pinhole adjusted to 0.6 Airy Unit and taken with a 40 nm pixel size and an 83 nm z step size were deconvoluted varying the R-parameter between 26, 28 and 30. Images show x, y-maximum projections of the raw data (A), and the deconvoluted data with R-parameter 26 (B), R-parameter 28 (C), and R-parameter 30 (D). Inserts show the position of the line profile used for fluorescence intensity measurements shown in the graphs next to each image (A'–D'). S/N-ratios indicated in (A'–D') are a value for the image intrinsic noise and were calculated as described in (2.8). The boxes indicate the regions of interest, where 5 main peaks can be separated after deconvolution. Note that R-parameter of 30 gave best results. Scale bar is 5  $\mu$ m.

2.5. Image acquisition of beads

Bead images were obtained as in [17] with the following modifications. GaAsP gain was set to 15% and the laser power adjusted so that the signal occupied the full dynamic range of the detector, but saturated voxels were carefully avoided. Beads were imaged starting and finishing the stack at least 5 μm below and above the bead centre. Beads were visually checked and improper stacks were discarded before determining the microscope PSF.

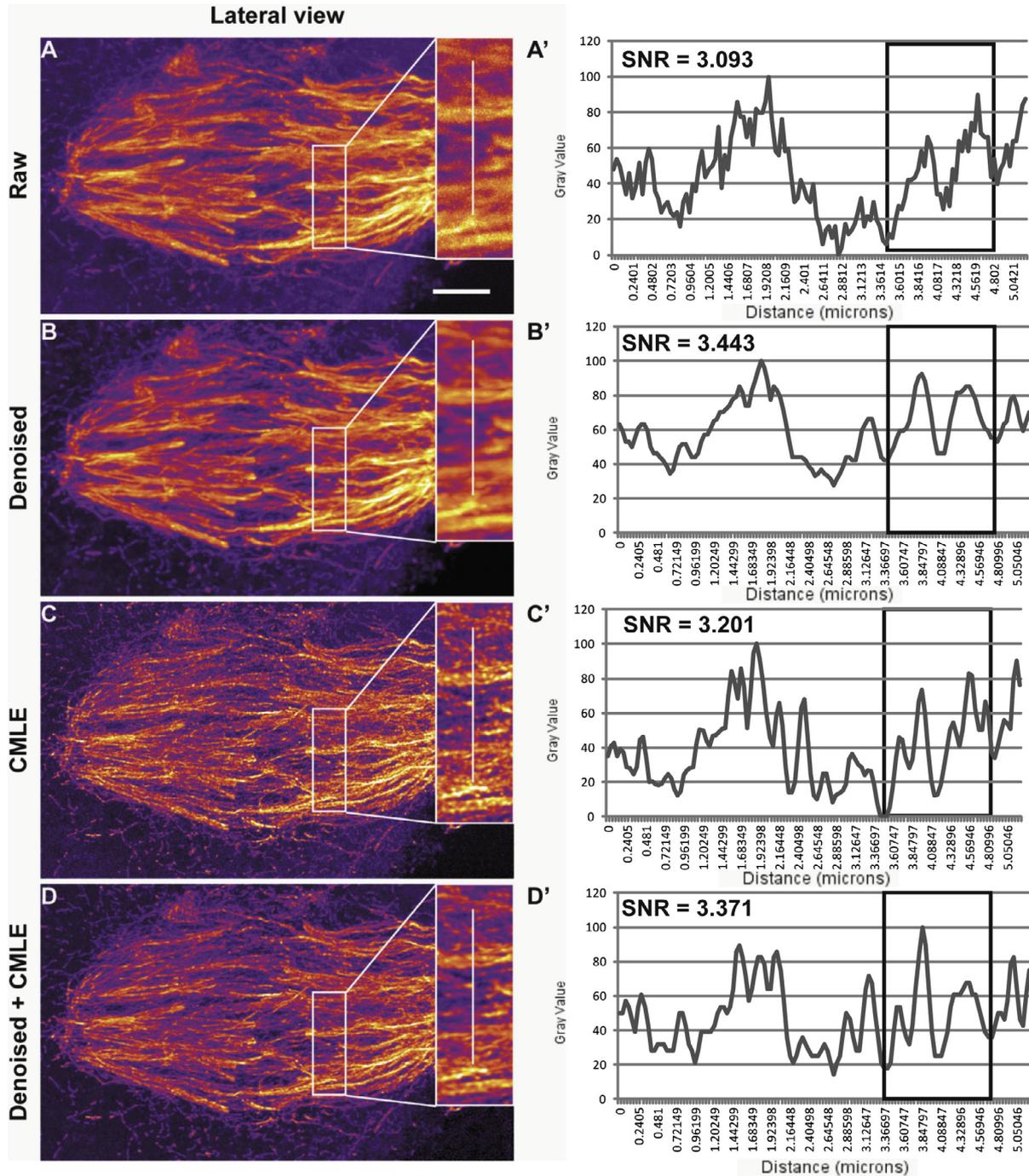
Beads used for resolution measurements (Fig. 1) were acquired at a sampling rate of 32 nm in x, y and 42 nm in z-direction. An

average of 60 beads was taken for each measurement. Beads used for deconvolution were acquired at a sampling rate of 40 nm in x,y and 83 nm in z-direction. At least 15 beads were registered and averaged in order to increase the SNR for deconvolution.

2.6. Resolution measurement with beads

Resolution measurements were carried out as in [18] using ImageJ with the following modifications:

We took the maximum intensity plane of the image stack along the x, y and z-axis resulting in 1D intensity profiles. A x,y- and

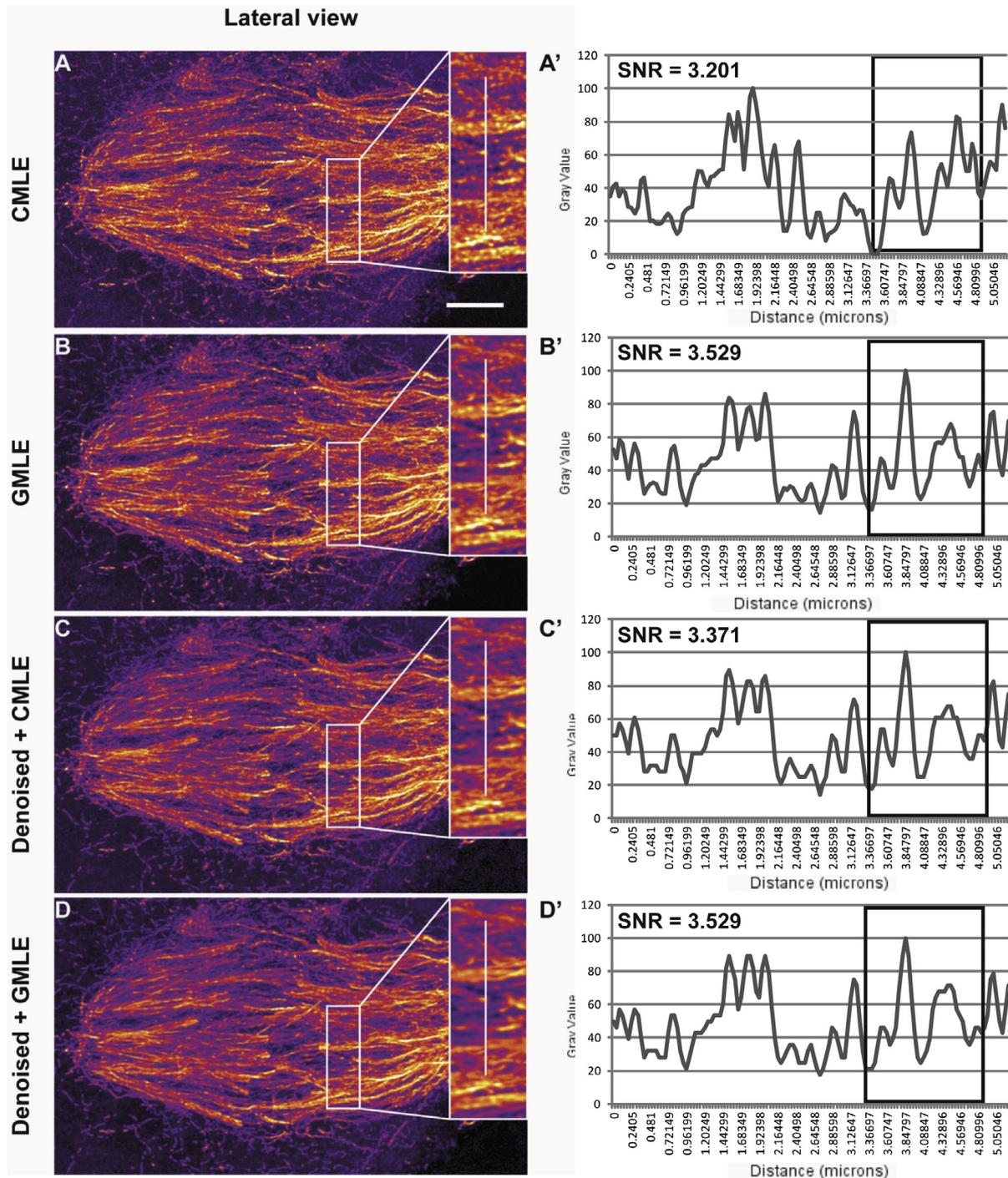


**Fig. 3.** Influence of denoising on the deconvolution result. Images show x, y-maximum projections of the raw data (A), the deconvolved data (B), the denoised data (C) and the denoised, deconvolved data (D). Deconvolution was carried using the CMLE-algorithm (R-parameter = 30, 500 iterations). Inserts show the position of the line profile used for fluorescence intensity measurements shown in the graphs next to each image (A'–D'). S/N-ratios are indicated in (A'–D'). The boxes indicate the regions of interest, where 5 main peaks can be separated after deconvolution. Note that denoising considerably improves the S/N-value of the raw data. Deconvolved data and a combination of denoising and deconvolution gave comparable results. However, denoising improves the S/N-value of the deconvolved data. Scale bar is 5 μm.

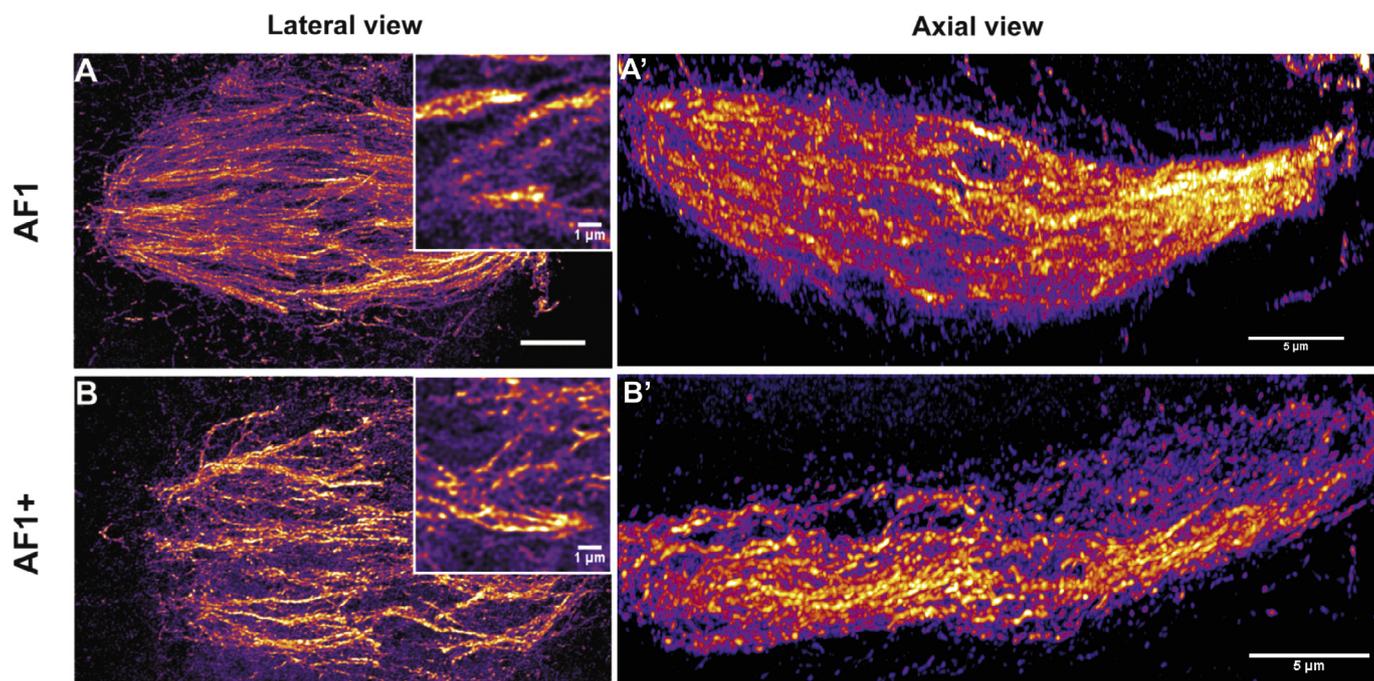
z-line-profile was generated manually along a line passing through the 2D maximum. From this image, fluorescence intensity profiles were fitted to a Gaussian curve, using ImageJ's built-in curve fitting function. The full width at half maximum (FWHM) of the gaussian curve was calculated for each profile, based on the parameters retrieved from the fitting (supp. Fig. 3A).

## 2.7. Deconvolution and image treatment

Confocal images of beads and biological data were deconvoluted with the Huygens 3.7 software (Scientific Volume Imaging, Hilversum, Netherlands) using a measured PSF (supp. Fig. 3B). We tested the Classical Maximum Likelihood Estimation (CMLE)



**Fig. 4.** Comparison of deconvolution algorithm CMLE and GMLE. Images show data deconvoluted with two different deconvolution algorithms implemented in the Huygens software, the Classical Maximum Likelihood Estimation (CMLE) and the Good's Maximum Likelihood Estimation (GMLE). Deconvolution was carried with the R-parameter set to 30 for both methods. Images show x, y-maximum projections of the CMLE-deconvoluted data (A), the GMLE deconvoluted data (B) the denoised, CMLE deconvoluted data (C), and the denoised, GMLE deconvoluted data (D). Inserts show the position of the line profile used for fluorescence intensity measurements shown in the graphs next to each image (A'–D'), the boxes indicate the position of the 5 main peaks compared before. Note that denoising improves the S/N of the CMLE-method, but not of the GMLE-method. The GMLE-algorithm gives the best visual result. Scale bar is 5  $\mu\text{m}$ .



**Fig. 5.** Impact of refractive index matching in mouse oocytes. Images show x, y-maximum projections (A, B, lateral view), a cut-out of a single plane in a depth of 10  $\mu\text{m}$  (insert of A, B) and x, z-projections (A', B') of mouse oocytes stained with FITC-tubulin and imaged at 488 nm in AF1-medium (A, A') and AF1+-medium (B, B'). All data in brackets are given as mean  $\pm$  SD and have the unit nm. Microtubule bundle diameters were measured in single images of raw data (images not shown) and deconvolved data (raw data/deconvolved data;  $390 \pm 131$  /  $218 \pm 45$ ). A 2-fold amelioration of resolution was noted after deconvolution. Lateral resolution of microtubule bundles in AF1 and AF1+-medium was comparable (AF1/AF1+;  $232 \pm 33$  /  $204 \pm 57$ ). Axial resolution of microtubule bundles was measured in AF1 and AF1+-medium (AF1/AF1+;  $402 \pm 10$  /  $7/292 \pm 49$ ). Note that refractive index matching with AF1+ improves considerably the axial resolution (A', B'). Scale bars are 5  $\mu\text{m}$ .

algorithm and the Good's (GMLE) algorithm. Signal-to-noise ratios, also called the R-parameter from 15 to 20 are recommended by the manufacturer for noisy confocal images R-values  $>20$  for low noise wide-field images. Since we used a confocal microscope with a low noise detector, we tested R-parameters of 26–34 and visually inspected the results. In images of Fig. 2–6 brightness and contrast were adjusted equally for all images after deconvolution and before 3D reconstruction and volume rendering. Image analysis was performed with ImageJ [23] and 3D-volume rendering was performed with ICY [24]. Deconvolution was performed with a measured PSF, with a quality threshold of 0.001, varying R-parameters as stated, performing 500 iterations in the optimized mode with CMLE-method and 125 iterations with the GMLE-method. Denoising was performed using the Pure Denoise Plugin [25–27] implemented in ImageJ with the cycle spin and the multi-frame parameter set to maximum.

### 2.8. Signal to noise ratio calculation

To calculate the SNR, we choose a threshold with the Shanbhag mode on ImageJ (supp. Fig. 4). We created a mask to measure the average intensity of our structure and with the inverted thresholded image; we create a second mask to measure the average intensities of the background. We finally calculate the SNR as average intensity/average background.

## 3. Results and discussion

### 3.1. Resolution measurement using sub-resolution microspheres

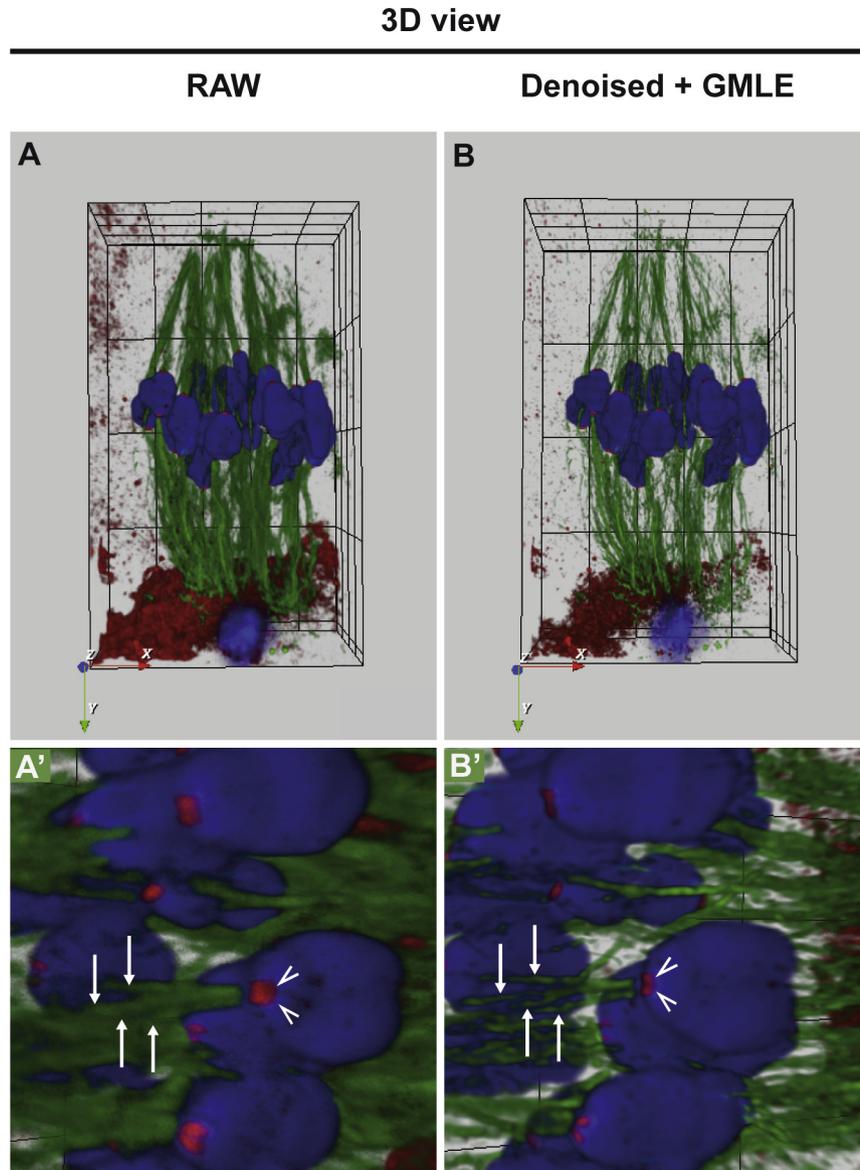
We first aimed at developing a workflow, consisting of optimized sample preparation, image acquisition and deconvolution

parameters to obtain the best resolution with our standard confocal microscope. We therefore investigated the lateral and axial resolution of the confocal microscope, using a high numerical aperture lens ( $63\times$ , NA 1.4, oil immersion).

We determined the Point Spread Function (PSF) using two different mounting media, and varying the detection pinhole aperture. The PSF is the impulse response of the focused optical system and gives information about the lateral and axial resolution of the optical system. The PSF was measured by imaging 3D stacks of 100 nm fluorescent beads at different depths. The lateral and axial resolutions were estimated from the full-width at half-maximum (FWHM) of the intensity profile of the PSF. We compared beads mounted in AF1 with a refractive index of 1.463 with beads mounted in AF1+, a modified AF1, harbouring a refractive index of 1.518. In the first case we introduced a mismatch of the refractive index ( $n_i$ ) between the sample ( $n_i = 1.463$ ) and the glass-oil-interface ( $n_i = 1.518$ ) thus producing spherical aberrations when focussing away from the coverslip. In the second case, we perfectly matched the sample ( $n_i = 1.518$ ) with the glass-oil-interface ( $n_i = 1.518$ ), thus minimizing spherical aberrations.

#### 3.1.1. Resolution improvement by deconvolution

Firstly, we determined the resolution of the confocal microscope before and after deconvolution of confocal data of beads in the  $n_i$ -mismatched and  $n_i$ -matched case with the detector pinhole set at 1 Airy Unit (AU). The resolution of the optical system taken from the raw images was  $225 \text{ nm} \pm 12 \text{ nm}$  in x,y and  $546 \text{ nm} \pm 18 \text{ nm}$  in z-direction in AF1-medium and  $223 \text{ nm} \pm 11 \text{ nm}$  in x,y and  $550 \text{ nm} \pm 19 \text{ nm}$  in z-direction in AF1+-medium (Fig. 1A, Table 1). After deconvolution, we observed a 1.7-fold increase in lateral resolution for both mounting media, a 2.8-fold increase in axial resolution for AF1 and a 2.5-fold increase for AF1+ for beads



**Fig. 6.** 3D-reconstruction of data. Mouse oocytes were triple stained with Hoechst (chromosomes), FITC (microtubules) and CY3 (kinetochores). Images show 3D-renderings of raw data (A–A') and deconvoluted data (B–B'). Magnifications of a region of (A, B) with arrows indicating microtubules (A', B') and arrowheads showing kinetochores (A', B') show the resolution improvement after deconvolution, where separate microtubules and two adjacent kinetochores can clearly be separated.

imaged next to the coverslip (Fig. 1 A, Table 1). The resolution values for AF1 and AF1+-mounted beads were comparable, corroborating previous findings, that close to the coverslip, a refractive index mismatch does not impair lateral and axial resolution since high numerical aperture objectives are corrected for this [13,18]. Having established that deconvolution improves indeed the resolution of confocal data, we secondly analysed how the refractive index mismatch would impact resolution when looking at beads at a depth of 60  $\mu\text{m}$  (Fig 1B, deconvoluted data, Table 1). The lateral

resolution of the beads was comparable for the two mounting media, which confirmed that the lateral resolution in depth is not impaired in a mismatched system [12,17]. However, axial resolution was 2.2-times inferior in AF1-medium compared to AF1+-medium. Besides, we observed a considerable loss of peak intensity at a depth of 60  $\mu\text{m}$  in AF1-medium. We had to adjust the laser power accordingly in order to exploit the full dynamic range of the image. This confirms the data, we published recently for beads measured in CFM3, another high refractive index

**Table 1**  
Resolution measurement with fluorescent subresolution beads.

	AF1 coverslip		AF 60 $\mu\text{m}$ depth		AF1+ coverslip		AF1+ 60 $\mu\text{m}$ depth	
	1 AU (nm)	0.6 AU (nm)	1 AU (nm)	0.6 AU (nm)	1 AU (nm)	0.6 AU (nm)	1 AU (nm)	0.6 AU (nm)
xy (raw)	207	177	309	294	223	182	214	203
xy (deconvoluted)	121	105	140	136	134	106	133	120
z (raw)	416	386	1532	1383	550	436	509	467
z (deconvoluted)	167	151	453	470	224	164	210	192

mounting medium ( $n_i = 1.518$ ) where we observed the same phenomenon [17]. The axial resolution of beads remained constant in depth compared to the coverslip when applying refractive index matching with AF1+. This means that spherical aberration is minimized when using AF1+ medium, and our observation corroborated previous theoretical and experimental findings of Hell and co-workers [12] and our own findings for a perfectly matched system [17].

### 3.1.2. Resolution improvement by closing the pinhole aperture

We then wanted to test if we could increase resolution of the confocal microscope by closing the detection pinhole. We compared resolution at the coverslip and in a depth of 60  $\mu\text{m}$  with the detector pinhole set to 0.6 AU (Fig. 1C, Table 1) and after deconvolution of the data. Our choice of the 0.6 AU pinhole size was based on several tests on our biological data. We acquired the same type of biological sample with different pinhole sizes, from 1 AU to 0.4 AU and observed that 0.6 AU is the threshold where we discard enough diffraction signal without photo-bleaching and with a good contrast. Since the result depends largely on the quality and the photo-stability of the biological sample the optimal pinhole value has to be evaluated for each biological sample.

We observed lateral resolutions of 120 nm  $\pm$  5 nm (coverslip) and 136 nm  $\pm$  9 nm (depth) and axial resolutions of 182 nm  $\pm$  10 nm (coverslip) and 471 nm  $\pm$  89 nm (depth) for the AF1 medium (Fig. 1C, Table 1). We had again to compensate for the loss of peak intensity by adjusting the laser power with the AF1 medium as described before. For the AF1+-medium, lateral resolutions of 106 nm  $\pm$  4 nm (coverslip) and 120 nm  $\pm$  7 nm (depth) and axial resolutions of 164 nm  $\pm$  8 nm (coverslip) and 192 nm  $\pm$  17 nm (depth) were measured.

Fig. 1D and E shows the direct comparison of the optical resolution at 1 AU and 0.6 AU close to the coverslip (Fig. 1D, Table 1) and in 60  $\mu\text{m}$  depth (Fig. 1E, Table 1), respectively. We observed a 1.3-fold increase in lateral and axial resolution measured in AF1-medium close to the coverslip (Fig. 1D, Table 1) when setting the detection pinhole to 0.6 AU. The use of AF1+ medium improved resolution 1.3-fold laterally and 1.4-fold axially at AU 0.6 at the coverslip (Fig. 1D, Table 1). In depth, a significant lateral and axial resolution improvement was observed only for AF1+-medium at 0.6 AU. Besides, the use of AF1+ allowed maintaining a good axial resolution of 192 nm  $\pm$  17 nm that is 2.5-times superior than with AF1-medium (471 nm  $\pm$  89 nm), when closing the pinhole. It is thus advantageous to use refractive index matching when closing the detector aperture and imaging in depth.

Closing the pinhole indeed increased lateral and axial resolution 1.3–1.4-fold. These results are in good agreement with data measured by Cox and Sheppard [28], who observed a 1.4-fold increase in resolution after closing the pinhole aperture to 0.5 using a Leica TSC SP2 confocal microscope.

By optimizing sample preparation, image acquisition parameters and performing deconvolution, our workflow allowed us to obtain a considerable gain in lateral and axial resolution throughout the sample thickness. Firstly, the lateral resolution attained with 100 nm sub resolution fluorescent beads close to the coverslip attained a value of 106 nm  $\pm$  4 nm. This is close to the real size of the beads and comparable to the values published earlier for 3D-SIM-imaging [4]. Secondly, the axial resolution we obtained with our workflow was 164 nm  $\pm$  8 nm and is thus approximately 1.6-fold better than the values published for 3D-SIM by Gustafsson [4]. However, these results are in good agreement with a paper by Schrader and colleagues [29], who demonstrated earlier that a resolution of 80 nm could be obtained axially on 50 nm gold particles with a confocal microscope and after deconvolution. Given that we used 100 nm beads, our results are consistent with these previous findings. Thirdly, refractive index matching of the sample allowed

to maintain 3D-SIM-like resolution of 120 nm  $\pm$  7 nm laterally and 192  $\pm$  17 nm axially in a depth of 60  $\mu\text{m}$ . This depth surpasses by a factor 3 the imaging depth of standard 3D-SIM [6].

### 3.2. Deconvolution optimization in biological data

Having established optimal image acquisition parameters, we then wanted to apply our workflow on biological samples and chose the meiotic spindle of the mouse oocyte labelled with alpha-tubulin coupled with FITC. We acquired 3D-images of meiotic spindles with a pinhole size of 0.6 Airy units at a sampling rate of 40 nm in x,y and 83 nm in z-direction. We wanted to determine the optimal deconvolution parameters and tested the Classic Maximum Likelihood Estimation (CMLE) and the Good's roughness Maximum Likelihood Estimation algorithm (GMLE), implemented in Huygens deconvolution software. CMLE and GMLE are iterative restoration method optimizing the likelihood of an estimate of the object given the measured image and the PSF. We also studied the influence of denoising prior to deconvolution using the Pure Denoising plugin implemented in ImageJ [26–28]. Finally we compared raw data and data optimized by our workflow after 3D-reconstruction.

#### 3.2.1. Establishment of the optimal Regularisation parameter (R-parameter)

The Huygens software uses what the manufacturer calls the Signal-to-Noise ratio as a Regularization Parameter (R-parameter), i.e. as a parameter that controls the sharpness of the restoration result. We will employ the term R-parameter in the following in order to avoid confusion with the image intrinsic Signal-to-Noise ratio (SNR) that we calculated to evaluate image quality.

The R-parameter is calculated as the square root of the number of photons in the brightest part of the image and controls the sharpness of the restoration result. The higher the R-parameter, the sharper is the deconvolution result. R-parameter values recommended by the manufacturer for noise-prone confocal data lie between 15 and 20. Since we used a low-level-noise detector for confocal imaging, we tested R-parameter values between 16 and 34. After adjusting the R-parameter, we visually inspected the deconvolution result and observed that at an R-parameter >26, microtubules were separated optimally, when compared to the raw image (Fig. 2A–D). We also carried out fluorescence intensity profile measurements to obtain a measurement of the separation of tubulin structures in a single plane (Fig. 2A'–D4) and calculated the image inherent SNR before and after deconvolution to estimate image improvement. We noticed that microtubule separation attained a maximum, setting the R-parameter to 30 (Fig. 2A'–D'). The SNR increased from 3.1 in the raw image to 3.32 after deconvolution with R-parameter set to 26. We started separating 5 peaks in the region of interest at an R-parameter of 28 (Fig. 2C', box). At higher R-parameter, the SNR decreased slightly to 3.24 (R-parameter 28) and 3.20 (R-parameter 30), showing that the R-parameter introduces sharpness into the image. When using R-parameters >30 we started introducing deconvolution artefacts such as patterning, indicating that an R-parameter of 30 is optimal for our dataset (Fig. 1, Supplementary Data 1 A–C).

#### 3.2.2. Influence of denoising on the deconvolution of confocal data

Having established the optimal R-parameter, we now aimed at studying the impact of denoising before deconvolution. It has been proposed that denoising may be crucial before deconvolution and may improve the restoration result [30]. We therefore denoised the confocal data with the "Pure Denoising" plugin implemented in ImageJ [25–27]. This plugin is well adapted for Poisson noise, which is characteristic of fluorescence images. It estimates automatically noise parameters such as detector gain, detector offset

and the standard deviation of the additive white Gaussian noise. We adjusted the number of cycle-spins and the number of adjacent frames used to maximum values in order to get the highest quality for our images. Denoising with the quality criteria set to maximum was rather computing intense and took several hours for a 3D-image stack. Fig. 3 shows a comparison of the raw image (A, A'), the denoised image (B, B'), the deconvolution of the raw image (C, C') and the deconvolution of the denoised image (D, D'). When inspecting the fluorescence intensity profiles, denoising alone did not improve the resolution of the image (Fig. 3B, B'), even if it improved considerably the SNR. We observed a smoothing of the fluorescence intensity profile (Fig. 3B', box). We then compared the deconvoluted dataset (Fig. 3C, C', box) with the denoised and deconvoluted image (Fig. 3D, D', box). We observed 5 main peaks in the region of interest of the deconvoluted image (Fig. 3C'). Denoising improved SNR of the restoration result. After denoising we were still able to distinguish 5 peaks (Fig. 3D, D' and box), even if the indentation was less prominent than in the non-denoised dataset (Fig. 3C, C', box). The overall aspect of the image was smoother and as a result microtubules appeared less spotty after denoising (Fig. 3 D, insert).

### 3.2.3. Comparison of CMLE and GMLE algorithm for the deconvolution of confocal data

In a last step, we wanted to compare the CMLE-algorithm with the Goods Maximum Likelihood Estimation (GMLE) algorithm. The GMLE-methods needs 4-times less iterations than the CMLE-method to obtain comparable restoration results, thus saving computation time. We compared the image after deconvolution with the CMLE-algorithm (Fig. 4A, A') with that deconvoluted with the GMLE-algorithm (Fig. 4B, B'). After visual inspection, we observed that the GMLE-algorithm smoothed the deconvolution result more than the CMLE-algorithm. Both algorithms gave comparable results in separating the microtubules and 5 main peaks could be distinguished in both cases (Fig. 4A', B', box), however, the SNR was higher for the GMLE-algorithm. Denoising did improve the SNR of the image deconvoluted with CMLE-algorithm, however, the SNR after GMLE-deconvolution remained constant, indicating that denoising did not further improve image quality in this situation (Fig. 4D, D').

### 3.2.4. Influence of refractive index matching on the 3D-image resolution

We have shown previously, that refractive index matching greatly improves axial resolution of confocal data in the mouse brain [17]. We therefore mounted mouse oocytes in AF1 and AF1+, respectively and performed 3D imaging. Fig. 5 shows lateral and axial views of the tubulin staining and after deconvolution. We observed a 2-fold increase of lateral resolution when comparing raw data to deconvoluted data. When considering the lateral maximum projection, the overall aspect of the image in the two mounting media was comparable (Fig. 5A, B). We then measured the diameter of microtubule bundles throughout the depth of the image stacks and detected comparable lateral resolution for AF1 and AF1+ (Fig. 5A, A'). However, axial resolution seemed to be better in the ri-matched sample with AF1+ after visual inspection (Fig. 5B'). We measured axial resolutions and found indeed a 1.4-fold increase of resolution in AF1+-medium, meaning that refractive index matching is crucial in the biological sample. This confirms the data obtained with beads, where a comparable gain of axial resolution has been observed in depth using the ri-matched system (Fig. 1C) and our data on biological samples published earlier [17].

### 3.3. 3D-reconstruction of the meiotic spindle

Having established all optimal parameters, we applied them to acquire 3D-stacks of triple stained mouse oocytes. We evaluated the overall aspect of the meiotic spindle resolution. We used alpha-tubulin coupled to FITC, to mark microtubules, Crest, a protein staining the kinetochore, which was revealed with a secondary antibody coupled to Cy3, and Hoechst to stain the chromosomes. Fig. 6 shows the 3D rendering of the raw data (A) and the deconvoluted data (B) of the meiotic spindle. We observed a clear improvement in resolution and much better separated microtubules (Fig. 6B', arrows) and the kinetochores (Fig. 6B', arrowheads) as compared to the raw data (Fig. 6A', arrows and arrowheads). With these results, we provide evidence that our workflow, consisting of refractive index matching, optimized image acquisition and deconvolution improves greatly the overall resolution of the meiotic spindle image.

## 4. Conclusion

The proposed workflow provides an important benefit to confocal imaging of fixed samples with a high numerical aperture lens. We show first with beads that deconvolution improves by 1.7-fold lateral and by >2.6-fold the axial resolution of confocal data. Secondly, we provide evidence that refractive index matching is particular important to improve resolution and signal strength of the confocal system, when imaging in depth, where we obtain a 2.2-fold axial resolution improvement for the matched system. Thirdly, we show that closing the pinhole improves lateral and axial resolution by a factor 1.3 close to the coverslip and that the spatial invariance of the PSF in depth is crucial to preserve resolution in depth. We observe that deconvolution of the biological confocal data improves 2-fold the resolution compared to raw confocal data and that refractive index matching helps to maintain optimal axial resolution in the biological sample. We lastly show that denoising is advantageous for image quality when using the CMLE-algorithm, however, it is time consuming. Denoising is not necessary when using the GMLE-algorithm, which also needs much less computing time. Summarized from these results, we were able to obtain confocal images having a 3D-SIM-like resolution in a rather bulky sample, the mouse oocyte. It would be of further interest to apply our workflow on other biological samples and see if we are able to maintain a 3D-SIM-like resolution at the full working distance of the objective (140  $\mu\text{m}$ ).

## Acknowledgements

The authors thank Jean-François Gilles and Nicolas Heck for helpful discussions and Marc Dos Santos for precious help with statistical analysis and Adobe Illustrator. Work on mouse oocytes was supported by an ANR Grant (ANR-12-BSV2-0005-01) to KW. Confocal work was carried out at the Institute of Biology Paris-Seine Imaging Facility that is significantly (or: which receives significant support from...) supported by the "Conseil Regional Ile-de France", the French national research council (CNRS) and Sorbonne University, UPMC Univ Paris 06.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jymeth.2016.11.003>.

## References

- [1] L. Schermelleh, R. Heintzmann, H. Leonhardt, A guide to super-resolution fluorescence microscopy, *J. Cell. Biol.* 190 (2) (2010) 165–175. 26.
- [2] M.G. Gustafsson, Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy, *J. Microsc.* 198 (Pt 2) (2000) 82–87.
- [3] R.S. Fischer, Y. Wu, P. Kanchanawong, H. Shroff, C.M. Waterman, Microscopy in 3D: a biologist's toolbox, *Trends Cell Biol.* 21 (12) (2010) 682–691.
- [4] M.G. Gustafsson, L. Shao, P.M. Carlton, C.J. Wang, I.N. Golubovskaya, W.Z. Cande, D.A. Agard, J.W. Sedat, Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination, *Biophys. J.* 94 (12) (2008) 4957–4970.
- [5] P.W. Winter, A.G. York, D.D. Nogare, M. Ingaramo, R. Christensen, A. Chitnis, G. H. Patterson, H. Shroff, Two-photon instant structured illumination microscopy improves the depth penetration of super-resolution imaging in thick scattering samples, *Optica* 1 (3) (2014) 181–191. 20.
- [6] K. Nienhaus, G.U. Nienhaus, Where do we stand with super-resolution optical microscopy?, *J. Mol. Biol.* 428 (2 Pt A) (2016) 308–322. 29.
- [7] A.G. York, S.H. Parekh, D. Dalle Nogare, R.S. Fischer, K. Temprine, M. Mione, A.B. Chitnis, C.A. Combs, H. Shroff, Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy, *Nat. Methods* 9 (7) (2012) 749–754. 13.
- [8] J.B. Pawley, Fundamental Limits of Confocal Microscopy, *Handbook of Biological Confocal Microscopy*, third edition, Springer, New York, Chapter 2, Pages 20–42.
- [9] M.B. Cannell, A. McMorland, C. Soeller, Image enhancement by deconvolution, in: *Handbook of Biological Confocal Microscopy*, third ed., Springer, New York, 2006, pp. 488–500. Chapter 25.
- [10] P.J. Shaw, Comparison of widefield/deconvolution and Confocal microscopy for Three-Dimensional Imaging, in: *Handbook of Biological Confocal Microscopy*, third ed., Springer, New York, 2006, pp. 453–467. Chapter 23.
- [11] N. Heck, S. Betuing, P. Vanhoutte, J. Caboche, A deconvolution method to improve automated 3D-analysis of dendritic spines: application to a mouse model of Huntington's disease, *Brain Struct. Funct.* 217 (2) (2012) 421–434.
- [12] A. Egner, S.W. Hell, Aberrations in confocal and multi-photon fluorescence microscopy induced by refractive index mismatch, in: *Handbook of Biological Confocal Microscopy*, third ed., Springer, New York, 2006, pp. 404–412.
- [13] M. Von Tiedemann, A. Fridberger, M. Ulfendahl, I. Tomo, J. Boutet de Monvel, Image adaptive point-spread function estimation and deconvolution for in vivo confocal microscopy, *Microsc. Res. Tech.* 69 (1) (2006) 10–20.
- [14] A. Dieterlen, C. Xu, M.-P. Gramain, O. Haerberle, B. Colicchio, C. Cudel, S. Jacquy, E. Ginglinger, G. Jung, E. Jeandidier, Validation of image processing tools for 3-D fluorescence microscopy, *Comptes Rendus Biologies Elsevier Masson* 325 (1) (2002) 327–334.
- [15] M. Booth, D. Andrade, D. Burke, B. Patton, M. Zurasukas, Aberrations and adaptive optics in super-resolution microscopy, *Microscopy (Oxf.)* 64 (4) (2015) 251–261.
- [16] T. Staudt, M.C. Lang, R. Medda, J. Engelhardt, S.W. Hell, 2,2'-thiodiethanol: a new water-soluble mounting medium for high-resolution optical microscopy, *Microsc. Res. Tech.* 70 (1) (2007) 1–9.
- [17] C. Fouquet, J.-F. Gilles, N. Heck, M. Dos Santos, R. Schwartzmann, V. Cannaya, M.-P. Morel, R.S. Davidson, A. Trembleau, S. Bolte, *PLoS ONE* (2015), <http://dx.doi.org/10.1371/journal.pone.0121096>.
- [18] M.W. Kirschner, T. Mitchison, Microtubule dynamics, *Nature* 324 (1986) 621.
- [19] R. Alfaro-Aco, S. Petry, Building the microtubule cytoskeleton piece by piece, *J. Biol. Chem.* 290 (28) (2015) 17154–17162.
- [20] S.A. Touati, E. Buffin, D. Cladière, K. Hached, C. Rachez, J.M. van Deursen, K. Wassmann, Mouse oocytes depend on BubR1 for proper chromosome segregation but not for prophase I arrest, *Nat. Commun.* 21 (6) (2015) 6946.
- [21] J.P. Chambon, K. Hached, K. Wassmann, Chromosome spreads with centromere staining in mouse oocytes, *Methods Mol. Biol.* 957 (2013) 203–212.
- [22] K. Hached, S.Z. Xie, E. Buffin, D. Cladière, C. Rachez, M. Sacras, P.K. Sorger, K. Wassmann, Mps1 at kinetochores is essential for female mouse meiosis I, *Development* 138 (11) (2011) 2261–2271.
- [23] W.S. Rasband, ImageJ, US National Institutes of Health, Bethesda, MD, U.S.A. <<http://rsb.info.nih.gov/ij/>> (1997–2016) (accessed 2016).
- [24] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, J.C. Olivo-Marin, Icy: an open bioimage informatics platform for extended reproducible research, *Nat. Methods* 9 (7) (2012) 690–696. 28.
- [25] F. Luisier, C. Vonesch, T. Blu, M. Unser, Fast Interscale Wavelet Denoising of Poisson-corrupted Images, *Sig. Process.* 90 (2) (2010) 415–427.
- [26] F. Luisier, The SURE-LET Approach to Image Denoising, Swiss Federal Institute of Technology Lausanne, EPFL Thesis no. 4566 (2010), 232.
- [27] F. Luisier, C. Vonesch, T. Blu, M. Unser, Fast Haar-Wavelet Denoising of Multidimensional Fluorescence Microscopy Data, in: *Proceedings of the Sixth IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'09)*, Boston MA, USA, June 28–July 1, 2009, pp. 310–313.
- [28] G. Cox, C.J. Sheppard, Practical limits of resolution in confocal and non-linear microscopy, *Microsc. Res. Technol.* 63 (1) (2004) 18–22. 1.
- [29] M. Schrader, S.W. Hell, H.T.M. Van der Voort, Potential of confocal microscopes to resolve in the 50–100 nm range, *Appl. Phys. Lett.* 69 (24) (1996) 3644–3646.
- [30] J. Boutet de Monvel, S. Le Calvez, M. Ulfendahl, Image restoration for confocal microscopy: improving the limits of deconvolution, with application to the visualization of the mammalian hearing organ, *Biophys. J.* 80 (5) (2001) 2455–2470.



## TED: A Tolerant Edit Distance for segmentation evaluation



Jan Funke<sup>a,b,c,\*</sup>, Jonas Klein<sup>c</sup>, Francesc Moreno-Noguer<sup>a</sup>, Albert Cardona<sup>b</sup>, Matthew Cook<sup>c</sup>

<sup>a</sup> Institut de Robòtica i Informàtica Industrial, UPC/CSIC, Barcelona, Spain

<sup>b</sup> Janelia Research Campus, VA, Ashburn, United States

<sup>c</sup> Institute of Neuroinformatics, UZH/ETH, Zurich, Switzerland

### ARTICLE INFO

#### Article history:

Received 25 July 2016

Received in revised form 29 December 2016

Accepted 30 December 2016

Available online 17 January 2017

#### Keywords:

Computer vision

Segmentation

Evaluation

Learning

Neuron segmentation

Electron microscopy

### ABSTRACT

In this paper, we present a novel error measure to compare a computer-generated segmentation of images or volumes against ground truth. This measure, which we call Tolerant Edit Distance (TED), is motivated by two observations that we usually encounter in biomedical image processing: (1) Some errors, like small boundary shifts, are tolerable in practice. Which errors are tolerable is application dependent and should be explicitly expressible in the measure. (2) Non-tolerable errors have to be corrected manually. The effort needed to do so should be reflected by the error measure. Our measure is the minimal weighted sum of split and merge operations to apply to one segmentation such that it resembles another segmentation within specified tolerance bounds. This is in contrast to other commonly used measures like Rand index or variation of information, which integrate small, but tolerable, differences. Additionally, the TED provides intuitive numbers and allows the localization and classification of errors in images or volumes. We demonstrate the applicability of the TED on 3D segmentations of neurons in electron microscopy images where topological correctness is arguable more important than exact boundary locations. Furthermore, we show that the TED is not just limited to evaluation tasks. We use it as the loss function in a max-margin learning framework to find parameters of an automatic neuron segmentation algorithm. We show that training to minimize the TED, i.e., to minimize crucial errors, leads to higher segmentation accuracy compared to other learning methods.

© 2016 Published by Elsevier Inc.

## 1. Introduction

In the computer vision literature, several approaches to assess the quality of contour detection and segmentation algorithms can be found. Most of these measures have been designed to capture the intuition of what humans consider to be two similar results. In particular, these measures are supposed to be robust to certain tolerated deviations, like small shifts of contours. For the contour detection in the Berkeley segmentation dataset [1], for example, the precision and recall of detected boundary pixels within a threshold distance to the ground truth became the widely used standard [2,3]. Contour error measures are, however, not a good fit for segmentations, since small errors in the detection of a contour can lead to the split or merge of segments. Therefore, alternatives like the Variation of Information (VOI), the Rand Index [4] (RI),

the probabilistic Rand index [5,6], and the segmentation covering measure [3], have been proposed.

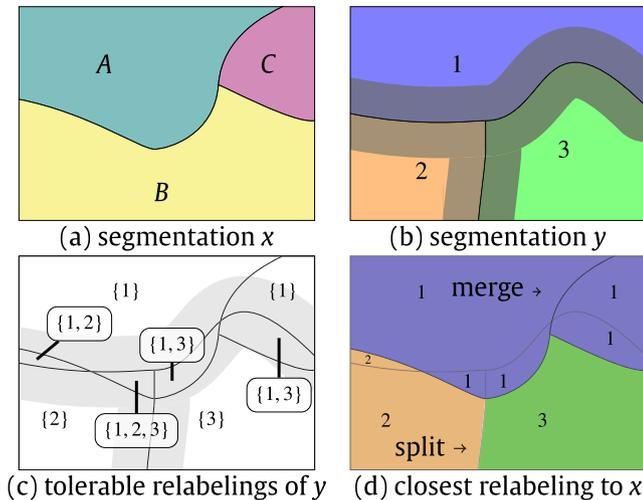
However, these measures do not acknowledge that there are different criteria for segmentation comparison, and instead accumulate errors uniformly, even for many small differences that are irrelevant in practice. Especially in the field of biomedical image processing, we are often more interested in counting true topological errors like splits and merges of objects, instead of counting small deviations from the ground truth contours. This is in particular the case for imaging methods for which no unique “ground truth” labeling exists. In the imaging of neural tissue with Electron Microscopy (EM), for example, the preparation protocol can alter the volume of neural processes, such that it is hard to know where the true boundary was [7]. Further, the imaging resolution and data quality might just not be sufficient to clearly locate contours between objects [8], resulting in a high inter-observer variability.

### 1.1. Contributions

The main contribution of this paper is a novel measure to evaluate segmentations on a clearly specified tolerance criterion to address the aforementioned issues. At the core of our measure,

\* Corresponding author at: Institut de Robòtica i Informàtica Industrial, UPC/CSIC, Barcelona, Spain.

E-mail addresses: [jfunke@iri.upc.edu](mailto:jfunke@iri.upc.edu) (J. Funke), [klein@ini.uzh.ch](mailto:klein@ini.uzh.ch) (J. Klein), [fmoreno@iri.upc.edu](mailto:fmoreno@iri.upc.edu) (F. Moreno-Noguer), [cardonaa@janelia.hhmi.org](mailto:cardonaa@janelia.hhmi.org) (A. Cardona), [cook@ini.uzh.ch](mailto:cook@ini.uzh.ch) (M. Cook).



**Fig. 1.** Illustration of the Tolerant Edit Distance (TED) between two segmentations  $x$  and  $y$ . By tolerating boundary shifts to a certain extend, shown as shadow in (b),  $y$  is allowed to be changed to match  $x$  as closely as possible. For that, we consider regions obtained by combining  $x$  and  $y$ , illustrated in (c). For each of these regions, we enumerate a set of labels used by  $y$  that are within a threshold distance to all locations inside the region (shown in curly brackets). This threshold is the maximally allowed boundary shift. Note that in this example, the region obtained from intersecting  $A$  and  $3$  can change its label to  $1$  (or keep  $3$ ), but not to  $2$ , since it contains points that are too far away from region  $2$ . Regions with only one possible label are too large to be relabeled by shifting their boundary and have to keep their initial label. From all the possible ways to relabel  $y$ , the relabeling (d) minimizing the number of split and merge errors compared to  $x$  is chosen by solving an integer linear program.

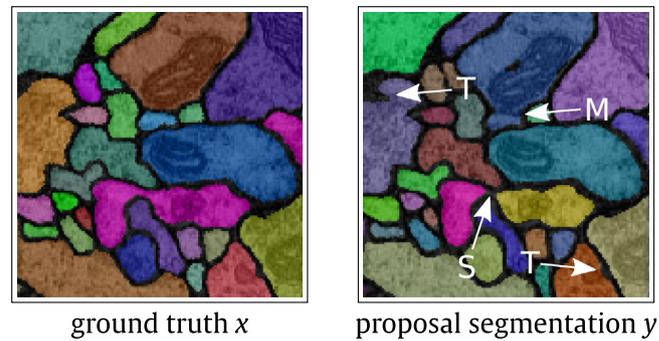
which we call *Tolerant Edit Distance* (TED),<sup>1</sup> is an explicit tolerance criterion (e.g., boundary shifts within a certain range). Using integer linear programming, we find the minimal weighted sum of split and merge operations to transform one segmentation into another, which is tolerably close to the ground truth. By setting the weights of the split and merge operations to the expected effort to perform these operations, the TED reflects the total effort needed to manually fix a segmentation. Similar to VOI and RI, our measure does not require voxels of the same object to form a connected component, and can thus be applied to volumes with missing data, known object connections via paths outside the volume, or on stitched volumes with registration artifacts. The reported numbers are intuitive (e.g., time or cost effort to fix a segmentation), easy to interpret (splits and merges of objects), and errors can be localized in the volume. An illustration of the TED can be found in Fig. 1.

## 1.2. Application to neuron segmentation

To demonstrate the usefulness of our measure, we present our results in the context of automatic neuron segmentation from EM volumes, an active field of biomedical image processing (for recent advances, see [9–13]). We argue that especially in this field there is a need for explicit and intuitive error measures. Furthermore, we show how the TED can be used to train neuron segmentation algorithms. Our findings (based on our previous work [14]) show that training to minimize the TED leads to higher segmentation accuracy on a range of error measures, compared to other methods.

### 1.2.1. Evaluation

As it is the case in many biological applications, the criterion to assess the quality of a neuron segmentation depends on the biological question one would like to answer. On one hand, *skeletons* of



**Fig. 2.** Example errors made by an automatic neuron segmentation algorithm. Errors like merges (M) and splits (S) dramatically change the reconstructed topology and should be avoided. Small disagreements in the boundary location (T) are however tolerable and should be ignored during evaluation.

neurons are sufficient to identify individual neurons [15], to study neuron types and their function [16], and to obtain the wiring diagram of a nervous system (the so-called *connectome*) [8]. In these cases, topological correctness is far more important than the diameter of a neural process or the exact location of its boundary (see Fig. 2 for examples). On the other hand, for biophysically realistic neuron simulation, *volumetric* information is needed to model action potential time dynamics, and to understand and simulate information processing capabilities of single neurons [17]. In this case, the segmentation should be close to the true volume of the reconstructed neurons. Only small deviations in the boundary location might still be tolerable.

Currently, reporting segmentation accuracy in terms of VOI or RI is the de facto standard [11,18,10,12,13]. Less frequently used [9,19] is the *Anisotropic Edit Distance* (AED) [9] and the *Warping Error* (WE) [20]. The AED is tailored to the specific error correction steps required for anisotropic volumes (splits and merges of 2D neuron slices within a section, connections and disconnections of slices between sections). The WE aims to measure the difference between ground truth and a proposal segmentation in terms of their topological differences. As such, the WE was the first error measure for neuron segmentation that deals with the delicate question of up to which point a boundary shift is not considered to be an error. However, since the WE assumes a foreground-background segmentation where connected foreground objects represent neurons, it is only applicable to volumes in which connectedness of neurons is preserved. Furthermore, only suboptimal solutions to the WE are found using a greedy, randomized heuristic, which makes it difficult to use for evaluation purposes. Consequently, the WE has found its main application in the training of neural networks for image classification [20].

In 2 we introduce the TED as an alternative to address some of the shortcomings of existing measures. Similar to the WE, the TED is designed to ignore small deviations from the ground truth and only count true topological errors, but is computed deterministically and to global optimality and does not impose constraints on the types of volumes being compared.

### 1.2.2. Training

Current state-of-the-art methods for automatic neuron segmentation can broadly be divided into isotropic [11,18,12,13] and anisotropic methods [9,10,19]. Assignment models constitute the current state of the art for the segmentation of neurons from anisotropic volumes, as obtained by serial section EM [9,10]. These models enumerate and price possible assignments of candidate segments across sections of EM stacks (see Fig. 8 for an overview and 3 for details). A final segmentation is found by selecting a cost minimal and consistent subset of all assignments.

<sup>1</sup> Source code available at <http://github.com/funkey/ted>.

Learning in this kind of models consists of finding a function that maps from features of the candidate segments to a cost. Currently, this function is either set by hand [21,10,22], learned from a random forest classifier based on positive and negative assignment examples [9,23], or found via grid-search by tuning weights of a small number of features [24]. Except for grid-search, which does not scale to larger sets of parameters, none of the currently used training methods implements real end-to-end learning. In 3, we show how to overcome these limitations by performing structured learning on a sensible loss function. For that, we solve two sub-problems: (1) We show how to generate a *training sample* suitable for structured learning from human annotated ground truth. (2) We introduce a loss for structured learning, which minimizes the TED during learning.

We show that our learning framework leads to consistently higher segmentation accuracy compared to other learning methods. Furthermore, we show that our learning framework can be used to train on skeleton annotations without big sacrifices in segmentation accuracy. Skeleton annotations are non-volumetric centerlines of neurons, which are in practice much faster to obtain.

## 2. Tolerant Edit Distance

In this section, we formally introduce the TED and its associated optimization problem. We will show how to compute the TED for a specific class of tolerance criteria, of which the boundary shift is an example. Finally, we will analyze some of the properties of the TED in the context of neuron segmentation and contrast them with conventional error measures used in this field.

### 2.1. Definition of the TED

The TED measures the distance<sup>2</sup> between two segmentations  $x: \Omega \mapsto K_x$  and  $y: \Omega \mapsto K_y$ , where  $\Omega$  is a discrete set of voxel (or supervoxel) locations in a volume, and  $K_x$  and  $K_y$  are sets of labels used by  $x$  and  $y$ , respectively. The distance is reported in terms of the minimal number of splits and merges appearing in a relabeling of  $y$ , as compared with  $x$ . How  $y$  is allowed to be relabeled is defined on a tolerance criterion, e.g., the maximal displacement of an object boundary.

We say that a label  $k \in K_x$  overlaps with a label  $l \in K_y$ , if there exists at least one location  $i \in \Omega$  such that  $x(i) = k$  and  $y(i) = l$ . If  $x$  and  $y$  represent the same segmentation, each label  $l$  overlaps with exactly one label  $k$ , and vice versa. Consequently, if a label  $k \in K_x$  overlaps with  $n$  labels from  $K_y$ , we count it as  $n - 1$  splits. Analogously, if a label  $l \in K_y$  overlaps with  $n$  labels from  $K_x$ , we count it as  $n - 1$  merges. For two labelings  $x$  and  $y$ , we denote as  $s(x, y)$  and  $m(x, y)$  the sum of splits and merges over all labels.

At the core of the TED lies a to-be-defined tolerance criterion  $T$ , which is meant to formalize our intuition about how a segmentation  $y$  is allowed to be relabeled. More formally,  $T(y, y')$  is supposed to evaluate to  $\top$  if  $y'$  is a tolerated relabeling, and to  $\perp$  otherwise. With  $\mathcal{Y}$  being the set of all labeling functions  $y': \Omega \mapsto K_y$ , (i.e., all possible labelings of  $\Omega$  using the labels of  $y$ ), we call the subset  $\mathcal{Y}^+(y) = \{y' \in \mathcal{Y} \mid T(y, y') = \top\}$  the set of all tolerated relabelings of  $y$ . The TED is the minimal weighted sum of splits and merges over all tolerable relabelings  $\mathcal{Y}^+(y)$ :

$$\text{TED}(x, y) = \min_{y' \in \mathcal{Y}^+(y)} \alpha s(x, y') + \beta m(x, y'), \quad (1)$$

where the weights  $\alpha$  and  $\beta$  represent the time or effort needed to fix a split or merge, respectively.

Without imposing restrictions on the tolerance criteria  $T$ , the optimization in Eq. (1) is intractable in general. Therefore, we restrict ourselves to what we call *local tolerance criteria* in the following. A local tolerance criterion is completely defined by providing relabel alternatives  $A_i \subseteq K_y$  for each location  $i$ , such that each relabeling  $y'$  using any label  $y'(i) \in A_i$  is tolerated. More formally,

$$T_{\text{local}}(y, y') = \bigwedge_{i \in \Omega} y'(i) \in A_i. \quad (2)$$

One example of such a local tolerance criterion is the boundary shift up to a distance threshold  $\theta$ , which we illustrate in Fig. 1(c). For this tolerance criterion,  $A_i$  of a location  $i$  comprises the union of labels of all other locations that are within a  $\theta$  distance from  $i$ .

For local tolerance criteria, Eq. (1) can be solved with the following integer linear program (ILP):

$$\min_{\mathbf{v}} \quad \alpha s + \beta m \quad (3)$$

$$\text{s.t.} \quad \sum_{l \in A_i} v_{i-l} = 1 \quad \forall i \in \Omega \quad (4)$$

$$\sum_{i \in \Omega} v_{i-l} \geq 1 \quad \forall l \in K_y \quad (5)$$

$$a_{kl} - v_{i-l} \geq 0 \quad \forall i \in \Omega : x(i) = k \quad (6)$$

$$a_{kl} - \sum_{i \in \Omega : x(i)=k} v_{i-l} \leq 0 \quad \forall k \in K_x \quad \forall l \in K_y \quad (7)$$

$$s_k - \sum_{l \in K_y} a_{kl} = -1 \quad \forall k \in K_x \quad (8)$$

$$m_l - \sum_{k \in K_x} a_{kl} = -1 \quad \forall l \in K_y \quad (9)$$

$$s - \sum_{k \in K_x} s_k = 0 \quad (10)$$

$$m - \sum_{l \in K_y} m_l = 0 \quad (11)$$

At the core of this ILP are binary indicator variables  $\mathbf{v} = (v_{i-l} \in \{0, 1\} \mid i \in \Omega, l \in A_i)$  to indicate the assignment of label  $l$  to location  $i$ . Constraints Eq. (4) and Eq. (5) ensure that exactly one of the labels gets chosen for each location and that each label of  $y$  has to appear at least once. Further, we introduce binary variables  $a_{kl}$  that indicate the presence of a joint assignment of label  $k$  from  $x$  and label  $l$  from  $y'$  at at least one location. With constraints Eqs. (6 and 7) we make sure that each  $a_{kl} = 1$  if and only if there is at least one location  $i \in \Omega$  such that  $x(i) = k$  and  $y'(i) = l$ . To count the number of times a label  $k \in K_x$  is split in  $y'$ , we further introduce integers  $s_k \in \mathbb{N}$ . These counts equal the number of times  $k$  was matched with any other label minus one, which we ensure with constraints Eq. (8). Analogously, we introduce integers  $m_l$  and constraints (9) for merges caused by label  $l$  in  $y'$ . The final split and merge numbers  $s$  and  $m$  are just the sums of the label-wise splits and merges, ensured by Eq. (10 and 11).

Once the optimal solution of this ILP has been found, the variables  $a_{kl}$  can be used to determine which labels got split and merged, and thus to localize errors.

<sup>2</sup> Note that, due to the intended tolerance to small deviations, the TED is not a proper metric on the space of segmentations. In slight abuse of nomenclature we use the term distance here anyway, which is sometimes used synonymously for metric.

## 2.2. Discussion of the TED

### 2.2.1. Parameters

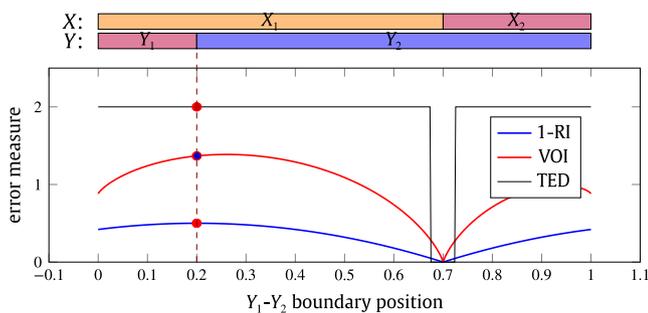
As formulated above, the TED and the boundary shift tolerance criterion introduce three parameters:  $\alpha$  and  $\beta$  to score differently split and merge errors, and a threshold  $\theta$  for the maximally permitted boundary shift.

$\alpha$  and  $\beta$  can be set straightforwardly as the effort or time needed to fix a split or merge error. This depends on the concrete application and the tools available to proof-readers. Since a study of the time needed to fix segmentations is beyond the scope of this paper, we will proceed as follows: We set  $\alpha = \beta = 1$  for this discussion, so as to count the number of errors. For the experiments presented in (3), we will set  $\alpha = 1$  and  $\beta = 2$  to reflect that merges are usually more difficult to fix than splits. Note that, up to scale, the TED will be the same for equally scaled  $\alpha$  and  $\beta$ . We allow them to be set independently anyway to obtain directly a time-to-fix estimate if  $\alpha$  and  $\beta$  reflect time.

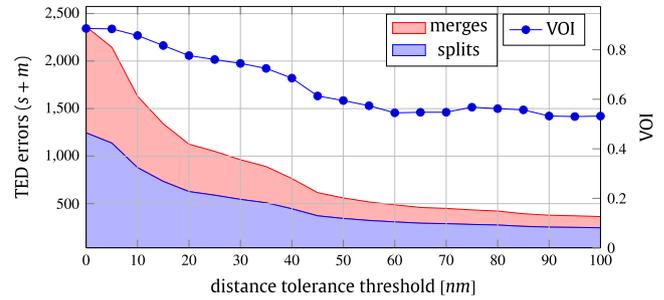
The distance threshold  $\theta$  might not be as obvious to set. Setting this value requires us to find an answer to the unpleasant question until which point a deviation from the ground truth is just a tolerable dent in a segment or a real error that should be counted. A single threshold alone is unlikely to provide an answer to this question. But, following a popular philosophy, we think that explicit is better than implicit. By explicitly setting this value, we achieve two things: First, we know exactly how to interpret the values measured by the TED. Second, we confront ourselves with the aforementioned unpleasant question, which we hope will encourage us to come up with more elaborate tolerance criteria, tailored to the needs of specific applications.

### 2.2.2. Shift of object boundary

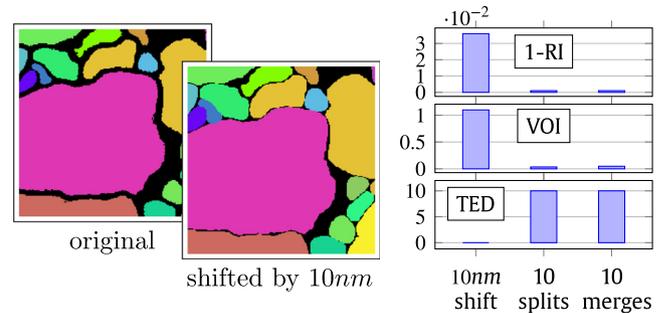
To illustrate the behavior of different error measures in the case of object boundary displacements, we created a simple artificial 1D labeling consisting of two regions. In Fig. 3, we show the errors of segmentations obtained by shifting the boundary between the objects. It can clearly be seen that TED assigns the same numbers (one split and one merge error) as soon as a given tolerance criterion is exceeded (0.025 in this example), regardless where the error happens. This is the desired outcome for applications like neuron segmentation, where it is important to count the number of topological errors regardless of how many voxels got affected.



**Fig. 3.** Comparison of Rand index (RI), variation of information (VOI), and Tolerant Edit Distance (TED) as functions of object boundary displacements. Given a ground truth labeling  $X$ , the error measures are plotted as functions of the split position between two objects in a reconstruction  $Y$ . It can clearly be seen that TED assigns the same numbers (one split and one merge error) as soon as a given tolerance criterion is exceeded (0.025 in this example). VOI is in bits (lower is better) and 1-RI is 1 minus the ratio of agreeing pairs over all pairs (lower is better). The advantage of the TED is that it explicitly counts the topological errors made, regardless where in the segment they occur. Furthermore, small boundary shifts are not counted at all, whereas for RI and VOI their contribution can not be distinguished from real errors.



**Fig. 4.** The Tolerant Edit Distance (TED) on an automatically generated reconstruction as a function of the tolerated boundary shift.



**Fig. 5.** Comparison of error measures between the original ground truth (left) and three modifications. For the boundary shift experiment, the labels of the ground truth were dilated by 10 nm. For the *split* and *merge* experiments, ten random locations were chosen where the ground truth neurons were manually split or merged, respectively. Both RI and VOI assign better scores (i.e., higher for RI, lower for VOI) to the *split* and *merge* experiments than to the *shift* experiment. The TED boundary shift tolerance was set to 20 nm and thus counts only the true morphological errors as false splits (FS) and false merges (FM).

### 2.2.3. Influence of distance threshold

In order to study the effect of the threshold distance for boundary shifts, we used an automatic segmentation result<sup>3</sup> and evaluated the TED for varying thresholds. Results are shown in Fig. 4. The TED reveals that most of the errors occur within the range of about 50 nm, corresponding to about 12 pixels in the x-y-plane of this dataset. Depending on the biological question, those errors might be tolerable. In the same plot, we show the VOI of the closest tolerable relabeling to the ground truth under the given boundary shift threshold (i.e., the equivalent of Fig. 1 (d) on the proposal segmentation). From this example, we can see that the errors <50 nm contribute quite significantly with 0.23 bits to the total VOI of 0.886, and thus can shadow true topological errors.

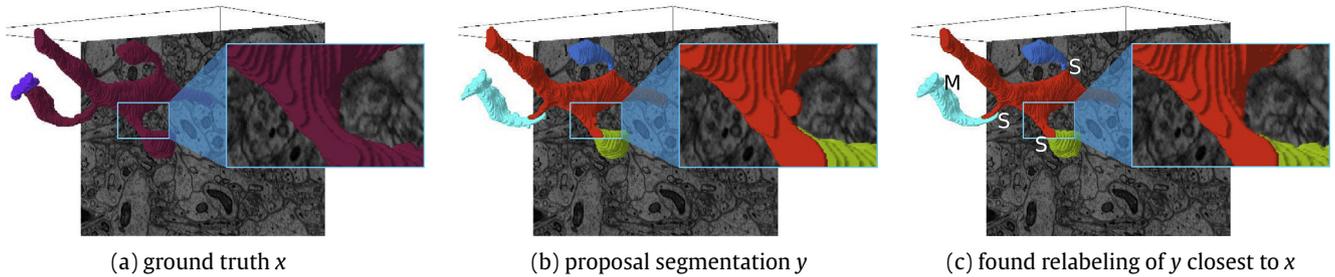
### 2.2.4. Comparison to RI and VOI

To demonstrate the main differences between TED and conventional error measures, we compare RI and VOI against TED for three manual modifications of the ground truth labeling of [25], shown in Fig. 5. For the *10 nm shift* experiment, we dilated the boundaries of neurons in the ground truth by 10 nm. For the *splits* and *merges* experiment, we split and merged neurons at 10 randomly selected locations, respectively. It can be seen that the small shifts of object boundaries can have a significant contribution to the measures RI and VOI, which confirms our previous observation.

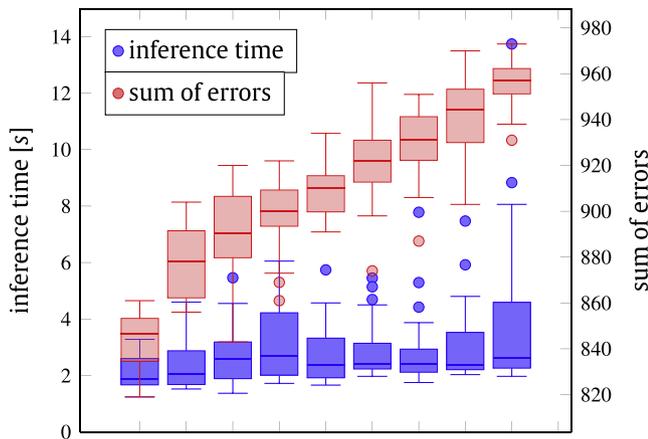
### 2.2.5. Localization of errors

Due to the explicit tolerance criterion of the TED, errors can be localized in the volume. In Fig. 6 we show example split and merge

<sup>3</sup> Obtained using SOPNET [9] on a publicly available EM dataset [25].



**Fig. 6.** Errors found by the TED between a human generated ground truth  $x$  (a) and a proposal segmentation  $y$  (b), illustrated on two neurons (purple and red in ground truth). Small errors, as the one shown in the magnification, are tolerated and consequently removed in the found relabeling of  $y$  (c). Remaining errors are considered real splits (S) and merges (M).



**Fig. 7.** Runtime analysis of the TED computation on several randomized segmentations of a  $1000 \times 1000 \times 100$  volume.

errors detected by the TED on an automatic segmentation result for the SNEMI dataset [26]. The boundary shift tolerance was set to 100 nm, which corresponds to  $16.6 \times 16.6 \times 3.3$  voxels for this volume with a resolution of  $6 \text{ nm} \times 6 \text{ nm} \times 30 \text{ nm}$ . With this setup, the TED reveals true topological errors made by the automatic segmentation method. This allows analyzing the weaknesses of a method, which is both useful for model design as well as to communicate the limits of what can be done with a method to neuroscientists.

### 2.2.6. Runtime

The runtime of the TED depends both on the size of the volumes and their discrepancy. The less similar two segmentations are, the more variables have to be introduced to represent the possible relabelings. This results in larger ILPs that are in general harder to solve.

We studied the impact of discrepancy on the runtime of the ILP by producing randomly generated segmentations. For that, we first created a reference segmentation by iteratively agglomerating supervoxels of a  $1000 \times 1000 \times 100$  volume, using an affinity-based scoring function to propose the next merge.<sup>4</sup> We stopped the agglomeration at a manually set threshold to produce  $\sim 800$  components. For the randomized segmentations, we added random noise of increasing intensity to the scoring function of the agglomeration to generate more and more discrepancies compared to the reference. Each segmentation obtained this way was compared against the reference segmentation. We measured the single-thread runtime on a Intel(R) Xeon(R) CPU with 2.2 GHz, using Gurobi to solve the ILP. Results for 9 noise intensities (with 20 repetitions each) are shown in Fig. 7. It can be seen that, although the number of errors

goes up as high as 960, the vast majority of runs finished in less than 4 s. The number of variables in the ILPs ranged from 59,397 (most similar segmentations) to 69,105 (most dissimilar segmentations).

These results match our observations so far and can be summarized in the following way: If two segmentations are similar enough, the runtime of the TED seems to be moderate and an exact solution of the ILP is tractable in practice. Although we have so far not encountered intractable instances, we can not exclude their existence. We hypothesize that in such a case the segmentations in question would be very dissimilar and an approximate solution to the ILP would suffice.

## 3. Learning of assignment models

In this section, we demonstrate that the TED can be used as a loss function to train neuron segmentation methods. Here, we focus on *assignment models* which gained popularity for the segmentation of anisotropic volumes of neural tissue. For that, we show first how the structured learning framework can be used to learn weights of a generic cost function. Second, we develop a tractable approximation of the TED that can be used as loss for structured learning. We report results on two publicly available datasets.

### 3.1. Assignment models

Assignment models for anisotropic neuron segmentation introduce  $n$  binary indicator variables  $\mathbf{z} \in \{0, 1\}^n$  to represent possible assignments of 2D neuron candidates across consecutive pairs of sections of a volume (for an illustration see Fig. 8, more details about assignment models can be found in [9,10]). Linear constraints are formulated on the binary assignment indicators to ensure that a solution is consistent. In particular, the following set of constraints ensures that no pair of overlapping candidates are chosen (see also Fig. 8(g)):

$$\sum_{c \in \mathcal{C}} \sum_{z_i \in \mathbf{z}_{-c}} z_i \leq 1 \quad \forall c \in \mathcal{C} \quad (12)$$

Here,  $\mathcal{C}$  denotes the set of all conflict cliques, i.e., sets of candidates that are mutually overlapping and  $\mathbf{z}_{-c}$  all assignment variables that link  $c$  to the previous section. For each conflict clique  $C$ , we require the number of assignment variables linking any candidate in it to the previous section to be at most 1. These constraints are accompanied by the following, which ensure a contiguous sequence of assignments (see also Fig. 8(h)):

$$\sum_{z_i \in \mathbf{z}_{-c}} z_i - \sum_{z_i \in \mathbf{z}_{c-}} z_i = 0 \quad \forall c. \quad (13)$$

Here,  $\mathbf{z}_{c-}$  denotes all assignments variables that link a candidate  $c$  to the next section. Noting that the above constraints are linear in  $\mathbf{z}$ , we can characterize the set of consistent solutions as

<sup>4</sup> We used the implementation <http://github.com/funkey/waterz> on a volume of neural tissue, for which we predicted voxel-wise affinities.

$$\mathcal{Z} = \{\mathbf{z} \in \{0, 1\}^n \mid \mathbf{A}\mathbf{z} \preceq \mathbf{b}\} \quad (14)$$

where we write  $\mathbf{a} \preceq \mathbf{b}$  to say that  $\mathbf{a}$  is element-wise less than or equal to  $\mathbf{b}$ . Given a cost vector  $\mathbf{c}$  for the assignment variables, the optimal assignment vector is the solution to the integer linear program

$$\min_{\mathbf{z} \in \mathcal{Z}} \langle \mathbf{c}, \mathbf{z} \rangle. \quad (15)$$

Without loss of generality, we assume that the cost  $c_i$  for selecting an assignment  $z_i$  is a weighted sum of features  $\phi_i$  extracted for this assignment:

$$\mathbf{c} = \Phi \cdot \mathbf{w} = [\phi_1, \phi_2, \dots, \phi_n]^\top \cdot \mathbf{w}. \quad (16)$$

### 3.2. Learning of model parameters

Using the structured learning framework [27], we find the optimal  $\mathbf{w}$  given annotated training data  $(\phi, \mathbf{z}')$ . More specifically, we use the margin rescaling variant to find the weights  $\mathbf{w}^*$  as the minimizer of

$$L(\mathbf{w}) = \lambda \|\mathbf{w}\|^2 + \max_{\mathbf{z} \in \mathcal{Z}} [\langle \Phi \mathbf{w}, \mathbf{z}' \rangle - \langle \Phi \mathbf{w}, \mathbf{z} \rangle] + \Delta(\mathbf{z}', \mathbf{z}), \quad (17)$$

where  $\lambda$  is the regularizer weight and  $\Delta(\mathbf{z}', \mathbf{z})$  is an application specific loss function. In order for this method to be successful, two problems need to be solved: (1) a representative training sample  $\mathbf{z}'$  has to be found, and (2) a sensible loss function  $\Delta(\mathbf{z}', \mathbf{z})$  has to be designed.

### 3.3. Training sample $\mathbf{z}'$

Even apart from the difficulties in obtaining unambiguous human generated ground truth for the neuron segmentation problem in the first place, the provision of  $\mathbf{z}'$  is not trivial: We have to find a member of  $\mathcal{Z}$ , i.e., the set of all possible assignment vectors using the found 2D neuron candidates, that is as close as possible to the human annotated ground truth. We have to note that the extracted 2D neuron candidates can be imperfect and thus there might not be a  $\mathbf{z} \in \mathcal{Z}$  that corresponds to the human annotated ground truth. Consequently, we have to accept that the training sample  $\mathbf{z}'$  will only represent a best-effort solution and not the ground truth.

In order to find this best-effort solution in a principled way, we assign a local ground truth matching score  $g_i$  to each assignment and then select a consistent solution that minimizes this score. Let  $\Omega = [1, W] \times [1, H] \times [1, D]$  be the set of all discrete pixel locations in a stack of size  $W \times H \times D$ . We assume a ground truth labeling  $x: \Omega \mapsto K$  that assigns a unique label  $k \in K$  to each ground truth segment in the volume. Let  $u(i)$  and  $v(i)$  denote the section indices that are linked by assignment  $z_i$ . We denote by  $A_i \subset \Omega$  the set of pixels of section  $u(i)$  and  $v(i)$  that are merged by the assignment  $z_i$ . Similarly, let  $G_i^k \subset \Omega$  denote the set of pixels that are labeled to belong to the same region  $k$  in the ground truth, limited to the sections  $u(i)$  and  $v(i)$ . For each pair of assignment  $i$  and ground truth label  $k$ , we compute a similarity  $g_i^k$  that rewards overlap between the sets  $A_i$  and  $G_i^k$  and punishes set differences:

$$g_i^k = \underbrace{|G_i^k \cap A_i|}_{\text{overlap}} - \underbrace{(|G_i^k \setminus A_i| + |A_i \setminus G_i^k|)}_{\text{set difference}}. \quad (18)$$

The final matching score  $g_i$  of an assignment  $z_i$  is the maximal similarity with any ground truth label:

$$g_i = \max_{k \in K} g_i^k. \quad (19)$$

The scores  $g_i$  reflect, for each assignment  $z_i$ , how well it locally fits to the ground-truth. We use these scores to find the overall best assignment  $\mathbf{z}$  by solving the following ILP:

$$\mathbf{z}' = \arg \max_{\mathbf{z} \in \mathcal{Z}} \langle \mathbf{g}, \mathbf{z} \rangle. \quad (20)$$

Note that this ILP is maximizing the sum of similarities for all assignments. This way we find a consistent solution (in terms of the constraints introduced in 3.1) that maximizes similarity with the provided ground truth.

### 3.4. Loss $\Delta(\mathbf{z}', \mathbf{z})$

Ideally, we would use the error measure that we use to evaluate the results of our automatic segmentation as  $\Delta(\mathbf{z}', \mathbf{z})$ . However, we have to make sure that the maximization in Eq. (17) is still tractable.

To this end, we suggest a first order approximation of the TED to be used as  $\Delta(\mathbf{z}', \mathbf{z})$ : For each assignment variable  $z_i$ , we estimate its contribution  $l_i$  to the TED score. If  $z_i = z'_i$ , no error was introduced by  $z_i$  and hence its contribution is 0. If, however,  $z_i \neq z'_i$ , the resulting segmentation will deviate from the best-effort solution. In order to estimate the contribution of an erroneous  $z_i$  to the TED score, we compute the TED score between two segmentations  $y_{\mathbf{z}'}$  and  $y_{\mathbf{z}(i)}$ :  $y_{\mathbf{z}'}$  denotes the segmentation obtained from the best-effort solution  $\mathbf{z}'$  and  $y_{\mathbf{z}(i)}$  denotes the segmentation obtained by  $\mathbf{z}$ , but with  $z'_i$  inverted.<sup>5</sup> More formally, we set

$$l_i = (1 - 2z'_i) \text{TED}(y_{\mathbf{z}'}, y_{\mathbf{z}(i)}) \quad \text{and} \quad c = \sum_{i: z'_i=1} -l_i, \quad (21)$$

where some of the contributions  $l_i$  turn into rewards (negative values) for using an assignment, i.e., when the corresponding  $z'_i = 1$ . This linearization allows us to model the loss as a linear function of  $\mathbf{z}$ :

$$\Delta(\mathbf{z}', \mathbf{z}) = \langle \mathbf{l}, \mathbf{z} \rangle + c \approx \text{TED}(y_{\mathbf{z}'}, y_{\mathbf{z}}), \quad (22)$$

which favorably plugs into Eq. (17). In fact, the loss augmented inference problem for a given  $\mathbf{w}$  has the same structure as the inference problem Eq. (15) itself, for which we already know that it is tractable in practice:

$$\max_{\mathbf{z} \in \mathcal{Z}} \langle \mathbf{l} - \Phi \mathbf{w}, \mathbf{z} \rangle + \underbrace{\langle \Phi \mathbf{w}, \mathbf{z}' \rangle + c}_{\text{constant}}. \quad (23)$$

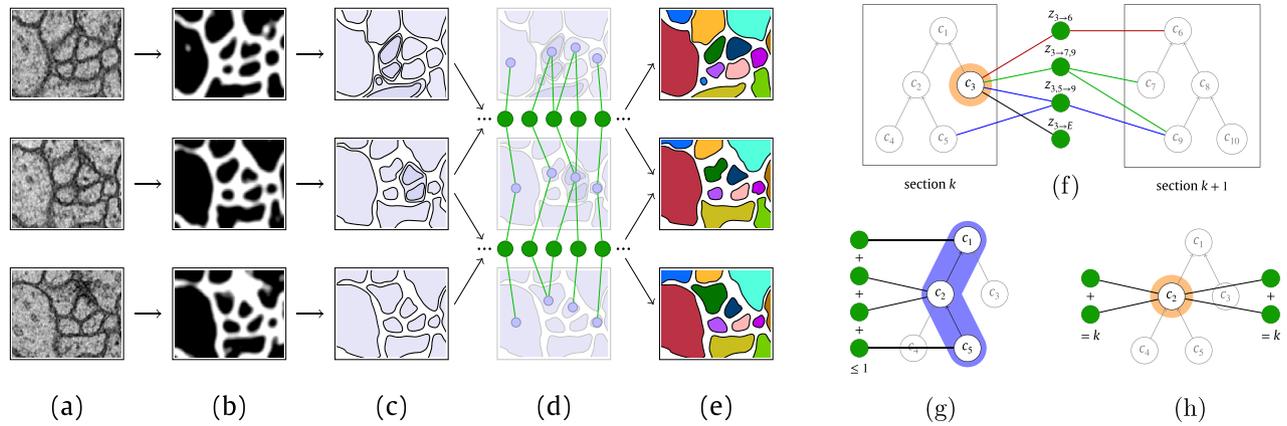
### 3.5. Results

We use two publicly available datasets for our experiments, which we refer to as DROSOPHILA [25], which consists of two stacks of 20 EM sections with  $4 \times 4 \times 40\text{nm}$  resolution ( $1024 \times 1024 \times 20$  pixels), and MOUSE CORTEX [26], which consists of two stacks of 100 EM sections with  $6 \times 6 \times 30\text{nm}$  resolution ( $1024 \times 1024 \times 100$  pixels).

We split the parts for which ground truth was available into two stacks of equal size ( $2 \times 10$  sections for DROSOPHILA and  $2 \times 50$  sections for MOUSE CORTEX). For each dataset, we trained all methods on a sample  $\mathbf{z}'$  (see 3.3) extracted from the first stack and report the results on the second stack.

We trained and evaluated the assignment model implemented in SOPNET [9], using membrane predictions from [28], and 2D neuron candidates extracted from component trees [9]. We used the default features implemented in SOPNET for  $\Phi$ .

<sup>5</sup> Since the constraints Eq. (14) might not allow inverting single variables in isolation, we identify a minimal group of variables that have to be inverted as well to obtain a consistent solution: for each assignment  $i$ , we find an assignment vector  $\mathbf{z}(i) \in \mathcal{Z}$  that has  $z_i \neq z'_i$  and minimizes the Hamming distance to  $\mathbf{z}'$ .



**Fig. 8.** Assignment model for anisotropic neuron segmentation. From a stack of EM sections (a), a pixel classifier is used to predict membrane locations (b). Several, possibly overlapping, 2D neuron candidates are extracted for each section (c), and possible assignments are enumerated between candidates of adjacent sections (d). In the model, each assignment of candidates between two sections (f) is represented by a binary variable  $z_i$  and has an associated cost  $c_i$  for selecting it. By finding a cost-minimal  $\mathbf{z}$  subject to constraints (g) and (h) the final segmentation is obtained (e).

3.5.1. Comparison of learning methods

We compare the structured learning method proposed in 3 to random forests (RF) as proposed in [9,19], support vector machines (SVM), and *overlap*. RF and SVM learn to score each assignment based on positive and negative examples provided by  $\mathbf{z}'$  (see 3.3). As a baseline, *overlap* uses the number of overlapping pixels of an assignment across sections as score. Since these methods need a prior for the selection of assignments, we trained RF and SVM on a subset of the training data (5 sections for DROSOPHILA, 40 sections for MOUSE CORTEX) and used the rest to validate a prior for RF, SVM, and *overlap* with a grid-search minimizing the Hamming distance to  $\mathbf{z}$ .

To study the performance of the structured learning method, we compare our loss SL-TED (see 3.4) against three baselines: SL-Ham, SL-VOI, and SL-RI. SL-Ham uses the Hamming distance of  $\mathbf{z}$

to  $\mathbf{z}'$  for  $\Delta(\mathbf{z}', \mathbf{z})$ . SL-VOI and SL-RI use the same linear approximation scheme we developed for the TED (see 21), but with VOI and RI as error measures instead of TED. For the computation of SL-TED, we evaluated the TED allowing boundary shifts up to  $\theta = 100$  nm, with weights  $\alpha = 1$  and  $\beta = 2$  to account for the fact that merges lose geometric information and thus usually take more time to repair than splits.

Results are shown in Table 1. We report errors for several commonly used measures for neuron segmentation: Rand Index (RI), Variation of Information (VOI), Anisotropic Edit Distance [9] (AED, note that we refer to the inter FP/FN as FS/FM), and TED. The TED counts topological errors that are not considered boundary shifts as false splits (FS) and false merges (FM). Splits of the ground truth background label are false positives (FP) and merges involving the reconstruction background label false negatives (FN).

**Table 1**  
Comparison of segmentation results of different learning methods on two anisotropic EM datasets.

Mouse Cortex Dataset														
method	Rand	split	VOI merge	total	FP	FN	AED FS	FM	total	FP	FN	TED FS	FM	TTF
overlap	0.9939	0.668	0.192	0.860	1,553	2,404	3,114	1,666	8,737	155	179	678	57	1,305
RF	0.9936	0.375	0.291	0.666	1,048	2,546	3,014	1,451	8,059	23	151	273	68	734
SVM	0.9572	0.507	1.434	1.940	2,998	3,761	5,155	4,587	16,501	4	147	129	167	761
SL-Ham	0.9933	0.348	0.309	0.657	895	2,258	2,735	1,333	7,221	23	138	243	82	706
SL-VOI	0.9870	0.525	0.899	1.424	799	2,466	2,884	1,325	7,474	14	127	161	141	711
SL-RI	0.9797	0.514	1.047	1.561	780	2,604	3,004	1,291	7,679	14	129	163	143	721
SL-TED	<b>0.9948</b>	0.331	0.275	<b>0.606</b>	838	2,297	2,752	1,268	<b>7,155</b>	18	135	229	82	<b>681</b>

Drosophila Dataset														
method	Rand	split	VOI merge	total	FP	FN	AED FS	FM	total	FP	FN	TED FS	FM	TTF
overlap	0.9906	0.309	0.340	0.648	179	517	648	254	1,598	13	58	201	99	528
RF	0.9864	0.934	0.518	1.452	181	585	556	252	1,574	1	175	108	35	529
SVM	0.9890	0.804	0.230	1.034	366	357	593	537	1,853	10	86	224	84	574
SL-Ham	0.9959	0.309	0.080	0.389	241	234	375	250	1,100	14	63	227	47	461
SL-VOI	0.9959	0.301	0.101	0.402	214	268	400	202	<b>1,084</b>	14	60	243	54	485
SL-RI	0.9958	0.301	0.109	0.410	202	288	419	192	1,101	16	59	243	60	497
SL-TED	<b>0.9960</b>	0.299	0.087	<b>0.386</b>	224	249	382	239	1,094	15	63	215	50	<b>456</b>

**Table 2**

Reconstruction results on MOUSE CORTEX after training on different ground truth types: volumetric uses the original ground truth, skeleton a skeletonized version. We show false splits and false merges (FS and FM), false positives and false negatives (FP and FN), and an estimated time-to-fix (TTF), as reported by the TED measure.

MOUSE CORTEX DATASET					
method	TED				
	FP	FN	FS	FM	TTF
volumetric ground truth	18	135	229	82	<b>681</b>
skeleton ground truth	17	114	188	152	737

For the time-to-fix (TTF) estimate, we again set the time needed for fixing a split to  $\alpha = 1$  and for fixing a merge to  $\beta = 2$ . The structured learning methods are in general superior to *overlap*, RF, and SVM, with the best results being obtained by training on SL-TED. Training on the TED-approximation SL-TED does indeed minimize the TTF. Furthermore, RI, VOI, and AED are minimized. Our results also reveal interesting differences between error measures: Although the best solutions in terms of TED have also best RI, VOI, and AED, we see a discrepancy in the mid-field: on DROSOPHILA, SVM scores much better than RF in terms of VOI and slightly better in terms of RI. However, TED on a clearly defined criterion shows that the numbers are misleading and in fact RF has less errors in total and shorter TTF.

### 3.5.2. Learning from skeletons

We show on MOUSE CORTEX that our method to find a training sample  $\mathbf{z}'$  allows us to train on skeleton annotations as well. Skeleton annotations are not volumetric, i.e., instead of labeling every pixel, only the centerline of the neuron is provided as training data. In practice, this saves a lot of manual labeling effort such that larger volumes can be annotated. To simulate skeleton annotations and compare them to the learning outcome of complete ground truth, we skeletonized each ground truth label of the training stack. For that, we shrunk each 2D connected component of one label in each EM section to a single pixel at its center of mass. Consequently, we adjusted the search for the training sample  $\mathbf{z}'$  to not consider the set difference term in Eq. (18). The results of training with SL-TED on the  $\mathbf{z}'$  obtained this way are shown in Table 2. Although significant, the loss in accuracy might be compensated by the time saved to annotate only skeletons for training.

### 3.5.3. Runtimes

The bottleneck of our method is the computation of the coefficients  $l_i$  needed for the TED approximations SL-TED, since for every binary variable in the  $\mathbf{z}'$  the TED has to be evaluated. For MOUSE CORTEX and DROSOPHILA,  $\mathbf{z}'$  contained 277,874 and 20,890 variables, respectively. Computing the coefficients took 64.3 h for MOUSE CORTEX and 4.8 h for DROSOPHILA on a 12 core Intel Xeon CPU with 3.47 GHz. By noting that the influence of a single variable flip is usually local, the computation of the TED could be limited to constant size subvolumes around the variable of interest, such that the effort of computing the coefficients scales linearly with the best-effort size. Structured learning with SL-TED took 30 m for DROSOPHILA and 1 h 45 m for MOUSE CORTEX on 10 cores of a Intel Xeon CPU with 2.6 GHz. We used an iterative cutting plane method<sup>6</sup> to minimize the convex learning

objective Eq. (17) to optimality. The maximization in 17 has been solved with an ILP to optimality (using the Gurobi solver) in each iteration as well.

## 4. Conclusions

We presented the TED, a novel measure for segmentation comparison, which tolerates small errors based on an explicit tolerance criterion and therefore focusses on counting true topological errors. As such, it is suited to report an effort or time to fix estimate.

A current limitation of the TED is the restriction to use local tolerance functions, e.g., a boundary shift up to a certain threshold. More complex tolerance criteria that do not factorize over regions are currently not expressible. Although they could in theory be incorporated into the ILP by adding auxiliary variables, it remains questionable whether the resulting problem is still tractable. Even though we did not observe that empirically, it is already conceivable in the current formulation that an optimal solution to the ILP can not be found in reasonable time. This could in particular be the case if ground truth and proposal segmentation differ a lot and a very lax tolerance criterion is used. In these cases, approximate solutions to the proposed ILP might be worth considering.

Besides being a tool to assess the quality of a segmentation, we also showed that the TED can be used to train a neuron segmentation algorithm.

We believe that the key for the superior performance of training using the TED compared to other losses is the consideration of topological errors. Previous attempts tried to correctly classify each assignment decision and did not take into account the severity of a wrong decision in terms of split and merge errors in the result. Training on a TED approximation overcomes this problem.

It is worth noting that the boundary shift we used as a tolerance criterion is just one example of how to use the TED for training and evaluation. Depending on the biological question, more or less deviations from the ground truth can be permitted. For example, boundary shifts could be tolerated to an extent that locally depends on the diameter of the ground truth neuron. In future work, it will be interesting to investigate the use of the TED for more general biomedical image processing problems with more specific tolerance criteria.

## Acknowledgements

We thank the reviewers for their valuable suggestions. This work was funded by SNF grant 149861.

## References

- [1] D.R. Martin, C.C. Fowlkes, D. Tal, J. Malik, A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics, in: ICCV, vol. 2, 2001, pp. 416–423.
- [2] D.R. Martin, C.C. Fowlkes, J. Malik, Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues, IEEE Trans. Pattern Anal. Mach. Intell. (2004).
- [3] P. Arbeláez, M. Maire, C.C. Fowlkes, J. Malik, From Contours to Regions: An Empirical Evaluation, in: CVPR, 2009.
- [4] W.M. Rand, Objective criteria for the evaluation of clustering methods, J. Am. Stat. Assoc. 66 (1971) 846–850.
- [5] R. Unnikrishnan, M. Hebert, Measures of Similarity, in: Seventh IEEE Workshop on Applications of Computer Vision, 2005.
- [6] R. Unnikrishnan, C. Pantofaru, M. Hebert, Toward objective evaluation of image segmentation algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 29 (6) (2007) 929–944.
- [7] G.E. Sosinsky, J. Crum, Y.Z. Jones, J. Lanman, B. Smarr, M. Terada, M.E. Martone, T.J. Deerinck, J.E. Johnson, M.H. Ellisman, The combination of chemical fixation procedures with high pressure freezing and freeze substitution preserves highly labile tissue ultrastructure for electron tomography applications, J. Struct. Biol. 161 (3) (2008) 359–371.
- [8] A. Cardona, Towards semi-automatic reconstruction of neural circuits, Neuroinformatics 11 (1) (2013) 31–33.

<sup>6</sup> Source code available at <http://github.com/funkey/sbmmr>.

- [9] J. Funke, B. Andres, F.A. Hamprecht, A. Cardona, M. Cook, Efficient Automatic 3D-Reconstruction of Branching Neurons from EM Data, in: CVPR, 2012, pp. 1004–1011.
- [10] V. Kaynig, A. Vazquez-Reina, S. Knowles-Barley, M. Roberts, T.R. Jones, N. Kasthuri, E. Miller, J. Lichtman, H. Pfister, Large-scale automatic reconstruction of neuronal processes from electron microscopy images, *Med. Image Anal.* 22 (1) (2015) 77–88.
- [11] J. Nunez-Iglesias, R. Kennedy, T. Parag, J. Shi, D.B. Chklovskii, Machine learning of hierarchical clustering to segment 2D and 3D Images, *Plos One* 8 (8) (2013) e71715.
- [12] T. Parag, S.M. Plaza, L.K. Scheffer, Small Sample Learning of Superpixel Classifiers for EM Segmentation-Extended Version, in: CoRR, vol. abs/1406.1, 2014.
- [13] G.B. Huang, V. Jain, Deep and Wide Multiscale Recursive Networks for Robust Image Labeling, in: arXiv preprint arXiv:1310.0354, 2014.
- [14] J. Funke, J. Klein, F. Moreno-Noguer, A. Cardona, M. Cook, Structured Learning of Assignment Models for Neuron Reconstruction to Minimize Topological Errors, in: ISBI, 2016.
- [15] H. Peng, P. Chung, F. Long, L. Qu, A. Jenett, A.M. Seeds, E.W. Myers, J.H. Simpson, BrainAligner: 3D registration atlases of drosophila brains, *Nat. Methods* 8 (6) (2011) 493–498.
- [16] W. Denk, K.L. Briggman, M. Helmstaedter, Structural neurobiology: missing link to a mechanistic understanding of neural computation, *Nat. Rev. Neurosci.* (2012).
- [17] M. London, M. Häusser, Dendritic computation, *Annu. Rev. Neurosci.* 28 (1) (2005) 503–532.
- [18] T. Kröger, S. Mikula, W. Denk, U. Köthe, F. A. Hamprecht, Learning to segment neurons with non-local quality measures, in: MICCAI, vol. 16, 2013, pp. 419–427.
- [19] J. Funke, J. Martel, S. Gerhard, B. Andres, D.C. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, H. Pfister, A. Cardona, M. Cook, Candidate sampling for neuron reconstruction from anisotropic electron microscopy volumes, in: MICCAI, 2014, pp. 17–24.
- [20] V. Jain, B. Bollmann, M. Richardson, D.R. Berger, M. Helmstaedter, K.L. Briggman, W. Denk, J.B. Bowden, J. Mendenhall, W.C. Abraham, K. Harris, N. Kasthuri, K.J. Hayworth, R. Schalek, J. Tapia, J. Lichtman, S.H. Seung, Boundary Learning by Optimization with Topological Constraints, in: CVPR, 2010.
- [21] H. J. H. Jiang, S. Fels, J. Little, A Linear Programming Approach for Multiple Object Tracking, in: CVPR, 2007.
- [22] F. Jug, T. Pietzsch, D. Kainmüller, J. Funke, M. Kaiser, E. van Nimwegen, C. Rother, E.W. Myers, Optimal Joint Segmentation and Tracking of Escherichia Coli in the Mother Machine, in: BAMBI, 2014.
- [23] A. Vazquez-Reina, D. Huang, M. Gelbart, J. Lichtman, E. Miller, H. Pfister, Segmentation Fusion for Connectomics, in: ICCV, 2011.
- [24] B.X. Kausler, M. Schiegg, B. Andres, M. Lindner, U.Köthe, H. Leitte, J. Wittbrodt, L. Hufnagel, F.A. Hamprecht, A discrete chain graph model for 3d+t cell tracking with high misdetection robustness, in: ECCV, 2012.
- [25] S. Gerhard, J. Funke, J. Martel, A. Cardona, R.D. Fetter, Segmented anisotropic ssTEM dataset of neural tissue, 2013.
- [26] I. Arganda-Carreras, S.C. Turaga, D.R. Berger, D. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, D. Laptev, S. Dwivedi, J.M. Buhmann, T. Liu, M. Seyedhosseini, T. Tasdizen, L. Kamentsky, R. Burget, V. Uher, X. Tan, C. Sun, T.D. Pham, E. Bas, M.G. Uzunbas, A. Cardona, J. Schindelin, H.S. Seung, Crowdsourcing the creation of image segmentation algorithms for connectomics, *Front. Neuroanat.* 9 (2015) 142.
- [27] I. Tsochantaridis, T. Joachims, T. Hofmann, Y. Altun, Y. Singer, Large margin methods for structured and interdependent output variables, *J. Mach. Learn. Res.* 6 (2005) 1453–1484.
- [28] D.C. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images, in: NIPS, vol. 25, 2012, pp. 2843–2851.



## TrackMate: An open and extensible platform for single-particle tracking



Jean-Yves Tinevez<sup>a,\*</sup>, Nick Perry<sup>a,1</sup>, Johannes Schindelin<sup>b,2</sup>, Genevieve M. Hoopes<sup>c</sup>, Gregory D. Reynolds<sup>c</sup>, Emmanuel Laplantine<sup>d</sup>, Sebastian Y. Bednarek<sup>c</sup>, Spencer L. Shorte<sup>a</sup>, Kevin W. Eliceiri<sup>b,e</sup>

<sup>a</sup>Imagopole, Citech, Institut Pasteur, 75724 Paris, France

<sup>b</sup>Laboratory for Optical and Computational Instrumentation, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>c</sup>Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>d</sup>Laboratory of Signaling and Pathogenesis, Centre National de la Recherche Scientifique, UMR 3691, Institut Pasteur, 75724 Paris, France

<sup>e</sup>Morgridge Institute for Research, Madison, WI 53719, USA

### ARTICLE INFO

#### Article history:

Received 15 August 2016

Received in revised form 29 September 2016

Accepted 30 September 2016

Available online 3 October 2016

#### Keywords:

Single-particle tracking

Phototoxicity

Clathrin-mediated endocytosis

Image analysis

Open-source software

Microscopy

### ABSTRACT

We present TrackMate, an open source Fiji plugin for the automated, semi-automated, and manual tracking of single-particles. It offers a versatile and modular solution that works out of the box for end users, through a simple and intuitive user interface. It is also easily scriptable and adaptable, operating equally well on 1D over time, 2D over time, 3D over time, or other single and multi-channel image variants. TrackMate provides several visualization and analysis tools that aid in assessing the relevance of results. The utility of TrackMate is further enhanced through its ability to be readily customized to meet specific tracking problems. TrackMate is an extensible platform where developers can easily write their own detection, particle linking, visualization or analysis algorithms within the TrackMate environment. This evolving framework provides researchers with the opportunity to quickly develop and optimize new algorithms based on existing TrackMate modules without the need of having to write *de novo* user interfaces, including visualization, analysis and exporting tools.

The current capabilities of TrackMate are presented in the context of three different biological problems. First, we perform *Caenorhabditis-elegans* lineage analysis to assess how light-induced damage during imaging impairs its early development. Our TrackMate-based lineage analysis indicates the lack of a cell-specific light-sensitive mechanism. Second, we investigate the recruitment of NEMO (NF- $\kappa$ B essential modulator) clusters in fibroblasts after stimulation by the cytokine IL-1 and show that photodamage can generate artifacts in the shape of TrackMate characterized movements that confuse motility analysis. Finally, we validate the use of TrackMate for quantitative lifetime analysis of clathrin-mediated endocytosis in plant cells.

© 2016 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Over the past 30 years, the biological sciences have enjoyed tremendous advances in imaging techniques that provided new insights into dynamic phenomena. Paramount to this significant progress has been the development of novel imaging hardware platforms and fluorescent probes for the visualization of biological

processes at the level of whole organisms, single cells, and subcellular. The development of analysis software to extract quantitative data on the dynamics of the processes examined by microscopy has been equally important. In particular, single-particle tracking software – computational tools that can follow objects in a time-lapse movie and quantify their dynamics – are crucial for almost any experiments involving live-cell imaging, and are a critical part of the researcher toolbox.

A number of programs have been designed and have undergone significant refinement over the past several years (some reviewed in [1–3]) for single particle tracking of live cell microscopy images. Furthermore, the importance of tracking in bio-imaging was highlighted by the 2014 IEEE International Symposium on Biomedical Imaging (ISBI) Grand Challenge [4] in which a number of the current approaches for particle tracking were objectively compared.

**Abbreviations:** GUI, Graphical User Interface; LSCM, Laser Scanning Confocal Microscope; SDCM, Spinning-Disk Confocal Microscope; TIRFM, Total-Internal Reflection Fluorescence Microscope; VAEM, Variable Angle Epifluorescence Microscopy; QE, Quantum Efficiency.

\* Corresponding author.

E-mail address: [tinevez@pasteur.fr](mailto:tinevez@pasteur.fr) (J.-Y. Tinevez).

<sup>1</sup> Present address: Palantir Technologies, USA.

<sup>2</sup> Present address: Microsoft GmbH, Germany.

<http://dx.doi.org/10.1016/j.ymeth.2016.09.016>

1046–2023/© 2016 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

One important and significant outcome of the ISBI Grand Challenge [4] was that there does not exist at this time a “one-size-fits-all”, universally accurate tracking method. Many specific dynamic biological processes require their own specialized tracking tools that can derive and exploit unique aspects of process such as the motion or shape of the tracked objects. This makes the problem of developing a universal tracking method even more challenging; as the accuracy of tracking is inherently dependent on specific parameters, one strategy that works well for a given problem is likely to fail for another one. Therefore, for many live imaging studies, the data analysis will involve generating *de novo* or adapting an existing tracking software method tailored for a single, specific application. However, the efforts required to build tracking software from scratch are considerable. Its development must include: 1) detection tools that can extract objects from images, 2) linking tools that can track these objects over time, 3) visualization tools that can overlay the raw data and tracking results, and 4) analysis tools that can evaluate the results. Additionally, if the goal is to develop a tracking tool that will be widely used within the research community, the software must be user-friendly, developer friendly, modular, validated and interoperable. Given all these hurdles, it is not surprising that only a few tracking algorithms exist [5–13] that are widely available and designed to be useful for the general needs of the greater Life-Science community.

TrackMate is a plugin within the Fiji ImageJ distribution [14] for tracking developed with several concomitant goals. First, it focuses on usability by providing Life-Science researchers with a turnkey user-friendly tracking solution, fulfilling the recommendations of [15]. It is openly available and well documented, and it houses several detection and tracking modules that allow combining manual and automated particle tracking approaches. TrackMate includes several visualization tools as well as other features that facilitate the export and exchange of data and results with other tracking tools and/or analysis applications. Second, TrackMate has been designed for maximal flexibility. The capabilities of TrackMate can be tailored by the user through the addition of specific tracking, detection, visualization, or analysis modules. As a result, researchers can quickly develop a solution suited for their specific application, while taking advantage of existing modules, thus accelerating and facilitating development. Finally, TrackMate uses a data model that makes it a useful tool for a wide range of tracking applications, ranging from single-particle tracking of subcellular organelles to cell lineage analysis. Below we describe the features of TrackMate in more detail and illustrate their application towards three biological use cases that involve tracking and quantification of dynamic live-cell events.

## 2. Material and methods

### 2.1. TrackMate: an open and extensible platform for single-particle tracking

#### 2.1.1. Purpose

TrackMate is a Fiji [14] plugin available through this software platform, and therefore benefits from the facilities offered by Fiji for image input/output and preprocessing. TrackMate’s main entry point is an interactive plugin where tracking is performed using a wizard-like GUI (Fig. 1). The user is guided through several stages, each of which constitutes a step in the tracking process, choosing what algorithm to use for detection and tracking, and then configuring them. The result of each step is displayed immediately. This allows the user to readily navigate back to readjust settings should the process output be deemed unsatisfactory. Various tools allow for the inspection of intermediate results: a 3D view of the data and the results can be launched, and each particle or track can

be annotated using specific colors to denote numerical values of interest, including *e.g.* particle quality, estimated diameter, track length, track displacement, *etc.* The use of TrackMate is documented in the [Supplemental Information](#) of this article.

Manual inspection of tracking results is critical, particularly in practical cases when quantitative metrics are missing, to assess the accuracy of the results. TrackMate contains several visualization tools to ease inspection and facilitate manual curation or editing of the results. The main visualization tool (Fig. 2, top left) overlays the tracking results on the raw image, re-using the ImageJ hyperstack display. A 3D view of the tracking and raw data can also be launched (Fig. 2, bottom left), based on the ImageJ 3D viewer [16]. Finally, TrackScheme was developed to inspect the higher-level structure of tracks (Fig. 2, top and bottom right), which is particularly handy for cell lineages. Much like a train schedule, this tool displays the tracks based only on their links, and lays out spots from bottom to top according to time. Each branching event generates a new vertical lane. Several synchronized views of the same data can be launched and selecting a subset of spots and links will highlight them in all the views.

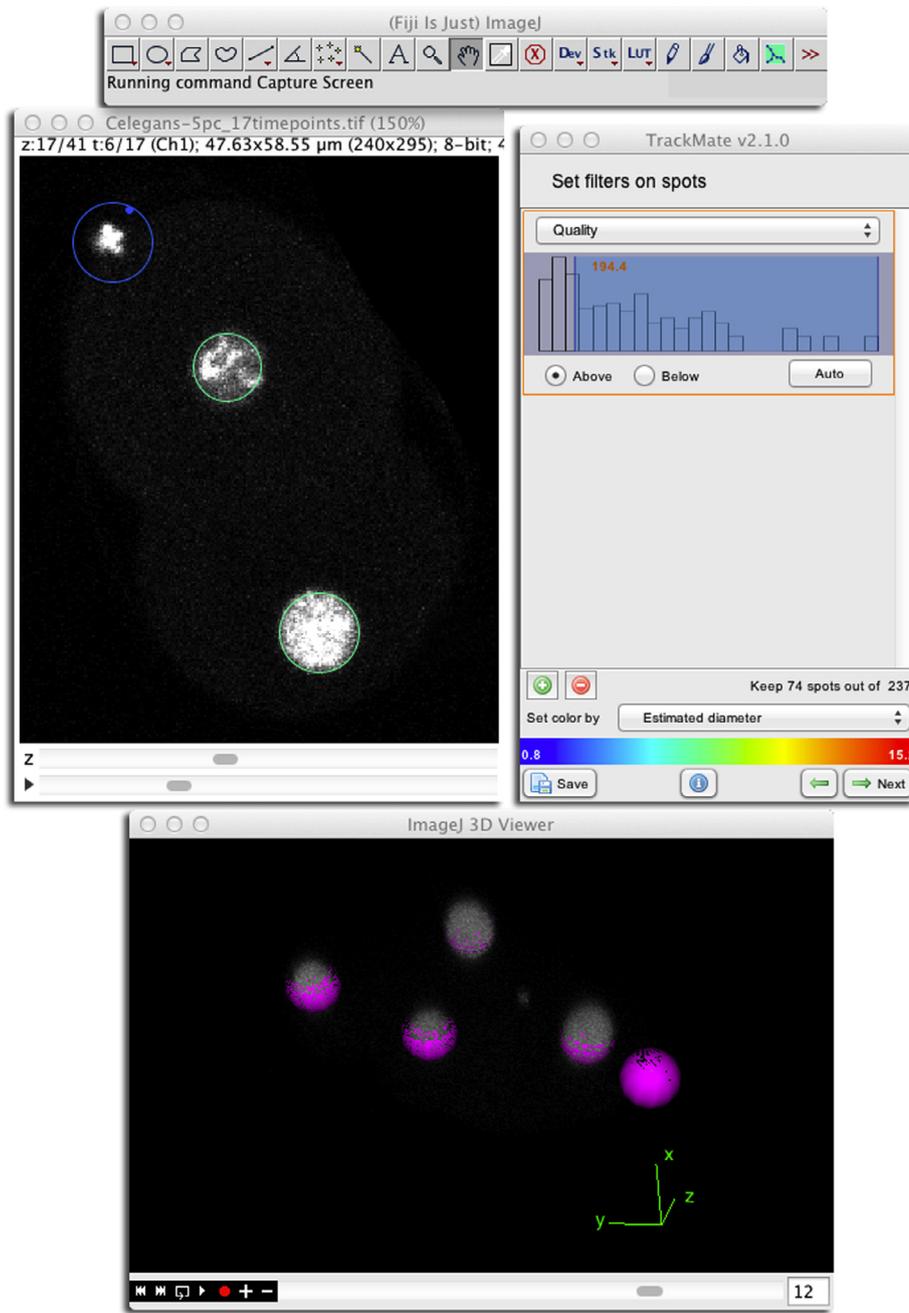
The main view and TrackScheme also allow direct manual editing of the tracking results. Spots can be moved, added or deleted, and linked or unlinked to tracks. In addition, the spot size can be adjusted, and their names can be changed in batch or individually. Two numerical features keep track of whether a spot or a link has been manually modified.

It is actually possible to entirely skip the detection step or the particle-linking step or both, and generate tracking results *via* manual annotation. A semi-automated tracking tool facilitates manual annotation, by creating the most likely track birthing from a single spot. Any manual modification of the data triggers a recalculation of all the numerical features, so that they are always in sync with the data. Though TrackMate is only a single-particle tracking tool, it ships with basic analysis facilities. Numerical features are calculated for spots, links and tracks automatically and kept up to date when the user manually edits tracks. These numerical features are also used in the GUI for filtering and display and can also be combined in various plots or data tables.

#### 2.1.2. Data model and capabilities

The data model used to store tracking results has constraints strongly affecting what kind of event can be detected at the particle-linking step. A linear data structure such as an array can only represent linear tracks that do not have division or merge events, where two objects fuse into one. Missing detections (also called gaps) however can be handled by storing spot indices in these arrays. In TrackMate the tracking results are represented as a directed simple graph, where vertices are filtered spots generated at the detection step, and edges are links generated at the particle-linking step. This data structure is very flexible and has limited restrictions making it well suited to single-particle tracking. A graph is a structure where vertices (in our case spots) are connected by edges (in our case links across time). In a simple directed graph, a link cannot go from one spot to the same spot and there cannot be more than one link between two spots. Links have a direction going from source spots earlier in time to target spots later in time, and links between two spots in the same frame are forbidden. With such a data structure, a spot can be the source and target to several links, possibly handling cell division events and particle merging events as well.

The generality of this data structure delegates specificity to the subsequent analysis tools. TrackMate can be configured to detect linking, gap-closing, splitting, or merging events and makes no assumption of the biological significance of these events. The subsequent analysis must therefore be tailored to the biological



**Fig. 1.** The TrackMate user interface for automated tracking, depicting the tracking the cells of a developing *C. elegans* embryo. The illustration displays from top to bottom the Fiji toolbar, the image data overlaid with the detection step result and the TrackMate GUI. The central part of the GUI contains the contextual commands for the current tracking step. Here, the user filters out spurious spots based on a quality threshold just after the execution of the detection step. Several filters can be stacked. Detection results, or spots, are represented as spheres with an initially constant radius. Here, the user chose a coloring scheme that reflects an estimate of the true spot radius. The bottom part of the GUI allows the user to navigate forward and backward through the tracking steps, to see a text log of the plugin activity and to save the data at anytime.

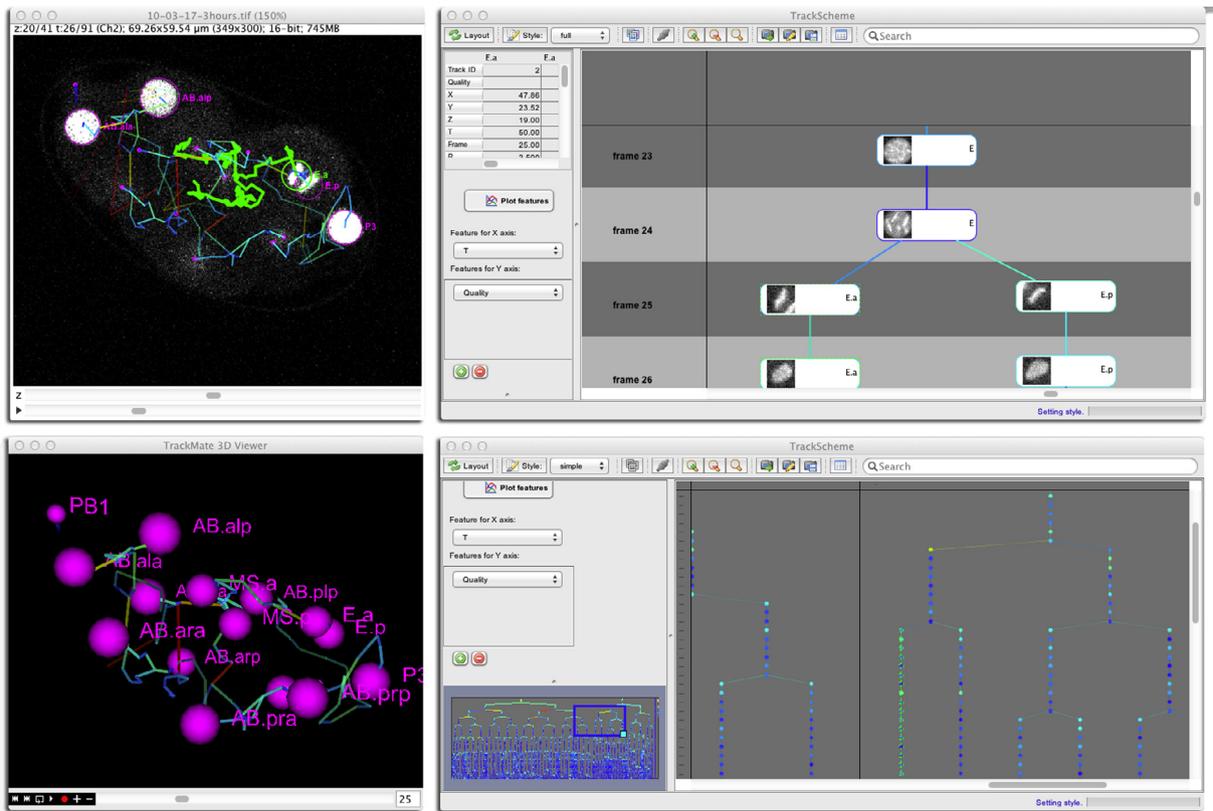
problem at hand. TrackMate aims at being as general as possible and encourages end-users to tailor algorithms and analyses specific to their biological problems.

### 2.1.3. Performance and accuracy

TrackMate ships three classes of particle-linking algorithms. A first class is derived from the LAP framework proposed by Jaqaman and colleagues [7]. Base linking costs are calculated from the square distance between particles, which makes it ideal to tackle Brownian motion [10]. However, costs can be modulated by feature value differences, penalizing the linking of particles that are

different in intensity distribution, rough shape, etc. A convenient GUI allows tuning these costs directly in TrackMate. A second particle-algorithm relies on the Kalman filter [17] to tackle linear motion. Finally, particle linking based on nearest-neighbor search is proposed as the simplest linking algorithm.

Accuracy measurements serve as a tool for end-users to choose the optimal algorithm for their specific biological applications. The accuracy of the particle-linking algorithms or spot trackers offered in TrackMate is documented in the [Supplemental Information](#). To establish it, we relied on the ISBI Grand Challenge single-particle tracking dataset [4].



**Fig. 2.** Four views of the same *C. elegans* embryo being tracked in TrackMate. From left to right then top to bottom: 1. The main TrackMate view, that overlays the tracking results on the raw image window in Fiji. The lineage of the *E.a* founder cell is selected and can be seen as a thick green line. 2. TrackScheme, the track visualization tool of TrackMate. It lays out the tracks in a time oriented hierarchical graph. Here it is centered on the *E* founder cell division and shows the cell names and heir thumbnail. 3. The 3D viewer here shows the TrackMate annotation with cell names. 4. Another instance of TrackScheme, with a higher-level view of the lineage. The edge color encodes the cell displacement in all views from 0  $\mu\text{m}$  (blue) to 5  $\mu\text{m}$  (red). Note that all views are synchronized and share the same selection. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

#### 2.1.4. Using TrackMate in scripts and other applications

TrackMate has a modular design that allows using its processing core without relying on the GUI, in scripts or other software. This allows TrackMate to be used for batch analysis, potentially running on a remote cluster over many images at once. The [Supplemental Information](#) details the core API of TrackMate and gives examples of scripts in Python and MATLAB. A subset of TrackMate capabilities can also be called using the ImageJ macro language.

#### 2.1.5. Interoperability with other software

TrackMate is deployed as a Fiji plugin, and therefore benefits from and interacts with the whole ImageJ ecosystem [18], whether for data import and preprocessing or export and scripting. TrackMate focuses on single-particle tracking, and therefore its end results take the shape of tracks, collections of spots and links. It offers basic track analysis and visualization capabilities, but specific applications are likely to require subsequent, finer analysis, which requires TrackMate files to be interoperable with specialized software. The Fiji distribution of TrackMate ships the files required to import TrackMate results in MATLAB, documented in the [Supplemental Information](#). A specific tool in TrackMate also allows the export of results to a format readable by the Icy software [8]. Alternatively, the Track Manager tool of Icy has an import filter for the TrackMate file format.

#### 2.1.6. Extending and reusing TrackMate

TrackMate is modular in design. Each step in the tracking application corresponds to a module in TrackMate, organized by common interfaces. There are seven types of TrackMate mod-

ules, including data visualization, particle detection, particle analysis, particle linking, link analysis, track analysis, and generic actions.

This modular design is made to be extensible. External developers can create their own TrackMate module to fill their own needs. By extending common TrackMate interfaces, they will be integrated in the GUI without distinctions from native modules. TrackMate discovers these new modules automatically thanks to the SciJava annotation framework, which considerably simplifies development. The new modules just need to be annotated with SciJava and compiled as java *jar* files. By dropping the *jar* file in the *plugins* folder of a Fiji installation, they will be automatically discovered by TrackMate and will appear in the GUI. Interestingly, this completely decouples the TrackMate developers from external developers, who can build new modules without requiring intervention. The [Supplemental Information](#) documents how to write custom TrackMate modules, covering each of the seven types of module.

#### 2.2. TrackMate and multidimensional imaging

TrackMate was designed for use with multidimensional light microscopy datasets from a wide range of modalities including simple 2D brightfield collection, 3D TIRF imaging, and 4D (space over time) laser scanning microscopy. The program was originally developed for *C. elegans* lineage analysis [19], but has been adapted and widely utilized for a number of other biological tracking applications. Given the wide number of current and possible dynamic imaging uses for TrackMate, in this manuscript we chose to focus

on three applications that showcase the different strengths, features, and potential of TrackMate.

In the first case study, we examine how laser scanning confocal microscopy (LSCM) induced phototoxicity affects cell division in *C. elegans*. By utilizing TrackMate-based cell lineage tracking we can determine whether different imaging doses or techniques can influence the normal cell lineage. This example shows how TrackMate can handle automatic analysis and manual correction. There are many use cases where a fully automatic strategy is not possible due to factors such as variable signal to noise (SNR) in the images. TrackMate was specifically designed to handle 3D data such as this and do automatic – and manual – based cell lineage tracking and analysis.

In the second example we use TrackMate to investigate the ability of NEMO (NF- $\kappa$ B essential modulator), a key component of the NF- $\kappa$ B signaling pathway, to assemble into punctate structures, that we refer as “NEMO dots”, upon cell stimulation by cytokines, and how phototoxic effects alter this process even at low illumination doses. NEMO has been shown to be involved in many physiological processes and we have shown that its assembly into punctate structures is an important event of the cell response to pro-inflammatory cytokine stimulation [20]. The formation of NEMO dots is a transient event that occurs rapidly and can be tracked in 2D (using SDCM). A key need for this application is the ability to use the tracking results in subsequent analyses using biophysics methods, revealing the key ability of TrackMate to interoperate with other analysis tools, such as MATLAB.

The third example focuses on clathrin dynamics at the plasma membrane of *Arabidopsis* hypocotyl epidermal cells. In particular, we assess the ability of TrackMate to track the transient recruitment of fluorescent fusion protein tagged clathrin light chain (CLC-FP) into diffraction-limited structures at the plasma membrane and their subsequent release using Variable Angle Epifluorescence Microscopy or TIRFM. This example illustrates an important trafficking process and the utility of TrackMate in both plant and mammalian models. This use case demonstrates the ability of TrackMate to handle 3D spatial and temporal data, as well as the complex modalities of VAEM/TIRFM time-lapse images. Importantly, in this case study we compare the ability of TrackMate to determine the lifetime of clathrin-coated plasma membrane structures relative to previously-published, manually-tracked CLC-FP foci. Thus, this case study provides a useful validation of TrackMate’s tracking accuracy.

### 2.3. *C. elegans* embryo imaging

*C. elegans* embryo imaging was done following the phototoxicity assessment protocol described previously [19]. Briefly: using a LSCM as imaging device (LSM700, Carl Zeiss, Jena, Germany) equipped with a 63x oil NA = 1.4 objective, embryos from the strain AZ212 were imaged at 21 °C. One acquisition made of 41 Z-slices spaced by 1  $\mu$ m were acquired every 2 min, for at least 2 h starting from the first anaphase. The pixel dwell time was chosen to be 1.58  $\mu$ s/pixel, and the laser power varied to probe different light doses. The light dose  $L$  is calculated as the total energy deposited on the sample for one acquisition; that is:

$$L = P \times dt \times N_{\text{equatorial}} \times N_{\text{Zslices}}$$

where  $P$  is the laser power measured before the objective, assuming nearly perfect transmittance;  $dt$  is the pixel dwell time,  $N_{\text{equatorial}}$  is the number of pixels in the embryo equatorial plane, and  $N_{\text{Zslices}}$  is the number of Z-slices scanned for a single acquisition.

### 2.4. NEMO punctate cluster imaging

GFP-NEMO-expressing cells were prepared for live imaging as in [20] and plated in 3.5 cm diameter glass-bottom petri dishes. Two different SDCMs were used for imaging: an UltraVIEW ERS and an UltraVIEW VOX (Perkin-Elmer), based on CSU-22 and CSU-X1 spinning-disks (Yokagawa) respectively, with a EM-CCD camera with a peak QE around 60% (C9100-50, Hamamatsu) and 92% (ImageEM-X2, Hamamatsu), again respectively. A single plane close to the coverslip was chosen for imaging, and the cells were imaged with a PlanApochromat 63x 1.4NA oil objective at 2 frames per second, at 37 °C and 5% CO<sub>2</sub>, using a 488 nm laser line and filters suited for the detection of the GFP emission. Stimulation by IL-1 was done at a final concentration of 10 ng/mL.

### 2.5. VAEM imaging of PM-associated Clathrin dynamics

Sterilized stable transgenic *A. thaliana* ecotype *Wassilewskija* seeds expressing functional Clathrin Light Chain 2-GFP under control of its native promoter [21] were plated on 0.5x Murashige and Skoog Basal Salts 0.5% agar. After vernalization at 4 °C for 72 h, plates were incubated under continuous light at 22 °C for 3 h before being grown in the dark for 5 days. Seedlings were imaged by Variable Angle Epifluorescence Microscopy as described [22] using a Nikon N-STORM Microscope with a motorized TIRF illuminator and Andor iXon Ultra 897 EMCCD camera. Using the TIRFM mode and the 100x oil NA = 1.49 objective, 2D time-lapse sequences were taken at 2 frames per second with 488 nm laser excitation.

## 3. Results

### 3.1. *C. elegans* embryo cell division synchrony and cell cycle length are robust against photodamage

#### 3.1.1. Methods

We recently proposed a quantitative method to assess the phototoxic impact of imaging devices [19]. *C. elegans* embryos have an invariant development pattern, with minimal variation across individuals for the cell lineage, their organization in space, and their division timing [23]. At a temperature of 21 °C and without external perturbation, an embryo develops into a 50-cells organism in two hours measured from the first anaphase. If this process is imaged using a fluorescent strain labeled for nuclei [24], harsh imaging conditions will perturb the normal development and hinder normal progression. For intermediate light doses, this manifests as a delayed development, and less than 50 cells are found after two hours of imaging. We exploited this effect to quantify phototoxic impact: the phototoxic light dose is defined as the energy per acquisition that yield an embryo with 25 cells after 2 h of development at 21 °C. To measure this dose, several embryos were imaged in controlled conditions, changing only the illumination power. The number of cells after 2 h of imaging was plotted against the light dose, and the phototoxic light dose was derived by a fit by a sigmoidal curve [19]. This protocol characterizes the phototoxic impact of an imaging device when imaging 3D specimens over several hours, revealing sensitivities to subtle phototoxic effects. With the laser line used for excitation and the fluorescent labeling of the sample, it reports mainly the fluorescence-sensitized phototoxicity and is not sensitive to any thermal effects that occur for near IR/IR excitation.

While this protocol benchmarks the performance of a system through a bulk measurement on a specimen, determining how photodamage affects specimens requires even deeper insights. We investigate here whether specific cells of the lineage have

different sensitivities to photodamage. This involves building the comprehensive cell lineages, as the embryos sample different light doses. While this task is nowadays greatly facilitated by the pioneering work of other groups [5,25–27], the tools they offer are unfortunately not applicable to our problem. Indeed, the images we generate have a quality that is not adjustable: it is the product of the sensitivity of the system tested and of the light dose probed. For a strong illumination power, the fluorophore quickly bleaches, while this does not happen or not as much for low incident power. A single dataset therefore may have movies with a very low and constant SNR for low light doses, and movies beginning with a strong SNR that quickly decays until the images are no longer exploitable for high light doses (Supplemental Movie 1). To add to this quality discrepancy, a single image of the embryo also shows varying SNR as the Z position of cells vary. Indeed, as we use an oil objective, the image quality quickly deteriorates as the focus moves away from the coverslip. This problem is classically solved by varying the illumination power as the focus moves [25], but again this is not applicable to our study since we want to use a constant illumination power throughout a single experiment.

It is important to note that the extremely high variability in image quality, both in space and time and within a single dataset, is there by construction. As the illumination power is an input variable in our experiments we do not have the liberty to tune it in order to reach a desirable image quality. Rather than develop a highly customized solution that would deal with this issue, we exploited the ability of TrackMate to combine automated and manual annotation approaches. Each movie was first segmented automatically for the cells, and the spurious or missing detections were manually corrected. The curated detection results were then tracked automatically, and tracking mistakes were again manually corrected. The data model of TrackMate directly allows generating lineages, cell divisions being represented by branching events with two links emerging from a mother cell spot. Cells were named following the Sulston and Horvitz convention [28]. Movies with quickly decaying SNR were those for which the automated processing was the less successful, but because they correspond to the toxic light doses, they yielded few cells and the manual correction

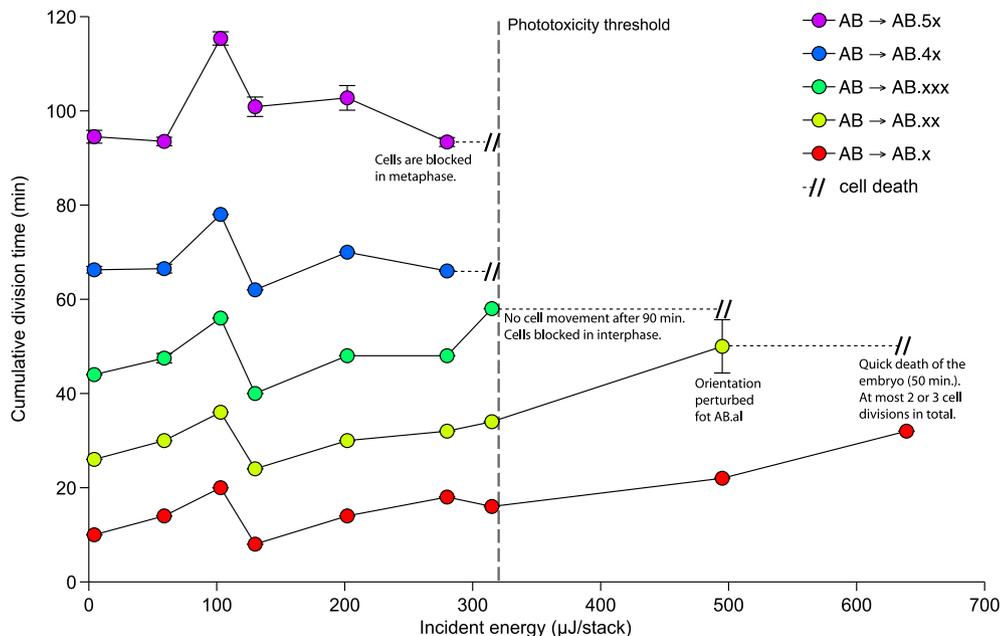
took little time. On average, a single acquisition of at least 2 h of development could be lineage in less than an afternoon. The dataset presented here includes 9 acquisitions, collected on a LSM, using light doses ranging from 13  $\mu\text{J}/\text{stack}$  to 650  $\mu\text{J}/\text{stack}$ , covering situations where no to extreme phototoxic effects can be observed. A full lineage from this dataset is presented in Supplemental Fig. 2.

### 3.1.2. Results

The phototoxicity threshold as defined in [19] was found to be 320  $\mu\text{J}/\text{stack}$  for this LSM, to be compared with the much larger value around 5 mJ/stack that can be measured for epifluorescence-based systems [19,29]. For values well below the phototoxicity threshold, up to 200  $\mu\text{J}/\text{stack}$ , the development of the embryo was virtually non-perturbed (Supplemental Movie 1 and Fig. 3). We investigated how photodamage impacts cell division synchrony and cell cycle length. To do so, we focused on the AB descendants. At 21 °C without imaging, the first two hours of the AB lineage was made of 5 successive divisions that gave rise to 32 descendants. Each of these 5 divisions was well synchronized across sibling cells [23], which allowed for the definition of the cumulative division time for  $AB_{n_x}$ , the time measured from the first anaphase to the  $n$ th division in the AB lineage. This time was characterized by then mean and standard deviation over the  $n$  times of division of the  $n$   $AB_{n_x}$  cells into  $2n$   $AB_{(n+1)_x}$  daughter cells.

We first investigated whether or not this synchrony is perturbed by harsh illumination. The cumulative division time is reported in Fig. 3 as a function of the incident light dose. We observed that division synchrony of the AB lineage resists phototoxicity well, as indicated by the very low standard deviation on the division time. Even for disruptive light doses, the AB descendants all divided within 10 min of their siblings. However, close to the phototoxicity threshold and beyond, some cell divisions did not happen at all. This shows that cell division synchrony is robust against photodamage, and more robust than cell division itself.

Surprisingly, cell cycle lengths only vary weakly with light doses. For doses lower than the phototoxicity threshold, the cumulative time to get a specific AB descendant division stayed roughly



**Fig. 3.** Cumulative cell division time for the AB lineage, measured from the first anaphase in *C. elegans* embryos up to two hours as a function of the incident light dose. Cell division time is defined with mean and standard deviation taken over the times of division of all  $AB_{n_x}$  cells into  $AB_{(n+1)_x}$ . A column in this plot represents a single embryo followed over time. Nine embryos were tracked to sample a wide-range of light doses. When a particular cell division does not happen, the assessed reason and notable events are shown as annotations in the plot.

constant (Fig. 3). The observation of [23] in non-invasive conditions still holds, and cell cycle length follows roughly a geometrical sequence, even if it increases slowly with incident light dose. Over the whole range of light doses sampled, the cell cycle time of the AB progenitor doubled. However, this immediate lengthening of division times is not the main feature of phototoxic effects. The embryo development was hindered because for large incident doses, some subsequent cell divisions did not happen. Looking closely at the AB lineage, we saw that for 289  $\mu\text{J}$ , the  $AB_{4x} \rightarrow AB_{5x}$  cell division took place, but the cells were blocked in metaphase. For 324  $\mu\text{J}$ , the  $AB_{3x} \rightarrow AB_{4x}$  did not happen. For 504  $\mu\text{J}$ , the embryo stopped after  $AB_x \rightarrow AB_{xx}$  and for 648  $\mu\text{J}$ , only the first division of AB happened before the embryo died. Phototoxic effects on a LSCM manifest as a brutal development halt at a certain time depending on the light dose. Before this time is reached, the cycles of cell divisions proceed more-or-less normally.

### 3.1.3. Discussion

The robustness of the synchrony and of the order of cell divisions indicates that there probably is not a cell-dependent sensitivity to phototoxic effects, at least for the AB descendants. Photodamage affects all cells in a similar manner. Embryonic cell proliferation involves rapid cell divisions through short cell cycles, most of which lack G1 and G2 phases [30]. By construction, our phototoxicity assay targets damage to the nucleus and is likely to cause first DNA damage. But the embryonic cell division misses robust checkpoints for DNA damage [30], which explains why the lengths of cell cycles only depend weakly on incident light doses. The catastrophic arrest of cell divisions might reflect harsher damage affecting the structural components of the mitosis. Therefore, even if we introduced as a sensitive specimen for our phototoxicity assay, our photodamage readout probably misses subtle DNA damage happening at low doses.

### 3.2. NEMO-IKK structures dynamics is adversely affected by subtle phototoxic damage

NEMO (NF- $\kappa\text{B}$  Essential Modulator) is a critical component of the NF- $\kappa\text{B}$  signaling pathway and is a key actor of many physiological processes such as immune response, inflammation, cell survival and proliferation [31]. We recently observed that upon stimulation by various cytokines, NEMO rapidly and transiently assembles into punctate structures [20]. Thanks to cell lines that stably express GFP-NEMO, this process can be investigated via imaging and single-particle tracking, opening the way to analyzing these structures via biophysical methods complementary to biochemical approaches. Using imaging we quantified the dynamics of the NEMO dots, showing that they experience a rapid turnover of their molecular content, and that they are anchored in the vicinity of the cell membrane [20]. These results are critical to our understanding of the process leading to the activation of the NEMO-IKK complex, as a part of the NF- $\kappa\text{B}$  pathway.

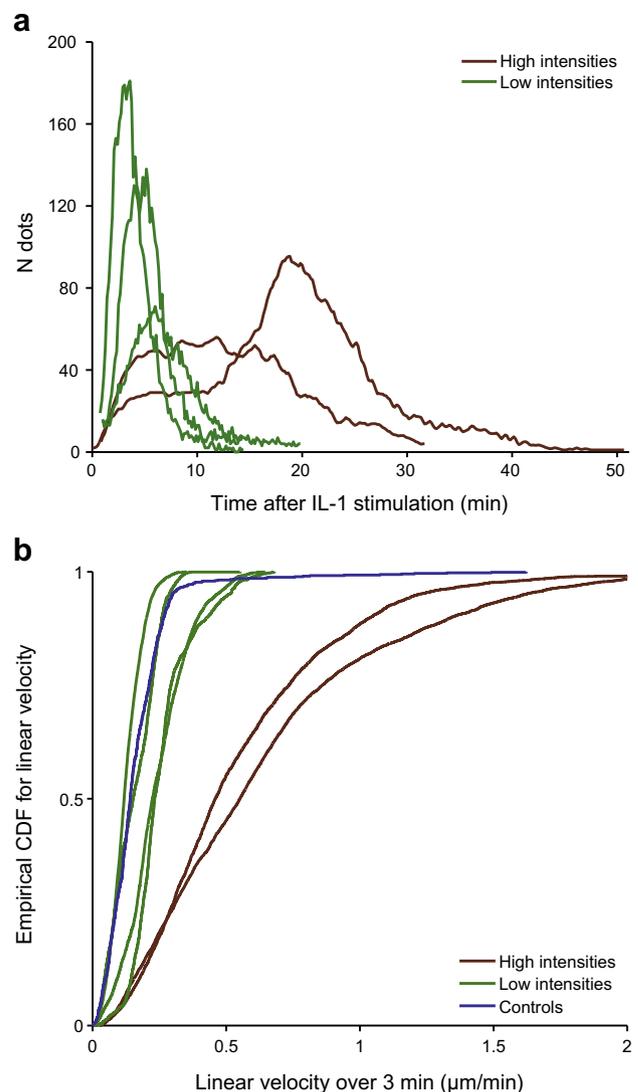
A downside of fluorescence microscopy as described in our first case study is that it can be invasive, altering the physiology of the organism and/or cells due to phototoxicity, thereby affecting the process under study. In particular it can affect the motion dynamics of organelles, potentially including the NEMO dots. We investigate here how phototoxicity impacts these dynamics and leads to an erroneous conclusion on the motion model of the NEMO subcellular structures.

#### 3.2.1. Results

Two SDCMs were used to image human cells constitutively expressing GFP-NEMO, each with a different light dose range (low intensities: 14–40  $\mu\text{J}$ , high intensities: 500–700  $\mu\text{J}$ ). Upon the addition of the IL-1 cytokine (10 ng/ml final), the NEMO

superstructures appear as bright, punctate, diffraction-limited dots over a fainter background. These dots were tracked using TrackMate until they disappeared. Subsequently, the tracks were imported and analyzed in MATLAB. Their time-course is displayed on Fig. 4a. In the case of high intensities, fewer dots were observed per cell, but they remained for a much longer time. For low intensities, they appeared and disappeared in less than 15 min measured from the addition of IL-1, whereas for high intensities, they lasted for more than 30 min.

Their motion characteristics also differed under high versus low intensity illumination. In both cases, they display random rapid movements, but over longer time scales in the case of high intensities, many of them display large, directed movements, sometimes over 10  $\mu\text{m}$ , preferably oriented towards the nucleus (Supplemental Movie 2). At low intensities, they all remained within 2  $\mu\text{m}$  of their average position for their whole lifetime (Supplemental Movie 3). The movement of the NEMO dots was quantified and



**Fig. 4.** a. NEMO dots dynamics over time, measured from the addition of IL-1 (10 ng/mL final), comparing high and low illumination intensities. Each line counts the number of NEMO dots in a single cell. b. Empirical cumulative distribution (CDF) of the linear velocity of NEMO dots for several cells after stimulation by IL-1. Linear velocity is defined for each NEMO dot as the median of all its 3 min-displacements divided by a 3 min interval. Each line is the CDF for all the NEMO dots of a single cell. The blue line, termed 'Controls', plots the CDF of the linear velocity of unspecific dots undergoing Brownian motion in a non-stimulated cell.

expressed in terms of linear velocity, a measure of how far a particle moves over a period of time. Here the linear velocity of a NEMO dot was defined as the median of all the 3 min-displacement of this dot, divided by an interval of 3 min.

The empirical cumulative distribution function for all cells is displayed in Fig. 4b. For low intensities, the linear velocity is in the range of 0.11–0.23  $\mu\text{m}/\text{min}$  (4 cells, 939 NEMO dots). This value is in the range of the linear velocity measured on dots of a control, non-stimulated cell, 0.14  $\mu\text{m}/\text{min}$  (1 cell, 17 dots). These dots are non-functional NEMO aggregates typically found in cells highly expressing GFP-NEMO, and undergo Brownian random motion. For high intensities the linear velocity ranges from 0.45 to 0.56  $\mu\text{m}/\text{min}$  (2 cells, 325 NEMO dots), with some values as large as 2  $\mu\text{m}/\text{min}$ . The fact that under high illumination NEMO dots can have large directed movements suggests that they are actively transported in these conditions.

To confirm this observation, we performed mean-square-displacement (MSD) analysis comparing NEMO dots under high and low illumination intensities. The fit of the log-log plot the MSD curves yielded a factor  $\alpha$  that determines whether the dots are bound to a fixed structure ( $\alpha < 1$ ), freely diffused ( $\alpha \approx 1$ ) or actively transported ( $\alpha > 1$ ) [32]. We retained dots for which the  $R^2$  value of the log-log fit is larger than 0.8. The values of all dots were pooled together for a cell, and we assessed the motion type for this cell using a *t*-test. We found that for high intensities, all cells have a  $\alpha$  value significantly larger than 1 ( $p = 10^{-3}$ – $10^{-4}$ ), indicating that NEMO dots are actively transported. However, the  $\alpha$  value for cells illuminated with low intensities have a  $\alpha$  value significantly lower than 1 ( $p < 10^{-10}$ ).

### 3.2.2. Discussion

The range of low versus high illumination intensities in this study is dictated by the sensitivity of the cameras equipping the two different microscope systems utilized. One was equipped with a very sensitive camera with high quantum efficiency, allowing for using low intensities to track NEMO superstructures. The other camera had a more modest quantum efficiency, and prompted for larger illumination intensities. However, even the high intensities used on the second system are not extremely high in absolute value. The phototoxic light dose measured on this system is about 21 mJ [19]. The equivalent light dose used in this study extrapolated over 2 min was in the range 60–80 mJ. We therefore used an illumination dose larger than the phototoxic threshold according to [19], but well within an order of magnitude. For low intensities, the equivalent light dose over 2 min is in the range 8.4–24 mJ. This difference was enough to separate results in two completely different categories, one of which is evidently corrupted by artifacts.

The NEMO dots display longer dynamics in cells exposed to high light intensities, which we attribute to artifactual cell stimulation by light-induced stress. The large displacements observed at high intensities are likely to be due to cell shrinking in response to phototoxicity. The dots may follow the cell membrane as the cell shrinks and display an artifactual motion, which resembles a directed movement towards the nucleus. An erroneous interpretation of these artifacts is that following stimulation by IL-1, NEMO is assembled into membrane-associated superstructures that are subsequently actively transported towards the nucleus. However, biochemical studies showed that the NEMO structures are most likely anchored to the cell membrane for their entire lifetimes, which is confirmed by MSD analysis when under non-invasive imaging condition [20]. Phototoxicity effects can give artifactual results that yield erroneous conclusions and strongly compromise interpretation of the role of a molecule.

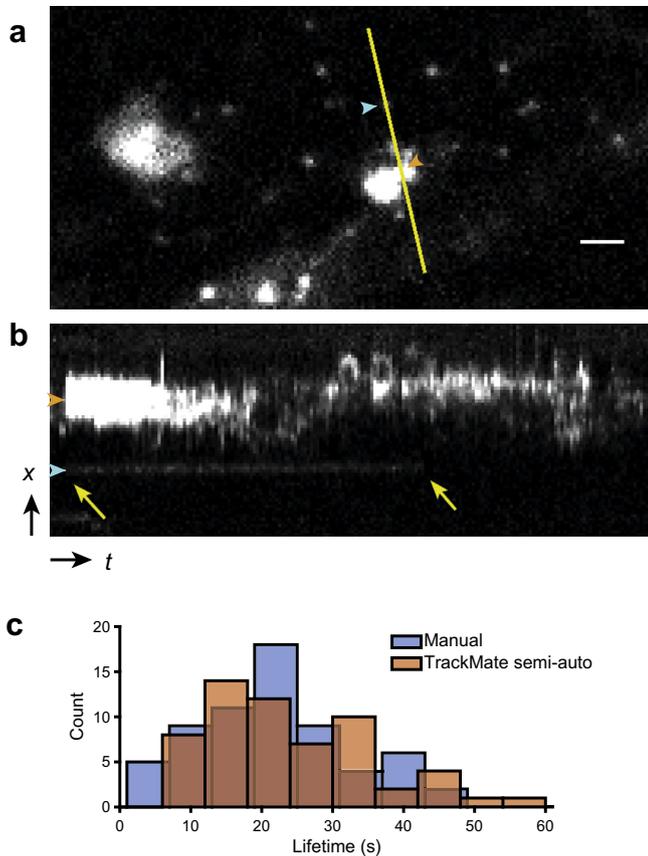
### 3.3. VAEM imaging of plasma membrane-associated Clathrin dynamics in Arabidopsis

Secretory and endocytic membrane trafficking are essential processes critical for growth and development of multicellular organisms including cell expansion and division, as well as responses to abiotic and biotic signals. Bidirectional trafficking of cargo materials between the discrete organelles of the secretory and endocytic pathways is primarily mediated by vesicular carriers. Clathrin, a vesicle coat protein complex, is responsible for cargo selection and formation of vesicles that traffic between the plasma membrane, endosome, trans-Golgi, and the vacuole [33,34]. Clathrin-coated vesicle (CCV) formation involves the temporal and highly ordered assembly of proteins following a poorly understood nucleation event. After this site-selecting initiation event, cargo begins to cluster together on the membrane through the action of various cargo receptors and adaptor complexes specific to each trafficking pathway (*i.e.* the adaptin 2, AP2, and TPLATE complexes in clathrin mediated endocytosis, CME, in plants [35]). While well characterized in yeast and mammalian systems, less is known of the molecular underpinnings of clathrin-dependent vesicle formation and targeting in plants.

Clathrin-mediated endocytosis (CME), cytoskeletal dynamics, and other events occurring at the cell cortex have been successfully imaged in mammalian and other systems using Total Internal Reflection Fluorescence Microscopy (TIRFM [36]). Likewise, TIRFM has been utilized to image molecular events at or adjacent to the cell cortex in plant cells [37]. However, plant cells are non-ideal samples for TIRFM analysis owing to the thickness of plant cell walls, which can exceed 500 nm, the theoretical limit of TIRFM illumination, and the non-uniform refractive index of cell walls due to variations in composition between cell types and stages of development. Additionally, plants do not readily adhere to glass surfaces, which, in combination with frequently non-uniform geometries of plant tissues, make the task of imaging molecular events at the cell cortex in single plant cells by the critical angle illumination necessary for TIRFM challenging. To accommodate these technical hurdles, Variable Angle Epifluorescence Microscopy (VAEM), which utilizes sub-critical laser angles to achieve oblique illumination of the plant cell cortex, was developed as an alternative application of TIRFM systems [22]. Essentially, the incident angle of the excitation beam utilized for VAEM is adjusted to generate an uneven ‘wedge’ of illumination radiating from the optical axis. The result is the maintenance of high S/N ratio at greater illumination depths (>1000 nm) than are achievable with TIRFM [22].

#### 3.3.1. Results

Here, we employ TrackMate to analyze VAEM data of fluorescently tagged clathrin foci at the cell cortex to compare the software’s particle tracking capabilities of manually chosen plasma membrane PM-associated clathrin foci with completely manually processed data. Representative images demonstrate CLC2-GFP foci labeling of PM foci and larger cortical Golgi structures (Fig. 5a and b) in Arabidopsis hypocotyl epidermal cells. The clathrin containing PM foci have a more uniform fluorescence and are diffraction limited in size compared to the larger CLC2-GFP labeled Golgi structures, displaying the characteristic fluorescence profile of sites of CME previously described [21]. Three independent hypocotyls were imaged and time-lapse images from each were analyzed manually and semi-automatically with TrackMate (Supplemental Movie 4). A histogram comparing the manual and TrackMate data displays the distribution of clathrin PM foci lifetimes observed by both methods (Fig. 5c). PM-associated clathrin foci lifetimes determined manually and via TrackMate range from 5–50 s and 7–55 s, respectively. The medians of the manual and TrackMate generated lifetimes are 20.5 s and 22.3 s, respectively.



**Fig. 5.** VAE M imaging of clathrin foci in *Arabidopsis* hypocotyl epidermal cells. a. Single image of a time-lapse movie following the lifetime of clathrin foci. The fluorescence signal shows clathrin foci along with larger, brighter Golgi structures. Yellow line: profile used for the kymograph in b. Scale bar: 2  $\mu\text{m}$ . Blue and orange arrowheads denote PM- and Golgi-associated clathrin, respectively. b. Kymograph extracted along the yellow line in a. The brightness has been adjusted to better show the clathrin focus, saturating the display of the large Golgi structure. Yellow arrows: start and end of the clathrin focus lifetime, determined manually. Blue and orange arrowheads denote PM- and Golgi-associated clathrin, respectively. c. Histograms of lifetimes for clathrin foci.

The average foci lifetimes for each individual hypocotyl did not vary significantly and the average foci lifetimes for manual and TrackMate data over all samples are  $21.7 \pm 10.5$  s (mean  $\pm$  standard

deviation,  $N = 64$ ) and  $24.2 \pm 11.6$  s ( $N = 59$ ), respectively. The Student's  $t$ -test and the Kolmogorov–Smirnov test gave both  $p$ -values greater than 0.1.

### 3.3.2. Discussion

Previously, the lifetimes of CLC2-GFP positive PM foci were determined by manual identification and quantitation of the fluorescence intensity of individual foci across a time-lapse series of images; a laborious and potentially biased process plagued by the difficult and somewhat arbitrary determination of where events begin and end. TrackMate helps in addressing these issues through its ability to follow a single focus throughout its whole lifetime from a single manual annotation. The slight difference between the manual and TrackMate average clathrin foci lifetimes is most likely due to TrackMate's capability to detect foci beyond what the human eye can discern via the quality threshold parameter. Despite this slight difference, the Student's  $t$ -test and the Kolmogorov–Smirnov test indicate that the manual and TrackMate average lifetime values and data distribution do not statistically differ from each other for PM associated Clathrin foci. This suggests that the software's tracking capabilities are compatible with manual tracking of clathrin dynamics at the PM. A number of laboratories studying clathrin and clathrin accessory proteins involved in plant CME have found the average lifetime of clathrin-associated PM foci to range from 17.7 s to 24 s, corroborating the average lifetime of clathrin foci at the PM found both manually and semi-automatically with TrackMate [21,38,39]. Inevitably, TrackMate's ability to objectively track foci is dependent upon parameters defined by the operator thereby introducing some level of bias into the results. Nevertheless, these parameters can be universally applied across all samples, eliminating unintentional precision errors and eye-constrained selection of track termination. Taken together, this data demonstrates the utility of TrackMate in alleviating a previously laborious analysis of dynamic events in microscopy data.

## 4. Discussion

TrackMate has been used in protein motility studies [40,41], molecular motor tracking [42], axonal transport [43], sperm cell tracking [44], Golgi bodies tracking [45] in plants, wound healing [46], bacterial biofilm micro-rheology [47], cell movement on stiffness-patterned substrate [48], cell tracking in zebrafish [49] and drosophila [50] embryos, infected cells tracking [51], colloid

**Table 1**  
User contributed modules of TrackMate v3.4.0.

Module name	Module type	Purpose	Author	Location
Linear tracker	Particle-linking	Linking transported particles by extrapolating their velocity	Ronny Szech	<a href="https://github.com/chicoronny/RonnyTrackMate">https://github.com/chicoronny/RonnyTrackMate</a>
Batch mode	Plugin	Runs TrackMate in batch, reading configuration from a file		
Binary detector	Spot detector	Detect spot using intensity thresholding and ImageJ particle analyzer		
CSV exporter	Generic action	Export tracks to CSV file		
SQLite exporter	Generic action	Export tracks to SQLite database		
Multi-channel intensity analyzer	Spot feature analyzer	Compute spot mean intensity in up to 10 channels	Benoit Lombardot	<a href="https://github.com/tinevez/TrackMate-extras">https://github.com/tinevez/TrackMate-extras</a>
Find maxima	Spot detector	Detection based on region growing, reimplementing ImageJ <i>Find maxima</i> command	Thorsten Wagner	<a href="http://imagej.net/Find_maxima_(TrackMate_module)">http://imagej.net/Find_maxima_(TrackMate_module)</a>
Trajectory classifier	Plugin	Classify 2D tracks into normal diffusion, subdiffusion, confined diffusion and directed/active motion by a random forest approach		<a href="https://github.com/thorstenwagner/ij-trajectory-classifier">https://github.com/thorstenwagner/ij-trajectory-classifier</a>
Close gaps	Generic action	Close gaps in tracks by creating spots in missing frame by linear interpolation of their coordinates	Robert Haase	Integrated into TrackMate v3.4.0

diffusion studies [52], carboxysome lineaging in cyanobacteria [53], magnetic bead aggregates tracking in arteries [54], diatom motility studies [55], and stretch measurements in biomedical materials [56]. It also enabled the biological studies presented here, investigating the presence of a cell-specific phototoxic effect on embryonic development, the impact of phototoxicity on the dynamics of NEMO dot assembly and the dynamics of clathrin-mediated endocytic events in plant cells.

This great breadth of applications highlights the vast applicability of TrackMate. Generality often has a toll on the ability to successfully handle very specific problems. As noted in [4], it is very likely that some tracking challenges will require the development of specific tracking tools. TrackMate positions itself not only as a tool for single-particle tracking, but also as a platform to facilitate and accelerate the development of such tools. Other developers can port their algorithm(s) to TrackMate as a module, reusing the facilities there that would otherwise consume countless and tedious time to develop (data model, visualization, etc.). The Scijava annotation mechanism used for plugin discovery ensures that new modules can be developed and distributed without requiring any interaction with the TrackMate developers, ensuring full autonomy and independence. As of today, we are aware of four groups of contributions made public, listed in Table 1.

## 5. Future directions

TrackMate was first developed as a lineage tool in *C. elegans* and largely extended to deal with other problems such as intra- and inter-cellular trafficking. As discussed above TrackMate is being widely used, but to remain most relevant, its development needs to continue, in order to deal with emerging data types, scales, modalities, and analysis workflows. Many of these problems can be readily solved by how TrackMate is developed and deployed. By being an open source Fiji plugin that uses modular libraries such as Bio-formats [57], Scijava [18], ImgLib2 [58], etc., it is relatively straightforward to adapt TrackMate to new modalities and use cases. However, there are some challenges that may necessitate augmenting TrackMate's capabilities. For example, it is interesting to not only look at varying temporal scales but also at varying spatial scales. As multiscale imaging methods are developed and deployed to look at process such as metastasis and cell growth, there will need to be corresponding modules that can track and analyze these processes over different spatial scales.

As well, TrackMate needs to continue to evolve in its interoperability and ability to handle heterogeneous data, including non-image data types. One strategy that has served TrackMate well that we will continue to leverage is to take advantage of existing software tools and also offer TrackMate as a flexible portable module. In future development, we hope to have TrackMate equally accessible not only as a Fiji plugin but as an Ops module [18]. The ImageJ Ops initiative aims at producing a unifying library for scientific image processing, and makes it widely accessible to any software framework, not just ImageJ. The Ops framework provides unified interfaces for basic image manipulations that can be called by any tool using this framework. This type of modularity has already shown considerable promise in the recent use of TrackMate in KNIME as a Scijava process [59]. By having TrackMate in KNIME, it is possible to create data pipelines that could integrate TrackMate with a number of other tools including statistical tools such as R. With the recently developed Ops framework being driven by the ImageJ2 and KNIME developer community, it should be possible to take TrackMate functionality even farther and to a wider audience. In this way, other programs such as databases and even other tracking programs could call TrackMate, much in the same way that programs invoke scripts and other algorithms.

Recent developments in light microscopy, such as light sheet microscopy [60], have resulted in massive datasets that requires large extensive annotations. These requirements will demand a new type of data model that can support this type of extension annotation in 2D and 3D. Our hope is that such a data model could be driven and utilized by a number of common tools, including TrackMate.

## Acknowledgements

The authors would like to extend their gratitude for financial support from the European commission FP7 ICT (project "MEMI", J.-Y.T, S.L.S.), the Laboratory for Optical and Computational Instrumentation at the UW-Madison and National Science Foundation award #1121998 (J.S., S.Y.B., K.W.E). N.P. was a Pasteur visiting student thanks to funds provisioned by the Stanford University. TrackMate development has been and is supported by the Fiji, ImageJ and KNIME communities. Their invaluable help particularly that of Fiji maintainer Curtis Rueden and KNIME image processing developer Christian Dietz is acknowledged here. The authors also thank Ellen T. Arena for critical reading of the manuscript. J.-Y.T thanks the members of the Imago pole of the Institut Pasteur and his family for support and patience.

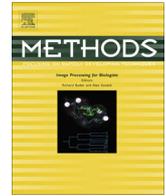
## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jymeth.2016.09.016>.

## References

- [1] E. Meijering, I. Smal, G. Danuser, Tracking in molecular bioimaging, *IEEE Signal Proc. Mag.* 23 (3) (2006) 46–53.
- [2] N. Chenouard, A. Dufour, J.C. Olivo-Marin, Tracking algorithms chase down pathogens, *Biotechnol. J.* 4 (6) (2009) 838–845.
- [3] Y. Kalaidzidis, Intracellular objects tracking, *Eur. J. Cell Biol.* 86 (9) (2007) 569–578.
- [4] N. Chenouard, I. Smal, F. de Chaumont, M. Maska, I.F. Sbalzarini, Y. Gong, J. Cardinale, C. Carthel, S. Coraluppi, M. Winter, A.R. Cohen, W.J. Godinez, K. Rohr, Y. Kalaidzidis, L. Liang, J. Duncan, H. Shen, Y. Xu, K.E. Magnusson, J. Jalden, H.M. Blau, P. Paul-Gilloteaux, P. Roudot, C. Kervrann, F. Waharte, J.Y. Tinevez, S.L. Shorte, J. Willemsse, K. Celler, G.P. van Wezel, H.W. Dan, Y.S. Tsai, C. Ortiz de Solorzano, J.C. Olivo-Marin, E. Meijering, Objective comparison of particle tracking methods, *Nat. Methods* 11 (3) (2014) 281–289.
- [5] J.I. Murray, Z. Bao, T.J. Boyle, R.H. Waterston, The lineaging of fluorescently-labeled *Caenorhabditis elegans* embryos with StarryNite and AceTree, *Nat. Protoc.* 1 (3) (2006) 1468–1476.
- [6] D. Sage, F.R. Neumann, F. Hediger, S.M. Gasser, M. Unser, Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics, *IEEE Trans. Image Process.* 14 (9) (2005) 1372–1383.
- [7] K. Jaqaman, D. Loeferle, M. Mettlen, H. Kuwata, S. Grinstein, S.L. Schmid, G. Danuser, Robust single-particle tracking in live-cell time-lapse sequences, *Nat. Methods* 5 (8) (2008) 695–702.
- [8] F. de Chaumont, S. Dallongeville, N. Chenouard, N. Herve, S. Pop, T. Provoost, V. Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, J.C. Olivo-Marin, Icy: an open bioimage informatics platform for extended reproducible research, *Nat. Methods* 9 (7) (2012) 690–696.
- [9] N. Chenouard, I. Bloch, J.C. Olivo-Marin, Multiple hypothesis tracking for cluttered biological image sequences, *IEEE Trans. Pattern Anal.* 35 (11) (2013) 2736–2750.
- [10] J.C. Crocker, D.G. Grier, Methods of digital video microscopy for colloidal studies, *J. Colloid Interface Sci.* 179 (1) (1996) 298–310.
- [11] A. Serge, N. Bertaux, H. Rigneault, D. Marguet, Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes, *Nat. Methods* 5 (8) (2008) 687–694.
- [12] I.F. Sbalzarini, P. Koumoutsakos, Feature point tracking and trajectory analysis for video imaging in cell biology, *J. Struct. Biol.* 151 (2) (2005) 182–195.
- [13] F.P. Cordelières, V. Petit, M. Kumasaka, O. Debeir, V. Letort, S.J. Gallagher, L. Larue, Automated cell tracking and analysis in phase-contrast videos (iTrack4U): development of Java software based on combined mean-shift processes, *PLoS ONE* 8 (11) (2013) e81266.
- [14] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (7) (2012) 676–682.

- [15] A.E. Carpenter, L. Kametsky, K.W. Eliceiri, A call for bioimaging software usability, *Nat. Methods* 9 (7) (2012) 666–670.
- [16] B. Schmid, J. Schindelin, A. Cardona, M. Longair, M. Heisenberg, A high-level 3D visualization API for Java and ImageJ, *BMC Bioinformatics* 11 (2010) 274.
- [17] R.E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng.* 82 (1) (1960) 35–45.
- [18] J. Schindelin, C.T. Rueden, M.C. Hiner, K.W. Eliceiri, The ImageJ ecosystem: An open platform for biomedical image analysis, *Mol. Reprod. Dev.* 82 (7–8) (2015) 518–529.
- [19] J.Y. Tinevez, J. Dragavon, L. Baba-Aissa, P. Roux, E. Perret, A. Canivet, V. Galy, S. Shorte, A quantitative method for measuring phototoxicity of a live cell imaging microscope, *Methods Enzymol.* 506 (2012) 291–309.
- [20] N. Tarantino, J.Y. Tinevez, E.F. Crowell, B. Boisson, R. Henriques, M. Mhlanga, F. Agou, A. Israel, E. Laplantine, TNF and IL-1 exhibit distinct ubiquitin requirements for inducing NEMO-IKK supramolecular structures, *J. Cell Biol.* 204 (2) (2014) 231–245.
- [21] C.A. Konopka, S.K. Backues, S.Y. Bednarek, Dynamics of arabidopsis dynamin-related protein 1C and a clathrin light chain at the plasma membrane, *Plant Cell* 20 (5) (2008) 1363–1380.
- [22] C.A. Konopka, S.Y. Bednarek, Variable-angle epifluorescence microscopy: a new way to look at protein dynamics in the plant cell cortex, *Plant J.* 53 (1) (2008) 186–196.
- [23] Z. Bao, Z. Zhao, T.J. Boyle, J.I. Murray, R.H. Waterston, Control of cell cycle timing during *C. elegans* embryogenesis, *Dev. Biol.* 318 (1) (2008) 65–72.
- [24] V. Praitis, E. Casey, D. Collar, J. Austin, Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*, *Genetics* 157 (3) (2001) 1217–1226.
- [25] Z. Bao, J.I. Murray, T. Boyle, S.L. Ooi, M.J. Sandel, R.H. Waterston, Automated cell lineage tracing in *Caenorhabditis elegans*, *Proc. Natl. Acad. Sci. U.S.A.* 103 (8) (2006) 2707–2712.
- [26] S.G. Megason, In toto imaging of embryogenesis with confocal time-lapse microscopy, *Methods Mol. Biol.* 546 (2009) 317–332.
- [27] R. Schnabel, H. Hutter, D. Moerman, H. Schnabel, Assessing normal embryogenesis in *Caenorhabditis elegans* using a 4D microscope: variability of development and regional specification, *Dev. Biol.* 184 (2) (1997) 234–265.
- [28] J.E. Sulston, H.R. Horvitz, Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*, *Dev. Biol.* 56 (1) (1977) 110–156.
- [29] J. Caron, C. Fallet, J.Y. Tinevez, L. Moisan, L.P. Braitbart, G.Y. Sirat, S.L. Shorte, Conical diffraction illumination opens the way for low phototoxicity super-resolution imaging, *Cell Adhes. Migr.* 8 (5) (2014) 430–439.
- [30] E.T. Kipreos, *C. elegans* cell cycles: invariance and stem cell divisions, *Nat. Rev. Mol. Cell Biol.* 6 (10) (2005) 766–776.
- [31] M.S. Hayden, S. Ghosh, NF- $\kappa$ B, the first quarter-century: remarkable progress and outstanding questions, *Genes Dev.* 26 (3) (2012) 203–234.
- [32] H. Qian, M.P. Sheetz, E.L. Elson, Single particle tracking. Analysis of diffusion and flow in two-dimensional systems, *Biophys. J.* 60 (4) (1991) 910–921.
- [33] X. Chen, N.G. Irani, J. Friml, Clathrin-mediated endocytosis: the gateway into plant cells, *Curr. Opin. Plant Biol.* 14 (6) (2011) 674–682.
- [34] G.A. Baisa, J.R. Mayers, S.Y. Bednarek, Budding and braking news about clathrin-mediated endocytosis, *Curr. Opin. Plant Biol.* 16 (6) (2013) 718–725.
- [35] Y. Zhang, S. Persson, J. Hirst, M.S. Robinson, D. van Damme, C. Sanchez-Rodriguez, Change your TPLATE, change your fate: plant CME and beyond, *Trends Plant Sci.* 20 (1) (2015) 41–48.
- [36] T. Kirchhausen, Imaging endocytic clathrin structures in living cells, *Trends Cell Biol.* 19 (11) (2009) 596–605.
- [37] G. Vizcay-Barrena, S.E.D. Webb, M.L. Martin-Fernandez, Z.A. Wilson, Subcellular and single-molecule imaging of plant fluorescent proteins using total internal reflection fluorescence microscopy (TIRFM), *J. Exp. Bot.* 62 (15) (2011) 5419–5428.
- [38] L. Bashline, S. Li, C.T. Anderson, L. Lei, Y. Gu, The endocytosis of cellulose synthase in arabidopsis is dependent on  $\mu$ 2, a clathrin-mediated endocytosis adaptin, *Plant Physiol.* 163 (1) (2013) 150–160.
- [39] A. Gadeyne, C. Sánchez-Rodríguez, S. Vanneste, S. Di Rubbo, H. Zaubner, K. Vanneste, J. Van Leene, N. De Winne, D. Eeckhout, G. Persiau, E. Van De Slijke, B. Cannoot, L. Vercrusse, Jonathan R. Mayers, M. Adamowski, U. Kania, M. Ehrlich, A. Schweighofer, T. Ketelaar, S. Maere, Sebastian Y. Bednarek, J. Friml, K. Gevaert, E. Witters, E. Russinova, S. Persson, G. De Jaeger, D. Van Damme, The TPLATE adaptor complex drives clathrin-mediated endocytosis in plants, *Cell* 156 (4) (2014) 691–704.
- [40] M.E. Tanenbaum, L.A. Gilbert, L.S. Qi, J.S. Weissman, R.D. Vale, A protein-tagging system for signal amplification in gene expression and fluorescence imaging, *Cell* 159 (3) (2014) 635–646.
- [41] M.F. Mossuto, S. Sannino, D. Mazza, C. Fagioli, M. Vitale, E.D. Yoboue, R. Sitia, T. Anelli, A dynamic study of protein secretion and aggregation in the secretory pathway, *PLoS ONE* 9 (10) (2014) e108496.
- [42] A. Reversat, M.I. Yuseff, D. Lankar, O. Malbec, D. Obino, M. Maurin, N.V. Penmarcha, A. Amoroso, L. Sengmanivong, G.G. Gundersen, I. Mellman, F. Darchen, C. Desnos, P. Pierobon, A.M. Lennon-Duménil, Polarity protein Par3 controls B-cell receptor dynamics and antigen extraction at the immune synapse, *Mol. Biol. Cell* 26 (7) (2015) 1273–1285.
- [43] S. Gluska, M. Chein, N. Rotem, A. Ionescu, E. Perlson, Tracking Quantum-Dot labeled neurotropic factors transport along primary neuronal axons in compartmental microfluidic chambers, *Methods Cell Biol.* 131 (2016) 365–387.
- [44] P. Denninger, A. Bleckmann, A. Lausser, F. Vogler, T. Ott, D.W. Ehrhardt, W.B. Frommer, S. Sprunck, T. Dresselhaus, G. Grossmann, Male-female communication triggers calcium signatures during fertilization in *Arabidopsis*, *Nat. Commun.* 5 (2014) 4645.
- [45] B. Sporsheim, A. Øverby, A.M. Bones, Allyl isothiocyanate inhibits actin-dependent intracellular transport in *Arabidopsis thaliana*, *Int. J. Mol. Sci.* 16 (12) (2015) 29134–29147.
- [46] M.S. Minns, G. Teicher, C.B. Rich, V. Trinkaus-Randall, Purinoreceptor P2X7 regulation of Ca(2+) mobilization and cytoskeletal rearrangement is required for corneal reepithelialization after injury, *Am. J. Pathol.* 186 (2) (2016) 285–296.
- [47] S.C. Chew, S.A. Rice, S. Kjelleberg, L. Yang, In situ mapping of the mechanical properties of biofilms by particle-tracking microrheology, *J. Vis. Exp.* 106 (2015) e53093.
- [48] K.A. Mosiewicz, L. Kolb, A.J. van der Vlies, M.P. Lutolf, Microscale patterning of hydrogel stiffness through light-triggered uncaging of thiols, *Biomater. Sci.* 2 (11) (2014) 1640–1651.
- [49] A. Pauli, M.L. Norris, E. Valen, G.L. Chew, J.A. Gagnon, S. Zimmerman, A. Mitchell, J. Ma, J. Dubrulle, D. Reyon, S.Q. Tsai, J.K. Joung, A. Saghatelian, A.F. Schier, Toddler: an embryonic signal that promotes cell movement via Apelin receptors, *Science* 343 (6172) (2014) 1248636.
- [50] K. Campbell, J. Casanova, A role for E-cadherin in ensuring cohesive migration of a heterogeneous population of non-epithelial cells, *Nat. Commun.* 6 (2015) 7998.
- [51] B. Spiesschaert, B. Goldenbogen, S. Taferner, M. Schade, M. Mahmoud, E. Klipp, N. Osterrieder, W. Azab, Role of gB and pUS3 in equine herpesvirus 1 transfer between peripheral blood mononuclear cells and endothelial cells: a dynamic in vitro model, *J. Virol.* 89 (23) (2015) 11899–11908.
- [52] P. Schattling, B. Thingholm, B. Städler, Enhanced diffusion of glucose-fueled Janus particles, *Chem. Mater.* 27 (21) (2015) 7412–7418.
- [53] A.H. Chen, A. Robinson-Mosher, D.F. Savage, P.A. Silver, J.K. Polka, The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo, *PLoS ONE* 8 (9) (2013) e76127.
- [54] R. Müller, O. Stranik, F. Schlenk, S. Werner, D. Malsch, D. Fischer, W. Fritzsche, Optical detection of nanoparticle agglomeration in a living system under the influence of a magnetic field, *J. Magn. Magn. Mater.* 380 (2015) 61–65.
- [55] K.G. Bondoc, J. Heuschele, J. Gillard, W. Vyverman, G. Pohnert, Selective silicate-directed motility in diatoms, *Nat. Commun.* 7 (2016) 10540.
- [56] C. Noble, N. Smulders, N.H. Green, R. Lewis, M.J. Carré, S.E. Franklin, S. MacNeil, Z.A. Taylor, Creating a model of diseased artery damage and failure from healthy porcine aorta, *J. Mech. Behav. Biomed. Mater.* 60 (2016) 378–393.
- [57] M. Linkert, C.T. Rueden, C. Allan, J.M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, C. Neves, D. Macdonald, A. Tarkowska, C. Sticco, E. Hill, M. Rossner, K. W. Eliceiri, J.R. Swedlow, Metadata matters: access to image data in the real world, *J. Cell Biol.* 189 (5) (2010) 777–782.
- [58] T. Pietzsch, S. Preibisch, P. Tomancak, S. Saalfeld, ImgLib2 – generic image processing in Java, *Bioinformatics* 28 (22) (2012) 3009–3011.
- [59] C. Dietz, M.R. Berthold, KNIME for open-source bioimage analysis: a tutorial, *Adv. Anat. Embryol. Cell Biol.* 219 (2016) 179–197.
- [60] N. Scherf, J. Huisken, The smart and gentle microscope, *Nat. Biotechnol.* 33 (8) (2015) 815–818.



## Ultrasound registration: A review



Chengqian Che<sup>a</sup>, Tejas Sudharshan Mathai<sup>b</sup>, John Galeotti<sup>a,b,\*</sup>

<sup>a</sup>Department of Biomedical Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA

<sup>b</sup>The Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA

### ARTICLE INFO

#### Article history:

Received 16 July 2016

Received in revised form 7 December 2016

Accepted 8 December 2016

Available online 11 December 2016

#### Keywords:

Registration

Ultrasound

IMAGE analysis algorithms and methods

Volume reconstruction

Image alignment

Matching

Warping

Transforms

Metrics

Optimization

Rigid

Non-rigid

Matrix

Elastic

Inelastic

2D

3D

### ABSTRACT

This article is a review of registration algorithms for use between ultrasound images (monomodal image-based ultrasound registration). Ultrasound is safe, inexpensive, and real-time, providing many advantages for clinical and scientific use on both humans and animals, but ultrasound images are also notoriously noisy and subject to several unique artifacts/distortions. This paper introduces the topic and unique aspects of ultrasound-to-ultrasound image registration, providing a broad introduction and summary of the literature and the field. Both theoretical and practical aspects are introduced. The first half of the paper is theoretical, organized according to the basic components of a registration framework, namely preprocessing, image-similarity metrics, optimizers, etc. It further subdivides these methods between those suitable for elastic (non-rigid) vs. inelastic (matrix) transforms. The second half of the paper is organized by anatomy and is practical in nature, presenting and discussing the complete published systems that have been validated for registration in specific anatomic regions.

© 2017 Elsevier Inc. All rights reserved.

### Contents

1. Introduction	129
2. Pre-processing	130
2.1. Noise removal	130
2.2. Multi-resolution and down-sampling	130
2.3. Image segmentation and feature detection	130
3. Registration metrics	131
3.1. Well established similarity metrics used in both rigid and non-rigid registration	131
3.2. Well established feature-based metrics	131
3.3. Similarity metrics used only in rigid registration	132
3.3.1. Novel	132
3.3.2. Adapted	132
3.4. Similarity metrics used only in non-rigid registration	132
3.4.1. Novel	132
3.4.2. Adapted	132
4. Registration optimizers	132

\* Corresponding author at: Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA.

E-mail addresses: [cche@andrew.cmu.edu](mailto:cche@andrew.cmu.edu) (C. Che), [tmathai@andrew.cmu.edu](mailto:tmathai@andrew.cmu.edu) (T.S. Mathai), [jgaleotti@cmu.edu](mailto:jgaleotti@cmu.edu) (J. Galeotti).

4.1.	Optimizers used only in rigid registration . . . . .	133
4.1.1.	Adapted . . . . .	133
4.2.	Optimizers used only in non-rigid registration . . . . .	133
4.2.1.	Novel . . . . .	133
4.2.2.	Adapted . . . . .	133
5.	Registration techniques and results by anatomy . . . . .	134
5.1.	Head . . . . .	134
5.2.	Neck . . . . .	136
5.3.	Breast . . . . .	136
5.4.	Heart . . . . .	136
5.5.	Liver . . . . .	137
5.6.	Kidney . . . . .	137
5.7.	Gall bladder . . . . .	137
5.8.	Bone . . . . .	138
5.9.	Prostate . . . . .	138
5.10.	Fetal imaging . . . . .	138
6.	Discussion, Shortcomings, and future opportunities . . . . .	139
6.1.	Diverse anatomy and ultrasound systems . . . . .	140
6.2.	Interactive nature of ultrasound acquisition . . . . .	140
6.2.1.	Physical probe tracking . . . . .	140
6.2.2.	Viewpoint specific . . . . .	140
6.2.3.	Interactive manipulation during ultrasound acquisition . . . . .	140
	Acknowledgements . . . . .	141
	References . . . . .	141

## 1. Introduction

This article is a review of registration algorithms for use between ultrasound images (monomodal registration). We review the algorithms that have been published during the years 1998–2016 to address rigid and non-rigid registration of ultrasound-to-ultrasound only. Ultrasound imaging is a non-invasive, real-time, and lower cost alternative to other medical imaging modalities, providing many advantages for clinical and scientific use on both humans and animals. Ultrasound does not expose the patient to ionizing radiation, provides continuous real-time video imaging, and is generally painless. Ultrasound imaging has been used in various fields of medicine such as imaging the human brain, heart, liver etc. [1–3], and ultrasound is increasingly being used in animal studies, including, e.g., porcine brain [4], porcine carotid artery [5], and mice [6]. In many cases, multiple ultrasound images or volumes are acquired by either free-hand scans of the anatomical region of interest, or following a specific pattern of scanning through the rotation and translation of the ultrasound transducer. Regardless of scan pattern, the transducer is usually pushed into the region of interest to obtain better signal-to-noise ratio (SNR), deforming tissue in the process.

As ultrasound imaging produces a continual stream of images (or volumes), it is often necessary to register these images together, either for clinical evaluation of the entire organ, or to be used intra-operatively [7]. Unfortunately, ultrasound images are also notoriously noisy and subject to several unique artifacts/distortions. *Unlike multimodal registration (which typically benefits from higher-quality CT or MRI images against which to register the ultrasound images), monomodal ultrasound image-based registration algorithms bring two ultrasound images into alignment with one another based solely on the ultrasound data.*

Monomodal ultrasound registration, though challenging, is important for a growing body of work. For longitudinal studies, ultrasound imaging avoids accumulating MRI cost, CT radiation, and exposure to potentially harmful MRI-or-CT contrast agents. Diagnosis may require searching for meaningful differences between ultrasound images that were acquired months or years apart. Population studies compare ultrasound images between large numbers of subjects, often to disambiguate normal vs.

pathological variation. Finally, it is desirable to combine multiple 2D ultrasound slices in a single 3D dataset, either when 3D ultrasound transducers are not an option, or when 2D transducers produce higher quality images.

Sometimes, the 3D spatial location and orientation of the ultrasound probe is also known through the use of ultrasound probe trackers [7,8]. However, probe tracking is also subject to error, especially for lower-accuracy electromagnetic (EM) trackers. Furthermore, probe tracking results in structural misalignment due to soft-tissue deformation, which is not directly measurable by probe tracking. For both reasons, image-based registration is still usually required to refine the tracker's initial registration estimate.

A registration process should result in the transform that best (however defined) aligns an image that needs to be registered (generally called the moving image) to a reference or fixed image. Registration transforms are typically grouped into two major classes: Rigid (or matrix) transformations and non-rigid (elastic) transformations. Within these classes, each transformation type has certain characteristic degrees of freedom (DOF). Degrees of freedom represent the number of independent parameters that are necessary to specify a transform, such as rotations and translations. Matrix transforms are generally of four types: strictly rigid, similarity, affine, and perspective. All these matrix transforms can perform rotations and translations. Similarity transforms can also perform scaling, and affine transforms can perform all of this plus shearing. Perspective transforms (which do not require parallel lines to remain parallel) are the most free-form matrix transform. These degrees of freedom for these matrix transforms, listed below in Table 1, constrain the warp of the moving image onto the fixed image. Closely related to matrix transforms are Quaternions, which are also a strictly rigid transform, but use a special representation so-as to avoid singularities in 3D rotation. Non-rigid transforms can have many more DOFs (e.g., vector fields

**Table 1**  
Degrees of freedom (DOF) for matrix transforms.

Dimensionality	Strictly rigid	Similarity	Affine	Perspective
2D	3	4	6	8
3D	6	7	12	15

can have thousands), and can thus model elastic deformations of anatomical objects. With so many DOFs, non-rigid transforms typically require regularization in order to maintain spatial order and regularity. Other terms for non-rigid include deformable and elastic. There are two common situations in ultrasound image registration for which a rigid-body or matrix transformation would be used. The first situation is registration of non-deformable or inelastic anatomical structures such as bone. The second situation is as a simple (and quick) approximation when changes are small, often for deeper organs such as liver or else to initialize a subsequent deformable registration. On the other hand, non-rigid transforms, such as a B-spline or Thin Plate Spline deformation, are frequently used in ultrasound image registration due to the deformable nature of soft tissue such as organs, vessels, or the brain.

The rest of the paper is organized as follows: in the pre-processing section, we elaborate the techniques employed for noise reduction, as well as image segmentation and feature detection approaches to guide rigid and non-rigid ultrasound registration. Next we describe the common registration metrics used in literature, and then we explain the different registration optimizers preferred by researchers over the years for rigid and non-rigid registration. We then detail complete registration methods specifically designed for different anatomical regions of the human body, and finally we summarize with a discussion of the methods used. This review paper is not an exhaustive review of the various registration algorithms proposed over the years that could potentially be applied to ultrasound. However, it serves to inform the reader of the different approaches that have been popularly used and validated for rigid and non-rigid ultrasound-to-ultrasound registration.

## 2. Pre-processing

Ultrasound images typically have substantial noise, shadowing, etc., making registration and analysis more difficult than for MRI or CT. Multiple image pre-processing algorithms have been applied on ultrasound images to aid registration. One common pre-processing technique is to apply noise removal filters (including interpolation techniques). Some registration algorithms incorporate image segmentation and/or feature extraction to guide the registration. These registration implementations are executed either on noise reduced ultrasound images or directly on raw ultrasound images. Also, if images' intensities are scaled differently, the image intensities can be adjusted in pre-processing, or else an additional intensity scaling parameter may be included in the registration model [9]. Below are methods used in the literature specifically for ultrasound-to-ultrasound registration. There are many other preprocessing methods that may also be useful, such as histogram equalization (intensity normalization) or bi-linear filtering (edge preserving smoothing).

### 2.1. Noise removal

One of the most common types of noise in ultrasound images is speckle noise, which is caused by the interference of coherent ultrasound waves scattered by tissue within each resolution cell (e.g., interference inside a single pixel) [10]. Although speckle can be used as a feature (speckle-based registration is discussed in Section 2.3), speckle degrades the quality of ultrasound images and can lead to poor registration if neither suppressed nor explicitly utilized. Researchers have applied different noise-removal techniques before registration. One technique is to apply Gaussian filtering [3–10] to reduce the speckle noise in the image by blurring. More significantly, certain characteristics of an image, such as edges and ridges, can be preserved and enhanced by

applying a bank of filters including Gaussian smoothing filter, the first derivative filter (gradient operator), the second derivative filter, the Laplacian filter, and threshold operators [11]. Oscillatory functions were used to reduce noise in [17]. Another common filter to reduce noise, especially speckle noise, is median filtering [18–22]. A median filter works as a non-linear low-pass filter, assigning to each pixel the median of its local (2D or 3D) neighboring values. Median filtering can potentially preserve image details while suppressing noise better than blurring, since the median is robust against outliers [20]. In [20], the S-Mean filter was proposed to more effectively remove speckle noise. This filter performs anisotropic diffusion to reduce speckle (SRAD) followed by a median filter. SRAD itself was introduced in [23] to preserve and enhance prominent edges. When used with anisotropic diffusion, the median filter was shown to reduce speckle noise with minimal edge degradation.

### 2.2. Multi-resolution and down-sampling

Multi-resolution (e.g. coarse-to-fine) approaches are common in image optimization tasks. Compared to “regular” (single-resolution) ultrasound registration algorithms, using a multi-resolution strategy usually has a higher convergence radius, is more robust to poor local optima, and progresses faster. Image pyramids are used to down sample both the fixed and moving images to a variety of scales, registering first on coarse, low-resolution images and then progressively refining that registration at increasing resolutions. In [12,15,24–27], authors used multi-resolution strategies with specific down-sampling ratios to get better registration results.

### 2.3. Image segmentation and feature detection

Many registration algorithms incorporate image segmentation and feature extraction methods in order to guide the registration. Segmentation approaches identify “foreground” regions of interest (ROI) in the ultrasound images, optimizing the registration primarily for those specific regions while the “background” regions are either ignored or else registered in a somewhat under-constrained fashion. Automatic feature detection utilizes points of interest in the image, along with their local neighborhoods, as key points for registration.

One example of registration based on coarse segmentation used structural information about the interface between bone and soft tissue. Their algorithm starts by segmenting the image using Otsu's thresholding method [28] to obtain a good initial, approximate segmentation of the separation between the echogenic zone and the shadow zone. Next, a Sobel filter is applied to detect the horizontal bone interface, followed by averaging to remove noise. Then, a fusion of the output of Otsu's method and the detected horizontal bone interface is performed. This fused result is dilated to produce the final segmented regions of interest with structural information [22]. An interactive live-wire segmentation algorithm was featured in [17] where the user selected seed points on breast and kidney ultrasound images in order to extract structural outlines of specific objects in the images. The cost functions for live-wire algorithm included the gradient magnitude, gradient direction, Canny edge features [29] and Laplacian zero crossing features. In [30], the local phase information of an image was extracted first, and then it was segmented to obtain the boundaries of objects with smooth borders in the images. The approach published in [5] segmented the lumen and media-adventitial boundaries using the B-snakes algorithm published in [31]. In [32], a fully automated segmentation algorithm based on a Generalized Hough Transform (GHT) and subsequent model adaptation with increasing degrees of freedom was used to segment volumes

of the heart of babies, and obtain a deformable model of the aortic root in a pediatric patient. In [33], both the Endocardium at end-diastole and the Epicardium were manually segmented.

Feature extraction can also be used to improve the accuracy of ultrasound image/volume registration. Researchers extract edges and texture features in ultrasound volumes to improve the smoothness of their metric functions, which leads to better optimization and results [34]. They model speckle noise as an irregular and complex texture pattern in ultrasound images, leading them to employ a Gabor filter bank to analyze textures. Their corresponding metric function is more stable at large scales, and useful information such as edges are well preserved [4]. In [17,35–37], SIFT features [38] were detected and extracted from the ultrasound images, and then utilized for matching of the ultrasound images. In [39,40], corner features were detected based on the principal curvature values of the Hessian matrices obtained for every pixel in the ultrasound image.

One special feature used for ultrasound registration is speckle noise. Speckle results from random back-scattering in a resolution cell of the ultrasound beam, and its intensity distribution is proven to be non-Gaussian. A fully formed speckle is known to have a Rayleigh distribution in the envelope detected image and Fish-Tippet distribution in the log-compressed image [41]. Speckles as special features can be extracted and used for ultrasound registration, especially when misalignment is either minimal or else almost entirely translation. [41–44]. For instance, in [41], a feature detector is employed based on statistical theoretical distributions of fully formed speckle in an ultrasound image to generate edge maps. They demonstrated their method to accurately register ultrasound images with speckled data and to be more robust to noise than standard gradient-based methods.

### 3. Registration metrics

A registration *metric* quantifies the similarity between two corresponding images based on a postulated transform between them. The metric, also known as a *similarity measure*, is used by the registration optimizer to solve for a transform (matrix and/or displacement field) that puts the two images into alignment. A robust metric will reach either its largest or smallest possible value when the two images are in alignment, and would ideally monotonically fall away for non-perfect alignments. Generally, registration metrics can be used to establish similarity between 2D images, 3D volumes, or even 4D/5D hyper-volumes. Pixel (or voxel)-based metrics have been developed and implemented widely in ultrasound registration. Pixel-based metrics compare the images' overlapping pixel values [13], as opposed to other metrics which compare non-pixel values (such as extracted features). Over the past few decades, a great number of similarity metrics have been proposed and developed in the medical imaging and computer vision communities. There is not yet a solid theoretical basis for choosing one similarity metric over another, and any given metric's performance depends significantly on other registration factors, including (1) the optimizer and its parameters, (2) the anatomical structures being imaged, (3) the dimensionality of the registered object, (4) the nature of the transformation (rigid/Quaternion, affine, non-rigid flow, non-rigid finite elements, etc.). Table 2 lists a variety of similarity metrics that have been used for ultrasound registration.

#### 3.1. Well established similarity metrics used in both rigid and non-rigid registration

Mutual Information (MI) is a popular image similarity metric for both rigid and non-rigid medical image registration. Mutual

**Table 2**  
Similarity metrics used for ultrasound registration.

Registration type	Similarity metric
Standard Pixel-Based Metrics, Well Validated for both Rigid and Non-Rigid Transforms:	Mutual Information (MI), Normalized Cross Correlation (NCC), Correlation Coefficient (CC), Sum of Squared Differences (SSD)
Standard Feature-Based Metrics for both Rigid and Non-Rigid Transforms:	Manual Key Points, Automatic SIFT Feature Detection, Speckle Utilization, Application of pixel-based methods locally around key points
Unique Metrics, Rigid Validation Only	Novel Metrics: Hellinger Distance + statistics-based Fuzzy Local Binary Patterns (FLBP), N-dimensional Mutual Information Matrix Correlation Ratio (CR), Sum of Absolute Differences (SAD)
Unique Metrics, Non-Rigid Validation Only	Adapted Metrics: Novel Metrics: SSD of pixel intensities + weighted SSD of local phase information, similarity metric $Sim(s_1, s_2)$ based on comparing attribute vectors
	Adapted Metrics: SIFT matching, sample variance

information was initially introduced as a similarity metric by [45] and [46]. This metric seeks a transform that aligns two (ultrasound) images or volumes by maximizing their mutual information. The metric measures how much information one variable (image or volume) contains about the other. As [45] proposed, both the joint entropy and the individual entropies are used to obtain the mutual information. MI will be optimal when the individual entropies are maximized while joint entropies are minimized. MI is robust to outliers, and it is efficient to use in optimization, making MI an excellent metric [19,21,24,46,47–52].

Normalized Cross-Correlation (NCC) is another widely used similarity metric. NCC calculates the correlation between two functions, and it is considered to work the best with mono-modality registration between two images acquired with the same characteristic curves (e.g., same gamma curve). The NCC metric has been incorporated in both rigid [22] and non-rigid registration algorithms [53,27] with good results.

Correlation Coefficient (CC) is another prevalent similarity measure in literature. CC measures the linear correlation between two variables, outputting a degree of correlation between 0 and 1. It is simpler than NCC, but can be less robust to variations in imaging parameters (slightly different characteristic curves, etc.). CC can still perform well when used as a similarity metric in ultrasound registration with rigid transformation [13] and non-rigid transformation [54].

Sum of Squared Differences (SSD) is perhaps the simplest standard similarity metric in image registration. This metric calculates the sum of squared differences of pixels' or voxels' intensity from both reference and moving images or volumes. As this metric requires both reference and moving images/volumes to have the same intensity range, SSD is best-suited for mono-modality ultrasound registration. SSD for rigid [55], and non-rigid [5,25,33,56–63] registration has been discussed in recent papers.

#### 3.2. Well established feature-based metrics

Feature Key Points were extracted using a wide variety of methods in [17,35–37,39,40], and the local neighborhoods around these keypoints were used in order to match them together prior to registration. The combination of feature detection and localized pixel intensity matching was a key component for the registration algorithms used in these papers. Identification of other features such as SIFT and Speckle was previously discussed in the preprocessing section. Once features have been identified, they can be mapped across images to create various metrics.

### 3.3. Similarity metrics used only in rigid registration

#### 3.3.1. Novel

A novel hybrid procedure for rigid registration was proposed in [64]. The novel similarity metric is based on Hellinger distance between the distributions in images on the global scale, and a statistics-based extension of Fuzzy Local Binary Patterns (FLBP) on a local scale. Many methods are proposed to measure the similarity between LBPs, and histogram intersection is considered to be the most common one. Since histogram intersection may yield local minima, the Hellinger distance is preferred, which measures the statistical similarity of two distributions. Statistical similarity is globally precise but is locally imprecise. In contrast, the statistics-based FLBP is globally imprecise and locally precise. Thus, the combination of both components is desirable to achieve a large convergence radius (global) and a precise final result (local).

Instead of using regular mutual information as similarity metric, [18] introduced a high-dimensional mutual information matrix obtained by calculating its eigenvalues. During registration, when two images are geometrically aligned, N-dimensional MI reaches its maximum 1. This non-negative metric can be extended to higher dimensions so that multiple images can be registered. It was shown to register multiple ultrasound images through simulation.

#### 3.3.2. Adapted

Correlation Ratio (CR) is another metric derived from Cross correlation. CR was proposed in [13], and measures the functional dependence between two variables. CR has been shown to be suitable for mono-modality ultrasound rigid registration [12,34].

Another similarity metric that has been used in rigid registration is the Sum of Absolute Differences (SAD). SAD is similar to SSD, but SAD does not quadratically emphasize the pixels with the largest intensity differences. SAD was utilized in [12].

### 3.4. Similarity metrics used only in non-rigid registration

#### 3.4.1. Novel

A novel similarity measure was proposed for deformable registration of ultrasound images in [65]. It is a modification of the HAMMER algorithm, which was originally proposed in [66] for elastic registration of brain MR images. This similarity metric uses an attribute vector, consisting of geometric-moment invariants that are defined on each voxel in a 3D image. The similarity measure of two voxels is obtained by comparing their attribute vector and it is defined as:

$$Sim(s_1, s_2) = \prod_i (1 - |av_i(s_1) - av_i(s_2)|)^{w_i}$$

where  $av_i$  is the  $i$  th element of the attribute vector.  $Sim(s_1, s_2)$  is 1 for similar voxels and zero for dissimilar voxels [9].

The similarity metric was elaborated in [30] as a *compatibility coefficient* between two feature points present in the local phase of two ultrasound images:

$$r_{p_i, q_i} = \alpha \cdot \beta \cdot \gamma$$

Here,  $p_i$  and  $q_i$  are the set of feature points extracted from the local phase information of each of the two ultrasound images. They also defined  $N_a^{p_i}$  and  $N_b^{q_i}$  with  $a = 1, 2 \dots A$ ;  $b = 1, 2 \dots B$  to be the points adjacent to the detected feature points. Thus,  $\alpha(p_i, N_a^{p_i}; q_i, N_b^{q_i})$  is the disparity in Euclidean distance between feature points  $p$  and  $q_i$ , while  $\beta(p_i, N_a^{p_i}; q_i, N_b^{q_i})$  is the disparity in the angle between the feature points.  $\gamma(p_i, N_a^{p_i})$  is the spatial smoothness measured by the distance between  $p_i$  and  $N_a^{p_i}$ .

The similarity metric was defined in [6] as the sum of two parts: the SSD between the pixel intensities of the source and target image, and the weighted SSD of the local phase information of the source and target image. They worked on in vivo cardiac ultrasound images.

#### 3.4.2. Adapted

In [67], the voxel intensity-based similarity metric proposed in [68,69] was used. Here the reference volume is not selected, so the similarity measure is the sample variance of a population, which represents the difference from the current mean intensity  $\bar{x}$  for each voxel  $x$  in volume. The similarity metric proposed in [65] was adopted by [8,9,50,55].

## 4. Registration optimizers

An optimizer plays an essential role in an image registration framework. The goal of an optimizer is to search for the transformation that produces the best alignment of a moving image or volume with a reference image or volume. This is usually done through the maximization of a similarity metric (metrics are detailed in Section 3). *With ultrasound, a similarity metric will typically not be robust enough to determine the best transformation between the reference and moving images or volumes.* This is because speckle noise, along soft tissue deformation, affects the ultrasound data, and therefore contributes to the degradation in the similarity between two corresponding pixels or voxels. To mitigate these effects, a cost function is normally used in registration of ultrasound volumes. *The cost function will typically consist of a data term that represents the similarity measurement between two voxels, and a regularization term that penalizes unlikely deformations (especially in non-rigid registration).* An optimizer will aim to reach its optimal goal of correctly aligning and registering two images or volumes with minimal deformation and maximum similarity. Two desirable properties of an optimizer are its robustness and short convergence time. Therefore, a good ultrasound registration framework will incorporate the best optimizer and metric for registration of images or volumes. Mathematically, the cost function  $J$  to be minimized is defined as:

$$J(u) = D(u) + \alpha S[u]$$

where  $D$  is the similarity measurement function,  $S$  is the smoothing term that penalizes unlikely deformations, and  $u$  is the deformation field or other transform applied to the moving image [61]. The influence of the regularizer depends on the value of the scalar  $\alpha$ .

Several optimization algorithms have been developed and validated for rigid and non-rigid ultrasound registration. We discuss the approaches proposed in the literature for rigid and non-rigid body transformations in the following subsections, which are summarized in Table 3.

**Table 3**  
Optimizers used for ultrasound registration.

Registration type		Optimizer
Rigid Only	Adapted	Least-Squares optimization using Horn's Quaternion-based method [71], Nelder-Mead Simplex method [72], Mean-Shift optimization combined with Powell's direction set method [34]
	Novel	Alpha-Expansion technique [73], Gradient Descent [63], Variational minimization [6,51]
Non-Rigid Only	Adapted	Nelder-Mead Simplex method [20,34–36,38,62], Gradient Descent [37,64,73–75], Conjugate Gradient Descent [54], Least-Squares minimization [26], Levenberg-Marquardt optimization [75], Broyden-Fletcher-Goldfarb-Shannon (LBFGS) optimizer [76]

#### 4.1. Optimizers used only in rigid registration

In a rigid registration, similarity measurement can be considered as a 6 dimensional function due to the 6 degrees-of-freedom transformation. Since it is not possible to search through the entire parameter space exhaustively, initial parameter estimation is essential for the efficiency and accuracy of the optimization process.

##### 4.1.1. Adapted

The most widely used optimization algorithm for rigid only ultrasound image registration is the simplex method of Nelder and Mead [19–22,55,72,77,78] due to its robustness and computational efficiency. Inspired by the concept of a simplex, the Nelder and Mead algorithm is commonly used to find the minimum or maximum of an objective function in a multidimensional space. Before determining the initial simplex, normalization is needed for both translations and rotations so that the unit step in parameter space is approximately the same as the displacement in the spatial domain. In order to avoid finding local minimums and maximums, the size of initial simplex should be greater than the unit dimension. To stop the optimization process, two conditions are considered. The first one is when the size of the simplex is smaller than a unit hypercube in the parameter space. The second one is when similarity measurement meets a pre-defined value [47].

In [34], a mean-shift based optimization algorithm was used with the Powell's direction set method [79]. There are two advantages to using the mean-shift algorithm: (1) the local fluctuation is eliminated on the surface of similarity function effectively, and (2) the optimization process can be robust and accurate using a multi-resolution approach. In a high dimensional space, the computational load will be reduced by using Powell's direction set method [34].

Another optimizer that has been validated with rigid body ultrasound registration is the least squares method [80]. Yip et al. [81] used Random Sample Consensus [82] (RANSAC) to reject outliers, and identify the best transformation for registration. The least squares optimization was performed using Horn's quaternion-based method [71].

#### 4.2. Optimizers used only in non-rigid registration

##### 4.2.1. Novel

In [25], the registration energy cost function was modeled using Markov Random Fields (MRF), and optimized using a parallelized alpha-expansion technique [73]. In [30], a fuzzy correspondence matrix was used in addition to the compatibility coefficient (see Section 3.3.1) in order to estimate the transformation that maps points in the moving ultrasound image to the fixed ultrasound image. The transformation was defined using the Thin Plate Spline (TPS) model. In [63], the optimizer of choice was a gradient descent optimizer with an automatic step-size update. This paper proposed a unique way of estimating the global spatio-temporal deformation field for a sequence of images using this optimization method and a B-spline parametric model. In [61], the cost function contained a diffusion regularizer, and a variational minimization approach was utilized in order to register the volumes together. Here, the cost function was minimized by applying calculus of variations to obtain a non-linear partial differential equation, which was then solved using the modified fix-point iteration with incremental updates being regularized [83]. A variational framework was also used in [6] to solve the problem of non-rigid ultrasound registration. Here, the cost function consisted of a data term and a smoothness term, which could be minimized for gradient descent using the Euler-Lagrange method. The Euler-Lagrange equations that

were derived in their paper were solved using an alternating minimization approach.

##### 4.2.2. Adapted

In [48], first semi-automated approach was proposed (the authors did not know of any previously published semi- or fully automated algorithm) for non-rigid registration of ultrasound images of the breast. The optimizer used in this paper was the Nelder simplex algorithm [72]. The user clicked 3 control points in the moving image, and the optimizer tried to maximize the MI between the fixed and the moving images. The three control points were used to define a standard rigid transform (rotate-translate), and the rigid transform defined an additional control point. The four control points were then used to estimate a full affine transform. Finally, the affine transform defined more points, which could in turn be used to estimate repeated TPS warps.

In [24], a sub-volume based volumetric registration (SURE) algorithm was proposed, which divided the three-dimensional volume into subvolumes. They then proceeded to compute the similarity of each subvolume to the target volume within a search window, while discarding non-matching subvolumes. Finally, they computed the translation only vector to target volume, and then used the TPS model to estimate the deformation field between original position of subvolumes and new position of subvolumes. The optimizer they used in this paper was the Nelder-Mead simplex method [72]. In [47], a modified version of the Nelder-Mead simplex method proposed in [21] was utilized.

In [50], the MIAMI-FUSE registration software [84] was utilized for non-rigid ultrasound volume registration of the breast. Initially, affine registration was done by manually selecting control points. After initial alignment, at least one additional control point was needed for elastic registration using TPS. The location of other grayscale voxels in the moving volume was interpolated using TPS, and the three-dimensional volumes were registered. Nelder-Mead simplex method [72] was used for optimization of the cost function.

In [54], a Bayesian regularization framework for non-linear registration was proposed with the minimization achieved through conjugate gradient descent. The deformation field was generated by fitting a cubic tensor product B-spline approximating mesh.

A non-rigid registration algorithm was proposed in [49] that operated by optimizing a cost function made up of global and local motion models. The global motion model described the motion of the brain using an affine transform. The local motion model was based on the B-spline free form deformation (FFD). They used a multi-resolution approach in order to estimate the FFD from coarse-to-fine levels. They used the gradient descent optimizer proposed in [74].

In [85], an approach was put forth which was an extension of the demons algorithm. They added an extra force to the optical flow equation of the demons algorithm called the inertia force. By adding this extra term, they were able to achieve better non-rigid registration results over the traditional demons algorithm. The optimizer used here was a second order gradient descent on the SSD criterion.

A near real-time algorithm (RESOUND) was put forth in [27] that incorporated the minimization a cost function comprised of a similarity term based on NCC, and a smoothness term. The deformation field was estimated using cubic B-splines. The optimization was done by taking the analytic derivative of NCC, and using a stochastic gradient descent algorithm as in [86,87]. The algorithm was implemented over a three level multi-resolution framework.

An algorithm was implemented in [67] that registered an entire 4D (3D + time) sequence of liver ultrasound volumes in a group-wise fashion, and avoided bias towards a specifically chosen reference time point. The algorithm utilized the 4D FFD B-spline model

in order to estimate the deformation field. The optimizer that was used in this paper was the stochastic gradient descent algorithm [86,87], and the parameters for this algorithm were from the elastix toolbox [88]. The algorithm was executed using a multi-resolution approach.

A multi-resolution approach was proposed in [4] that had an embedded multi-grid registration framework. This meant that they divided a volume into partitions at each level. At each of the various scales, the image was divided into specific sub blocks within the same grid, and the deformation field was estimated for that block in the grid. Then, for a cube at each grid level, a 12-parametric increment field was estimated. The deformation field at the grid level was then used to initialize the deformation field over the next grid level. The optimizer used here was a multi-grid reweighted least squares minimization algorithm employing an iterative Gauss-Seidel scheme [26].

In [5], an algorithm for Intra-Vascular Ultrasound (IVUS) image registration was implemented by transforming the images from image coordinates into polar coordinates. Here, the registration energy functional to be minimized comprised the sum of squared differences of intensities, and the sum of squared differences in radial gradient in each pixel of the region of interest (ROI). The minimization was done based on the Levenberg-Marquardt optimization [75]. Bicubic B-splines were used to represent the deformation field.

In [52], a non-rigid registration algorithm was proposed which incorporated a twisting and bending model. They used the Powell optimizer [77,79] to estimate the six parameters of the rigid transformation along with the six parameters of the non-rigid transform. This optimizer did not require the calculation of gradient parameters in order to perform the optimization, and hence it was a suitable choice in this paper.

In [17], two methods were utilized to extract point pair correspondences describing the shape of objects in ultrasound images. First, they employed a global shape extraction scheme through segmentation to generate point pair correspondences between the fixed and moving images. Then, they used SIFT feature detection and matching to obtain salient feature points that defined the structure of the object of interest in the image. They used the point pair correspondences generated from the two aforementioned steps in a Bayesian framework, where the displacement field was estimated using MAP framework. The fast optimization algorithm proposed in [89] to obtain the velocity field by simple scale-space convolution was used for registering the images together.

Similarly, the method proposed in [35] utilized SIFT-based feature detection and matching at a single scale to register volumes. The volumes were acquired with small displacements, which allowed them to adopt a Register-To-Global strategy where features detected in newly acquired volumes were registered to a combination of features from all previous volumes encountered in the sequence. Their group-wise registration scheme was faster than [36,37], while maintaining accuracy.

The algorithm in [62] used the diagonal (d), vertical (v), and horizontal (h) components of the discrete Meyer wavelet transform to generate energy maps. They applied wavelet decomposition through a multi-resolution approach, minimized an energy functional with SSD as similarity metric and diffusion regularizer. They used the calculus of variations to minimize the energy functional leading to the Euler-Lagrange equation. This equation was approximated as a linear system by a finite difference method. The fixed point iteration method was used to find the approximate solution of this linear system.

Inter-session registration of three-dimensional *trans*-rectal ultrasound (TRUS) was achieved in [51]. They utilized two different non-rigid algorithms for comparison of performance. The first was surface-based registration using an initial ICP-based rigid

registration of three-dimensional TRUS volumes followed by non-rigid registration using TPS [90]. The second was image-based registration with a block matching approach [91] with MI as the similarity metric, followed by non-rigid registration using cubic B-splines. The optimizer that was used here was the Broyden-Fletcher-Goldfarb-Shannon (LBFGSB) optimizer [76].

A non-rigid registration algorithm was proposed in [60] to track the mitral valve annulus in three-dimensional ultrasound volumes. The optimizer that was used in this algorithm was the LBFGSB optimizer. The deformation field was estimated using three-dimensional third order B-splines. A multi-resolution approach was used.

An algorithm was proposed in [32] that registered volumes through the use of a model-based segmentation. They created a model of the aortic root by manually segmenting three-dimensional volumes of pediatric patients. Next, they acquired new 4D echo images of pediatric hearts, from which two three-dimensional volumes were extracted, and set as the reference and target respectively. These volumes were segmented using the deformable aortic root model that was generated. Next, 3D meshes were generated from the segmentations of the three-dimensional volumes. The three-dimensional meshes were then registered using TPS, thereby registering the extracted three-dimensional echo volumes.

A two-step approach was developed in [53] for elastic registration of IVUS frames in a sequence. First, rigid registration was performed to align coronary artery IVUS ultrasound B-scans. Then, the lumen contour was detected, and the IVUS image was transformed into polar coordinates. Finally, TPS interpolation was used for elastic registration of the lumen contour points detected between successive two frames in a sub-sequence.

An alternating minimization strategy that was proposed in [92] was utilized in [93] for the registration of pre- and post-biopsy volumes. The cost function that was proposed in [54,66] was modified, and utilized in [14–16,65,70], and there was no optimization performed in these implementations.

## 5. Registration techniques and results by anatomy

Ultrasound imaging has been used to image different organs of the human body over the past decades. In the following sections, the results of ultrasound-to-ultrasound (monomodal) registration algorithms are discussed based on the type of organ or anatomy for which they were implemented and validated. Table 4 summarizes some key anatomic results and references.

### 5.1. Head

In [34], a novel robust method is presented to register three-dimensional ultrasound fetal head images. They used Gabor filters

**Table 4**  
Representative ultrasound registration systems for specific anatomy.

Anatomy	Reference	Registration type	Accuracy of results
Head	[34]	3D Rigid	92% successful registration
Neck	[64]	2D Rigid	5.1 pixel registration error
Breast	[50]	3D Non-Rigid	1.2 ± 0.9 mm
Heart	[53]	2D Non-Rigid	0.5 ± 1.535 pixels
Liver	[62]	2D Non-Rigid	6.9700 pixels SSD error
Kidney	[30]	2D Non-Rigid	0.131 pixel RMS error
Gall bladder	[13]	3D Rigid	<10 mm registration error
Bone	[22]	3D/4D Rigid	<1 mm registration error
Prostate	[51]	3D Non-Rigid	1.96 ± 0.85 mm registration error
Fetal imaging	[25]	3D Non-Rigid	0.78 ± 0.633 mm mosaicking error

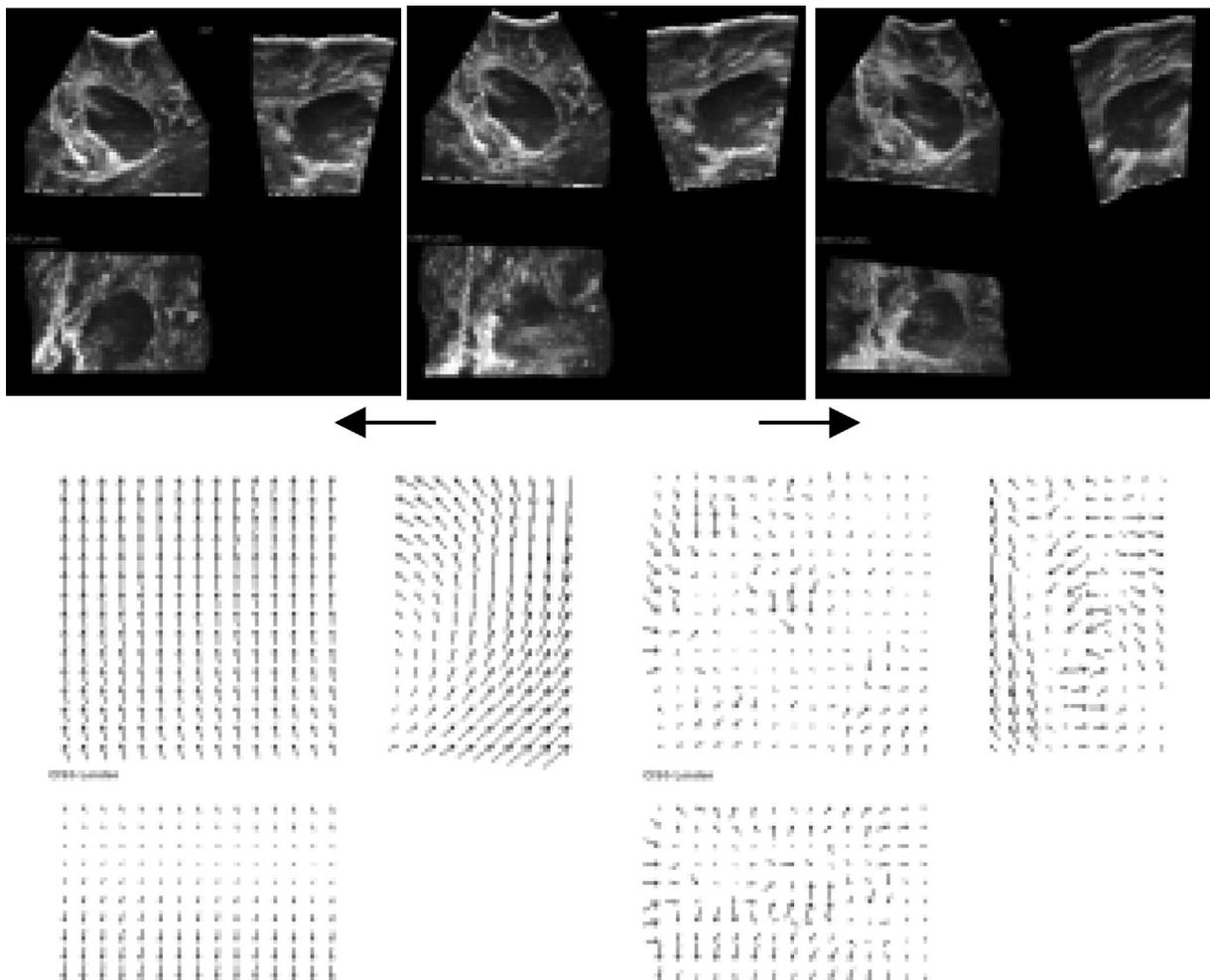
to extract textures and edge features, the correlation ratio as the similarity metric and a mean-shift based optimizer with Powell's direction set method. As a result, the proposed method was able to achieve a successful registration rate of 92% for large initial misalignment with correlation ratio metric. The algorithm was evaluated on two three-dimensional volumes of fetal heads. Among the different metrics that were used, CR performed the best with the greatest capture range.

The minimization of a non-rigid three-dimensional registration problem in a multi-resolution framework was done in [4]. The optimizer used in this approach was a multi-grid reweighted least squares minimization algorithm employing an iterative Gauss-Seidel scheme. They had applied their embedded multi-grid and multi-resolution approach on artificially deformed volumes of the brain of an 8-month old baby, and on the brain of a pig. With their regularized Perona and Malik (P&M) approach, they achieved a better metric score, lowering the baby-brain Mean Squared Error (MSE) metric score by about 5.7% and reducing the angular error by about 2.3% to 13.8°, in comparison to an approach that does not incorporate any multi-scale framework.

A non-rigid registration algorithm designed in [49] that was applied to three-dimensional ultrasound volumes acquired before and after opening the dura of the brain of two patients with brain

tumors. The algorithm operated by optimizing a cost function made up of global and local motion models. They used the gradient descent optimizer proposed in [74] to minimize the cost function. The results before and after registration were compared; the registration method involved an affine plus free-form deformation (FFD) with a control point spacing of 4. They computed the overlap of the segmented tumor volume before and after registration, and found that the overlap after registration jumped to 96% from a 76% overlap before registration. Fig. 1 shows results that the affine registration is applied to ultrasound volume after opening the dura.

A near real-time registration algorithm (RESOUND) was developed in [27] for registering ultrasound volumes acquired before and after the resection of brain tumors. The algorithm incorporated the minimization a cost function comprised of a similarity term based on NCC, and a smoothness term. The optimization was done by taking the analytic derivative of NCC, and using a stochastic gradient descent algorithm as in [86,87]. The deformation field was estimated using cubic B-splines. The authors validated their results by manually providing corresponding landmarks in each volume, and compared the volumes before and after registration by estimate the mean target registration error. The initial error before registration was 3.7 mm, while the error after registration dropped to 1.58 mm.



**Fig. 1.** Deformation field. **Top:** left: Ultrasound volume after opening the dura, middle: Ultrasound volume after opening the dura + affine registration, right: Ultrasound volume before opening the dura. **Bottom:** left: Deformation field for affine registration, right: Deformation field for free-form deformation (control point spacing 4 mm) without the affine component. (Medical Image Computing and Computer-Assisted Intervention – MICCAI 2003, Non-rigid Registration of 3D Ultrasound Images of Brain Tumours Acquired during Neurosurgery, Volume 2879, 2003, p414, Marloes M. J. Letteboer, Peter W. A. Willems, Max A. Viergever, and Wiro J. Niessen, © Springer-Verlag Berlin Heidelberg 2003, with permission of Springer.) [49].

## 5.2. Neck

In [64], a novel hybrid procedure for registering envelop detected radio frequency ultrasound data of human neck was proposed, using global statistics and local textural features. Globally, Hellinger distance between distributions is utilized and Fuzzy Local Binary Patterns (FLBP) is used on a local scale to perform registration. As a result, the median registration error on five datasets using the methodology put forth in [64] is much lower than other similarity measures such as SSD and NCC.

## 5.3. Breast

In [54], a fully automated non-rigid registration algorithm was proposed for registering free-hand ultrasound volumes of the female breast. A Bayesian regularization framework incorporating a block matching scheme was developed with the minimization achieved through conjugate gradient descent. The deformation field was generated by fitting cubic tensor product B-spline meshes. The boundary of the cysts and background texture were well preserved in in vivo ultrasound images of the breast after registration. Fig. 2 shows results of registering and compounding two sweeps of a breast fibroadenoma.

A sub-volume based volumetric registration (SURE) algorithm was developed in [24], which was used to register synthetically warped ultrasound volumes of the female breast. The algorithm used the TPS model to estimate the deformation field between original position of subvolumes and new position of subvolumes. The optimizer they used in this paper was the Nelder-Mead simplex method [72]. The synthetic deformations that were applied to a reference volume were in the range of 1.5–2.5 mm, and this was reduced by over 85% when using their registration algorithm. This was significantly higher than the 59% and 50% reduction in deformation obtained by using rigid and affine transforms respectively.

In [48], the first semi-automated approach was proposed for non-rigid registration of ultrasound images of the breast. The optimizer used in this paper was the Nelder simplex algorithm. TPS warps were computed by having the user manually click control

points in the moving image, and using the optimizer to maximize the MI between the fixed and the moving images. They prove that their algorithm achieves a 27% increase in mutual information by using the nine point TPS solution over the full affine registration.

In [50], the MIAMI-FUSE registration software [84] was utilized for non-rigid ultrasound volume registration of the female breast. Initially, affine registration was done by manually selecting control points. After initial alignment, at least one additional control point was needed for elastic registration using TPS. The location of other grayscale voxels in the moving volume was interpolated using TPS, and the three-dimensional volumes were registered. Nelder-Mead simplex method was used for optimization of the cost function. The algorithm successfully registered automated whole breast ultrasound (ABU) volumes, and reduced the mean registration between manually annotated landmark points after registration to  $1.2 \pm 0.9$  mm.

In [85], an approach was put forth which was an extension of the demons algorithm. They added an extra force to the optical flow equation of the demons algorithm called the inertia force. By adding this extra term, they were able to achieve better non-rigid registration results over the traditional demons algorithm. The optimizer used here was a second order gradient descent on the SSD criterion. The paper showed that the SSD ratio computed over multiple iterations dropped quickly and was lower than comparable registration algorithms, thus indicating successful registration.

## 5.4. Heart

In [78], a registration framework was proposed to register real-time volume ultrasound images of the heart. This rigid registration framework is based on mutual information and uses Nelder-Mead method as an optimizer. Over a range of image sets and parameters, the reported translation error range was 0.32 mm to 2.58 mm, and the rotation error range was 0.17 degree to 9.25 degree.

In [19], a technique was developed to temporally align two sequences of pre-stress and post-stress three-dimensional images for stress echocardiography, followed by a spatial registration process. This registration uses mutual information as similarity metric and a downhill simplex method as optimizer. As a result, working with three-dimensional stress echocardiography it is shown that this registration framework can potentially “improve the diagnostic accuracy of stress testing”.

In [5], an algorithm for Intra-Vascular Ultrasound (IVUS) image registration was implemented by transforming the images from image coordinates into polar coordinates. Here, the registration energy functional to be minimized comprised the sum of squared differences of intensities, and the sum of squared differences in radial gradient in each pixel of the region of interest (ROI). The minimization was done based on the Levenberg-Marquardt optimization [75]. The mean and standard deviation for the X error with a control point spacing sequence (CPSS) of 30–20, and a weighting coefficient of 0.5 was  $0.023 \pm 0.156$  pixels.

In [63], the paper proposed a unique way of estimating the global spatio-temporal deformation field for a sequence of left ventricle ultrasound images using a gradient descent optimization method and a B-spline parametric model. In the ultrasound sequence containing realistic noise, the geometric error in pixels was 1.265 pixels (corresponding to 5% displacement). This result suggests that the spatio-temporal algorithm works better than a previously published algorithm in [94].

In [52], a non-rigid registration algorithm was proposed which incorporated a twisting and bending model to register three-dimensional volumes of the carotid artery. They used the Powell optimizer [58,60] to estimate the six parameters of the rigid

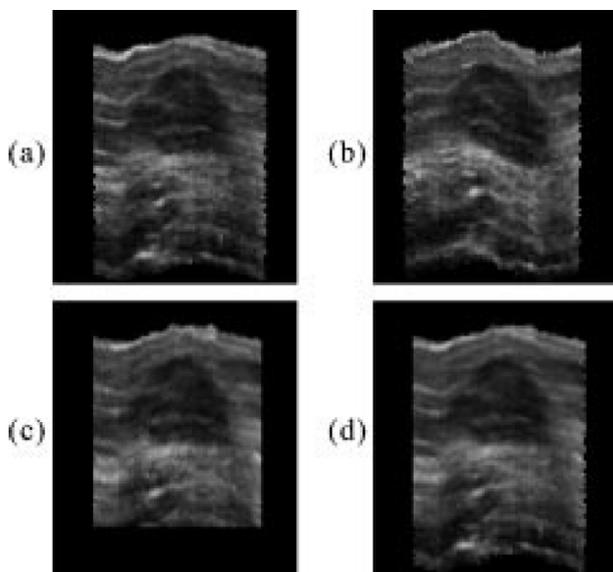


Fig. 2. Registration and compounding of two sweeps of a breast fibroadenoma. (a) Sweep 1 (the reference image). (b) Sweep 2 (the floating image). (c) The result of registering Sweep 2 to Sweep 1. (d) Compounding of Sweep 1 and 2 after registration. © [2002] IEEE. Reprinted, with permission, from [54].

transformation along with the six parameters of the non-rigid transform. The overall mean and standard deviation for the mean registration error between the fixed and moving volumes was  $1.03 \pm 0.23$  mm for non-rigid registration over  $1.50 \pm 0.50$  mm for rigid registration.

A variational framework was used in [6] to solve the problem of non-rigid ultrasound registration. Here, the cost function consisted of a data term and a smoothness term, which could be minimized for gradient descent using the Euler-Lagrange method. The Euler-Lagrange equations that were derived in their paper were solved using an alternating minimization approach. The SSD between the fixed and moving images after intensity-based registration algorithm was 37.37 pixels, while the SSD after intensity + local phase information based non-rigid registration was 34.76 pixels.

An algorithm that registered three-dimensional volumes through the use of an aortic root model-based segmentation was described in [32]. The performance of the registration algorithm was illustrated by the use of the Canny edge detector to extract boundaries in the image sets, and it was clear that the algorithm did well in aligning the edges and boundaries in the images.

A multi-resolution non-rigid registration algorithm was developed in [60] to track the mitral valve annulus in three-dimensional ultrasound volumes. The optimizer that was used in this algorithm was the LBFGS optimizer. The deformation field was estimated using three-dimensional third order B-splines. The tracking algorithm produced an average root mean squared error (RMSE) of  $1.96 \pm 0.46$  mm over all the datasets. The same non-rigid registration algorithm was also used in [33,56–59].

A two-step algorithm for elastic registration of IVUS frames in a sequence was developed in [53]. First, rigid registration was performed to align coronary artery IVUS ultrasound B-scans. Then, the lumen contour was detected, and the IVUS image was transformed into polar coordinates. Finally, TPS interpolation was used for elastic registration of the lumen contour points detected between successive two frames in a sub-sequence. The elastic registration error was measured between five image pairs and compared against the iterative closest point (ICP) algorithm [95]. The mean and standard deviation of the elastic registration error of one image pair was  $0.519 \pm 1.464$  pixels, and this was lower than the ICP-based registration error  $0.778 \pm 2.327$  pixels.

### 5.5. Liver

In [55], a methodology for image-based real time tracking for 4D ultrasound data was proposed using rigid image registration to deduce the positioning of each ultrasound frame in a global coordinate system. They used sum of squared differences (SSD) as similarity metric, a non-linear Nelder-Mead algorithm as optimizer and validated their methodology using liver scan. As a result, they proved that their method is able to provide results in agreement with previously developed magnetic-based tracking technique, which is widely used in clinical industry nowadays.

In [20], a registration framework was developed using mutual information to register 2-D liver images rigidly. Dividing rectangles and Nelder-Mead method are used as optimization algorithm. The results demonstrate that their automatic registration framework can reach minimum value of 0.719 and 0.98 using DIRECT and Nelder-Mead methods respectively after 50 iterations. Thus it is accurate, robust and well suited for clinical applications.

In [96], a fast affine registration framework was developed to compensate in real-time for liver three-dimensional motion/displacement due to breathing using block-matching method. Normalized cross correlation is used as similarity metric. The optimization problem is mapped from non-homogeneous to homogeneous and solved by using replicator dynamics efficiently.

Validation was performed on 91 subjects, with two ultrasound volumes from each, with a “mean registration error of 1.8 mm”.

The diagonal (d), vertical (v), and horizontal (h) components of the discrete Meyer wavelet transform was used in [62] to generate energy maps from liver ultrasound images. They minimized the energy functional with SSD as similarity metric and diffusion regularizer. They used fixed point iteration method as the optimizer. The algorithm outperformed an intensity-based registration with a SSD of 6.97 over 29.7 for the intensity-based registration.

An algorithm was implemented in [67] that registered an entire 4D (three-dimensional + time) sequence of liver ultrasound volumes in a group-wise fashion, and avoided bias towards a specifically chosen reference time point. The optimizer that was used in this paper was the stochastic gradient descent algorithm, and the parameters for this algorithm were from elastix toolbox [88]. Their paper showed that the average 75% percentile of the registration error (1.0 mm) is lower than the average 75% percentile of the bias-corrected inter-observer error (1.4 mm).

In [61], the non-rigid registration algorithm utilized a cost function that had a diffusion regularizer, and a variational minimization approach was used to register the three-dimensional volumes of the liver together. Here, the cost function was minimized by applying calculus of variations to obtain a non-linear partial differential equation, which was then solved using the modified fix-point iteration with incremental updates being regularized [83]. The gain in similarity was measured by computing the ratio of the differences in similarity before and after registration to the original similarity. For the SSD, the ratio ranged between 32.82% and 48.47% with an average of 40.78% for 5 datasets.

A three-dimensional registration algorithm was implemented in [15] to register human liver volumes using a novel similarity metric defined in Section 3.3.1 and a cost function that was proposed in [54,66]. For validation of the results, they manually selected points in the fixed and moving data sets, and estimated the distance between them after automatic registration. The error was about 1.4 mm. Similar results were proposed in [14,16,65,70].

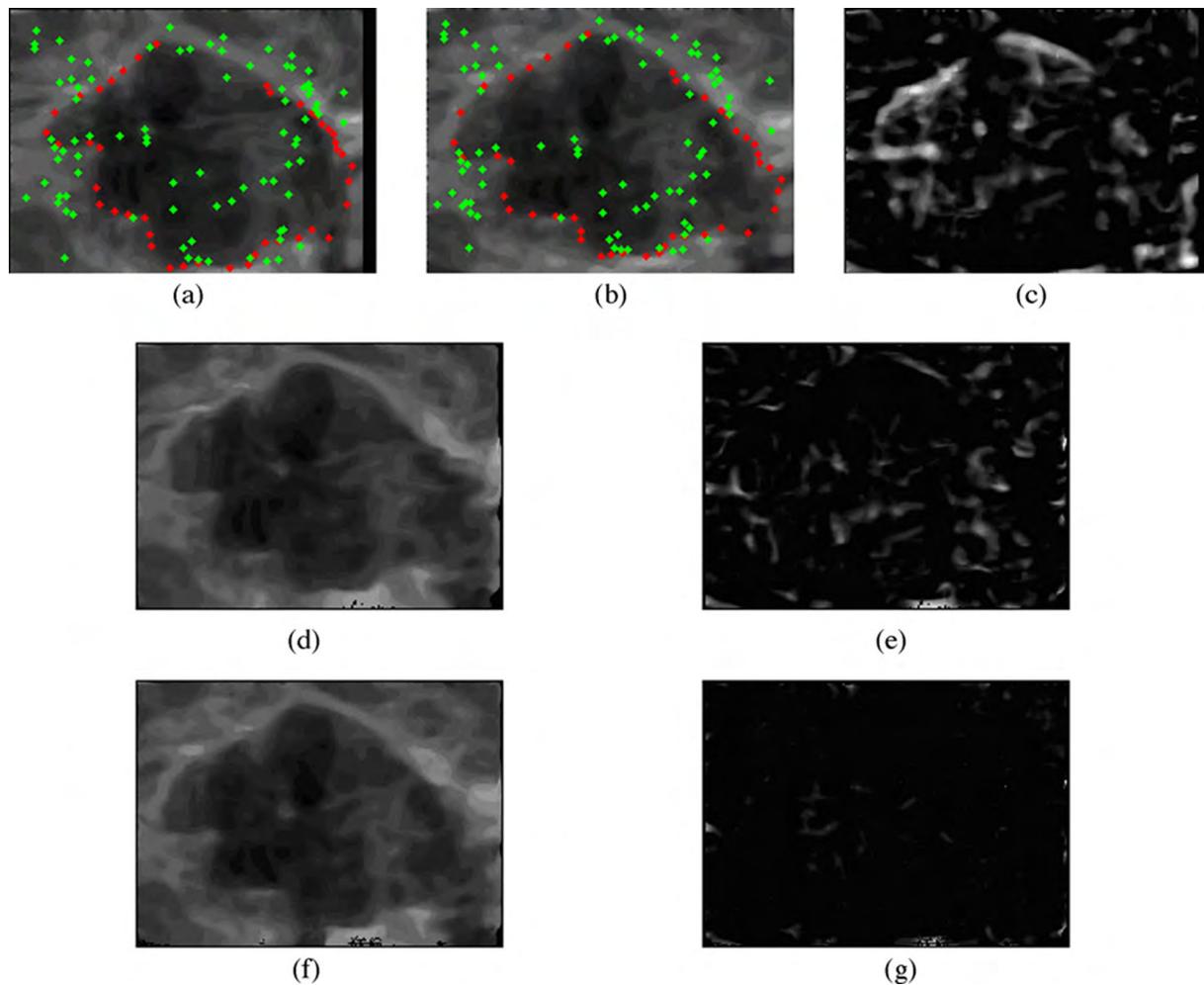
### 5.6. Kidney

In [17], point pair correspondences were extracted that described the shape of objects in ultrasound images, and then used in a Bayesian framework, where the displacement field was estimated using MAP framework. The fast optimization algorithm proposed in [89] to obtain the velocity field by simple scale-space convolution was used for registering the images together. They achieved a cross-correlation (CC) value of 0.979 by using their registration algorithm over the CC value of 0.92 before registration. See in Fig. 3.

In [30], a fuzzy correspondence matrix was used in addition to the compatibility coefficient (see Section 3.3.1) in order to estimate the transformation that maps points in the moving ultrasound image to the fixed ultrasound image. The transformation was defined using the Thin Plate Spline (TPS) model. The mean Root Mean Square (RMS) error in a 9 frame separation was  $1.31 \times 10^{-1}$ , compared to  $2.22 \times 10^{-1}$  in TPS – Robust Point Matching (TPS – RPM) proposed in [97].

### 5.7. Gall bladder

In order to accurately perform three-dimensional spatial compounding for improvement in quality of ultrasound data, [13] adapted a previously developed multi-model CT to MRI registration algorithm to three-dimensional ultrasound data. Using a multi-resolution approach, they developed a correlation-based registration framework and it has been validated with gall bladder



**Fig. 3.** (a) The template image ( $220 \times 160$ ) with 135 landmark points. (b) The target image ( $220 \times 160$ ) with 135 landmark points. (c) The difference image between template and target images. (d) The deformed template image using Christensen's fluid method. (e) The difference image between d and b. (f) The deformed template image using [the] method. (g) The difference image between (f) and (b). Reprinted from Publication [17] with permission from Elsevier.

ultrasound data. As a result, the rigid registration translational and rotational errors are approximately maintained within 10 mm and 10 degrees for a total of 600 B-scans.

### 5.8. Bone

In [22], a method was developed for 3D/4D ultrasound registration of the bone in order to reduce the invasiveness of Computer Assisted Orthopaedic Surgery (CAOS). Normalized cross correlation is used as similarity metric and the simplex method of Nelder and Mead is selected for optimization process. As results, 65% of the cases show a successful registration result with average time of 10 s for each.

### 5.9. Prostate

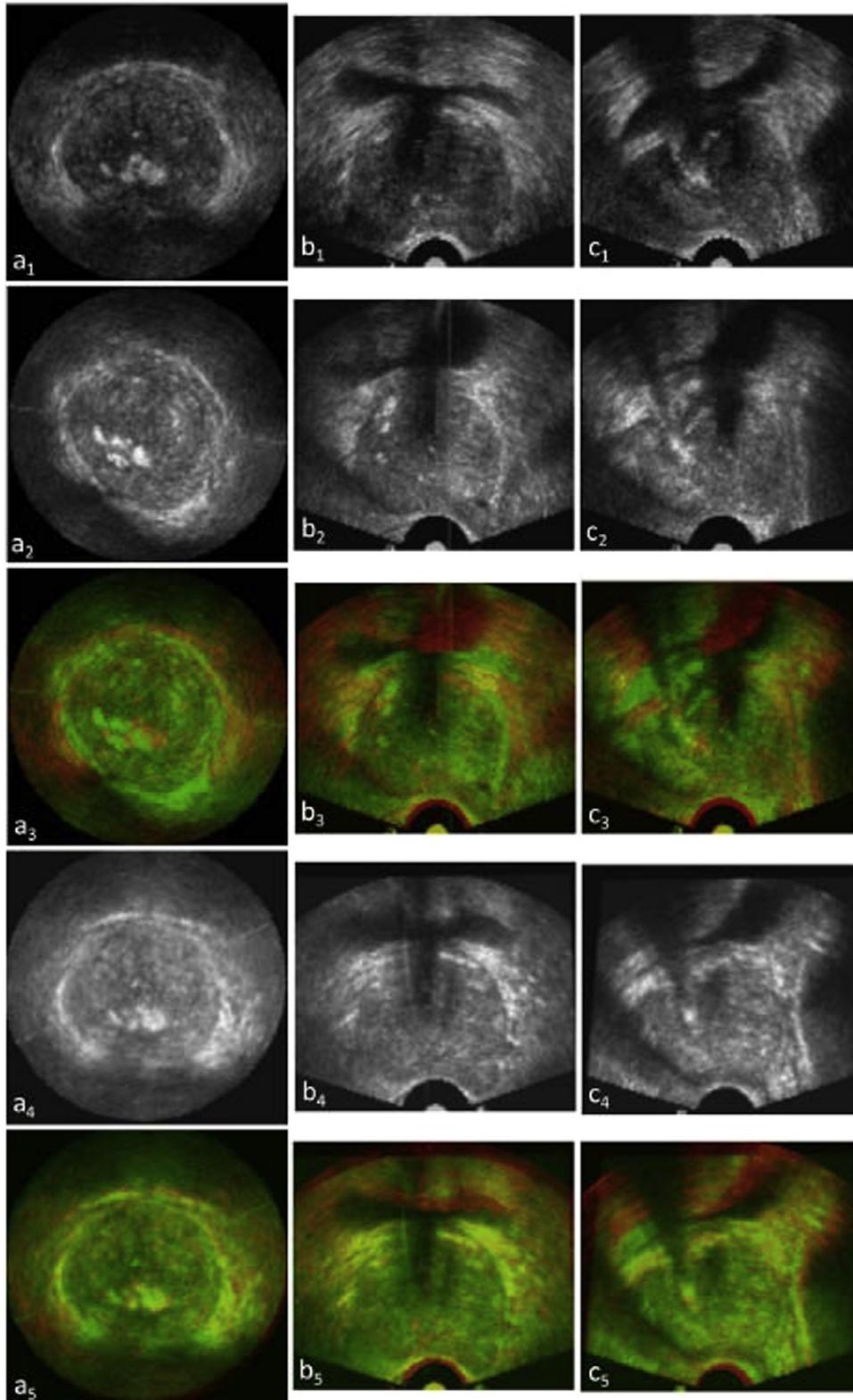
Inter-session registration of three-dimensional *trans*-rectal ultrasound (TRUS) was achieved in [51]. They utilized two different non-rigid algorithms for comparison of performance. The first was surface-based registration using an initial ICP-based rigid registration of three-dimensional TRUS volumes followed by non-rigid registration using TPS. The second was image-based registration with a block matching approach with MI as the similarity metric, followed by non-rigid registration using cubic B-splines. The optimizer that was used here was the Broyden-Fletcher-Goldfarb-Shannon (LBFGSB) optimizer [76]. The pre-registration mean target reg-

istration error (mTRE) for the whole gland was  $7.36 \pm 4.17$  mm, while the mTRE after non-rigid image-based registration was  $1.96 \pm 0.85$  mm. Results are shown in Fig. 4.

For the guidance of needle biopsies in prostate image-guided radiation therapy, an alternating minimization strategy that was proposed in [92] was utilized in [93]. The algorithm used B-splines in order to transform and register the post-biopsy volumes to the pre-biopsy volumes. In [46], three registration methods were evaluated for the registration of three-dimensional-transabdominal ultrasound volumes. These volumes were acquired for the setup of postprostatectomy patients during radiation therapy. The similarity metric used for the registration was mutual information, and since they were only estimating translations, the three translation parameters were optimized using the adaptive stochastic gradient descent optimizer [83].

### 5.10. Fetal imaging

In [25], the registration energy cost function was modeled using Markov Random Fields (MRF), and optimized using a parallelized alpha-expansion technique [73]. The clinical application of this paper was in the construction of fetal mosaics for use in a training simulator, the proposed method outperformed the spatial compounding techniques [98] that required almost perfect alignment between volumes. See Fig. 5.

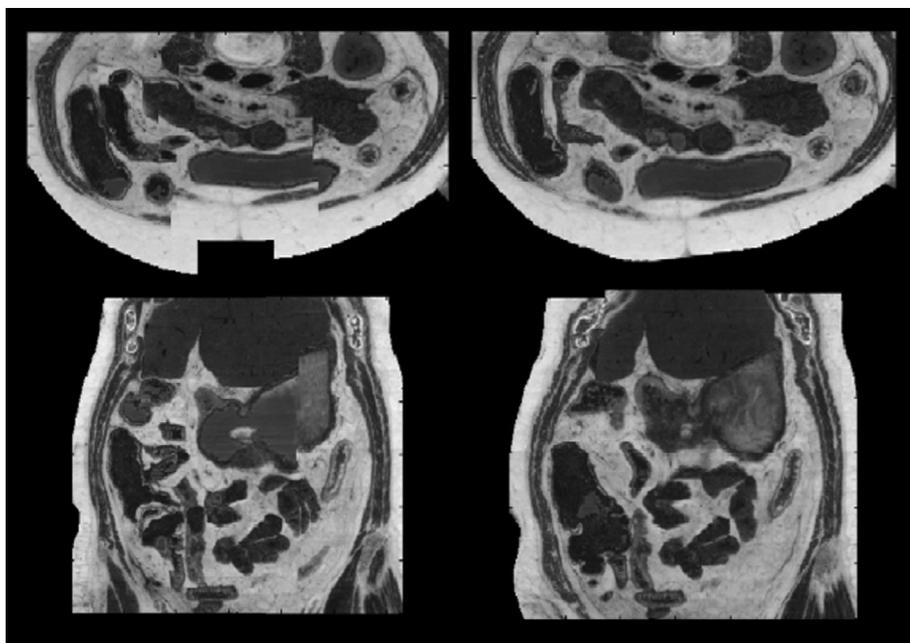


**Fig. 4.** Visual assessment of three-dimensional TRUS registration.  $a_1$ ,  $b_1$  and  $c_1$  are post-biopsy images in three directions.  $a_2$ ,  $b_2$  and  $c_2$  are pre-biopsy images, and  $a_3$ ,  $b_3$  and  $c_3$  are the fusion images between pre- and post-biopsy images.  $a_4$ ,  $b_4$  and  $c_4$  are the 3D registered pre-biopsy images; and  $a_5$ ,  $b_5$  and  $c_5$  are the fusion images between registered pre- and post-biopsy images. Reprinted from Publication [93] with permission from SPIE.

## 6. Discussion, Shortcomings, and future opportunities

Many of the cited algorithms are freely available from the author's websites. An increasingly large number of registration

algorithms also have free open-source implementations in the Insight Toolkit (ITK) [99] and its associated Insight Journal. Although a variety of mono-modality ultrasound registration frameworks were proposed and developed over the past few



**Fig. 5.** Mosaicking of multiple volumes through the registration approach in [25]. The left half shows slices from the mosaiced volume when no registration is performed. The right half demonstrates how [the] algorithm can reconstruct the anatomy seamlessly, using multiple partially overlapping volumes. Reprinted from Publication [25] with permission from Elsevier.

decades, there are still shortcomings and unsolved issues in the field of ultrasound registration.

We have not summarized how rapidly, nor in what fraction of cases, reasonable non-rigid registration can be achieved. That assessment should be done periodically, particularly with continually evolving software packages such as commercial software.

Opportunities for future research are suggested below.

### 6.1. Diverse anatomy and ultrasound systems

With all the papers that have been reviewed, most registration frameworks were validated on human liver, cardiac and kidney anatomical structures. Ultrasound registration on other structures such as bone, bladder and peripheral anatomy are not as well studied, nor are animal models such as mice and pigs. More attention is needed on less common ultrasound configurations, such as intravascular acquisition or high-frequency small-animal imaging.

### 6.2. Interactive nature of ultrasound acquisition

As discussed in previous sections, ultrasound registration algorithms are robust to some changes, such as gain, focal depth etc. However, other differences remain problematic. Ultrasound is a uniquely interactive imaging modality. The probe is freely positioned during image acquisition, allowing imaging from different viewpoints while providing visual feedback that can be used to physically manipulate the patient (or animal). Careful manipulation and choice of viewpoint are often crucial to acquiring diagnostic quality images, leading to engineering complexities and untapped potential.

#### 6.2.1. Physical probe tracking

In many clinical applications, free-hand ultrasound volumes are acquired, and the position of the ultrasound probe is tracked directly using either a magnetic or optical tracking system. However, existing probe tracking systems are somewhat restrictive and/or not extremely accurate. More importantly, the true three-dimensional location of the probe is obtained relative to the track-

er's coordinate system, and not relative to the patient's anatomy. This can be a hindrance to the registration of images since the internal structures of the patient's anatomy can deform while being imaged, but the tracker cannot accurately quantify these movements as it provides only global motion estimates relative to the tracker coordinate system. Prior work has been done to instead directly track an ultrasound transducer relative to a patient's anatomy [8,35,100–103]. Most researchers use probe tracking as (an often poor) surrogate for rigid registration between consecutive images, in many cases followed by non-rigid registration as a fix-it step. However, given the under-constrained nature of non-rigid registration, it is preferable to begin with as accurate of a rigid transform as possible.

#### 6.2.2. Viewpoint specific

Ultrasound allows image acquisition from different viewpoints. Even when imaging the same anatomy without any physical manipulation, different viewpoints will still produce different pixel intensities for the exact same micro-volume of tissue. Other effects such as refraction, visual occlusion and shadowing due to the complex internal tissue structures can make it difficult to register ultrasound images acquired from different viewpoints using only image-based approaches.

#### 6.2.3. Interactive manipulation during ultrasound acquisition

Ultrasound imaging routinely requires physical manipulation of the subject being scanned, typically to visually judge the effects of compression or to push intervening anatomy out of the way. Often, the transducer itself is used as the primary manipulation device to compress or displace the tissue. Eventually, independent tracking of the ultrasound probe (as discussed in Section 6.2.1) might be used to generate good a priori models of deformation for use during non-rigid registration algorithms. To accurately predict both tissue compression and lateral tissue displacement, these models should take into account both the transducer's compression force and the total trajectory the transducer has followed. Open problems include continuous measurement of transducer compression, tracking the trajectory of the transducer relative to the anatomy,

and building predictive models from these measurements. With 2D array transducers, which are capable of real-time 3D acquisition, live tracking of tissues in the images between acquired image volumes may also provide useful preregistration.

## Acknowledgements

This review was funded in part by U.S. National Institutes of Health R01 grant 1R01EY021641, U.S. National Library of Medicine contract HHSN27620100058OP, and U.S. Department of Defense Peer Reviewed Medical Research Program award PR130773 (HRPO Log No. A-18237).

## References

- [1] T. Selbekk, A.S. Jakola, O. Solheim, T.F. Johansen, F. Lindseth, I. Reinertsen, G. Unsgård, Ultrasound imaging in neurosurgery: approaches to minimize surgically induced image artefacts for improved resection control, *Acta Neurochir. (Wien)* 155 (2013) 973–980, <http://dx.doi.org/10.1007/s00701-013-1647-7>.
- [2] U. Wagnetz, M. Atri, C. Massey, A.C. Wei, U. Metser, Intraoperative ultrasound of the liver in primary and secondary hepatic malignancies: comparison with preoperative 1.5-T MRI and 64-MDCT, *Am. J. Roentgenol.* 196 (2011) 562–568, <http://dx.doi.org/10.2214/AJR.10.4729>.
- [3] K.L. Hansen, M.M. Pedersen, H. Møller-Sørensen, J. Kjaergaard, J.C. Nilsson, J.T. Lund, J.A. Jensen, M.B. Nielsen, Intraoperative cardiac ultrasound examination using vector flow imaging, *Ultrasound. Imaging* 35 (2013) 318–332, <http://dx.doi.org/10.1177/0161734613505552>.
- [4] I. Pratikakis, C. Barillot, P. Hellier, E. Memin, Robust multiscale deformable registration of 3d ultrasound images, *Int. J. Image Graphics* 3 (2003) 547–565.
- [5] K. Oakeson, Quantification of cross-sectional artery wall motion with IVUS image registration, *Med. Imaging* 2004 (5373) (2004) 119–130, <http://dx.doi.org/10.1117/12.536523>.
- [6] J. Woo, B.W. Hong, C.H. Hu, K.K. Shung, C.C.J. Kuo, P.J. Slomka, Non-rigid ultrasound image registration based on intensity and local phase information, *J. Signal Process. Syst.* 54 (2009) 33–43, <http://dx.doi.org/10.1007/s11265-008-0218-2>.
- [7] R.C.G. Martin, D.A. North, Enhanced ultrasound with navigation leads to improved liver lesion identification and needle placement, *J. Surg. Res.* 200 (2015) 420–426, <http://dx.doi.org/10.1016/j.jss.2015.09.003>.
- [8] J. Wang, V. Shivaprabhu, J. Galeotti, S. Horvath, V. Gorantla, G. Stetten, Towards Video Guidance for Ultrasound, Using a Prior High-Resolution 3D Surface Map of the External Anatomy, 2014, pp. 51–59.
- [9] J. Ashburner, K. Friston, Rigid Body Registration and Interpolation, in: *Hum. Brain Funct., Second ed.*, 2003, pp. 635–653, <http://dx.doi.org/10.1016/B978-012264841-0/50034-2>.
- [10] M. Sarode, P. Deshmukh, Reduction of speckle noise and image enhancement of images using filtering technique, *Int. J. Adv. Technol.* 2 (2011) 30–38, <http://www.ijct.org/index.php/ijct/article/viewArticle/speckle-noise>.
- [11] K.W.J. Segen, M. Kulbacki, Optimization of joint detector for ultrasound images using mixtures of image feature descriptors, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9012, 2015, pp. 1–2, <http://dx.doi.org/10.1007/978-3-319-15705-4>.
- [12] A.H. Gee, G.M. Treece, R.W. Prager, C.J.C. Cash, L. Berman, Rapid registration for wide field of view freehand three-dimensional ultrasound, *IEEE Trans. Med. Imaging* 22 (2003) 1344–1357, <http://dx.doi.org/10.1109/TMI.2003.819279>.
- [13] R.N. Rohling, A.H. Gee, L. Berman, Automatic registration of 3-D ultrasound images, *Ultrasound Med. Biol.* 24 (1998) 841–854, <http://www.sciencedirect.com/science/article/pii/S030156299700210X>.
- [14] S. Khalilgh, C.G.M. Leung, K. Hastrudi-Zaad, P. Foroughi, C. Nguan, P. Abolmaesumi, Experimental validation of an intrasubject elastic registration algorithm for dynamic-3D ultrasound images, *Med. Phys.* 39 (2012) 5488–5497, <http://dx.doi.org/10.1118/1.4742056>.
- [15] P. Foroughi, P. Abolmaesumi, K. Hashtrudi-Zaad, Intra-subject elastic registration of 3D ultrasound images, *Med. Image Anal.* 10 (2006) 713–725, <http://dx.doi.org/10.1016/j.media.2006.06.008>.
- [16] C. Leung, K. Hashtrudi-Zaad, P. Foroughi, P. Abolmaesumi, A real-time intrasubject elastic registration algorithm for dynamic 2-D ultrasound images, *Ultrasound Med. Biol.* 35 (2009) 1159–1176, <http://dx.doi.org/10.1016/j.ultrasmedbio.2008.12.018>.
- [17] X. Lu, S. Zhang, W. Yang, Y. Chen, SIFT and shape information incorporated into fluid model for non-rigid registration of ultrasound images, *Comput. Methods Programs Biomed.* 100 (2010) 123–131, <http://dx.doi.org/10.1016/j.cmpb.2010.03.005>.
- [18] B. Wang, Y. Shen, A method on calculating high-dimensional mutual information and its application to registration of multiple ultrasound images, *Ultrasonics* 44 (2006).
- [19] R. Shekhar, V. Zagrodsky, M.J. Garcia, J.D. Thomas, Registration of real-time 3-D ultrasound images of the heart for novel 3-D stress echocardiography, *IEEE Trans. Med. Imaging* 23 (2004) 1141–1149, <http://dx.doi.org/10.1109/TMI.2004.830527>.
- [20] S.R.R. Suganya, R. Kirubakaran, Registration of ultrasound liver images using mutual information technique, *Adv. Intell. Syst. Comput.* 246 (2014) 155–162, <http://dx.doi.org/10.1007/978-81-322-1680-3>.
- [21] V. Zagrodsky, R. Shekhar, J.F. Cornhill, Multi-Function Extension of Simplex Optimization Method for Mutual Information-Based Registration of Ultrasound Volumes, *Image (Rochester, N.Y.)*, vol. 4322, 2001, pp. 508–515, <http://dx.doi.org/10.1117/12.431124>.
- [22] J. Schers, J. Troccaz, V. Daanen, C. Foudard, C. Plaskos, P. Kilian, 3D/4D ultrasound registration of bone, *Proc. - IEEE Ultrason. Symp.*, 2007, pp. 2519–2522, <http://dx.doi.org/10.1109/ULTSYM.2007.634>.
- [23] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1990) 629–639, <http://dx.doi.org/10.1109/34.56205>.
- [24] J.F. Kruecker, G.L. LeCarpentier, J.B. Fowlkes, P.L. Carson, Rapid elastic image registration for 3-D ultrasound, *IEEE Trans. Med. Imaging* 21 (2002) 1384–1394.
- [25] J. Kutarnia, P. Pedersen, A Markov random field approach to group-wise registration/mosaicing with application to ultrasound, *Med. Image Anal.* 24 (2015) 106–124, <http://dx.doi.org/10.1016/j.media.2015.05.011>.
- [26] E. Memin, P. Perez, Dense estimation and object-based segmentation of the optical flow with robust techniques, *IEEE Trans. Image Process.* 7 (1998) 703–719, <http://dx.doi.org/10.1109/83.668027>.
- [27] H. Rivaz, D.L. Collins, Near real-time robust non-rigid registration of volumetric ultrasound images for neurosurgery, *Ultrasound Med. Biol.* 41 (2015) 574–587, <http://dx.doi.org/10.1016/j.ultrasmedbio.2014.08.013>.
- [28] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man. Cybern.* 9 (1979) 62–66, <http://dx.doi.org/10.1109/TSMC.1979.4310076>.
- [29] J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 8 (1986) 679–698, <http://dx.doi.org/10.1109/TPAMI.1986.4767851>.
- [30] J.-H. Lee, Y.K. Seong, M. Park, K.-G. Woo, J. Ku, H.-J. Park, Non-rigid ultrasound image registration using generalized relaxation labeling process, *Image Process. Mach. Vis. Appl.* VI 8661 (2013) 1–6, <http://dx.doi.org/10.1117/12.2003233>.
- [31] S. Menet, P. Saint-Marc, G. Medioni, B-Snakes: Implementation and Application to Stereo, 1990.
- [32] B. Matinfar, L. Zagorchev, Non-rigid Registration of 3D Ultrasound Images Using Model-based Segmentation, in: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 323–328, <http://dx.doi.org/10.1109/CVPRW.2014.132>.
- [33] B. Heyde, M. Alessandrini, J. Hermans, D. Barbosa, P. Claus, D. Jan, Conservation to Assess Cardiac Deformation From 3D Ultrasound Recordings, vol. 35, 2016, pp. 501–511.
- [34] F. Cen, Y. Jiang, Z. Zhang, H.T. Tsui, T.K. Lau, Robust Registration of 3-D Ultrasound Images Based on Gabor Filter and Mean-Shift Method, 2004, pp. 304–316.
- [35] R.J. Schneider, D.P. Perrin, N.V. Vasilyev, G.R. Marx, P.J. Del Nido, R.D. Howe, Real-time image-based rigid registration of three-dimensional ultrasound, *Med. Image Anal.* 16 (2012) 402–414, <http://dx.doi.org/10.1016/j.media.2011.10.004>.
- [36] D. Ni, Y.P. Chui, Y. Qu, X. Yang, J. Qin, T.T. Wong, S.S.H. Ho, P.A. Heng, Reconstruction of volumetric ultrasound panorama based on improved 3D SIFT, *Comput. Med. Imaging Graph.* 33 (2009) 559–566, <http://dx.doi.org/10.1016/j.compmedimag.2009.05.006>.
- [37] D. Ni, Y. Qu, X. Yang, Y.P. Chui, Volumetric ultrasound panorama based on 3D SIFT, *Med. Image Comput. Comput. Assist. Interv.* (2008) 52–60, [http://dx.doi.org/10.1007/978-3-540-85990-1\\_7](http://dx.doi.org/10.1007/978-3-540-85990-1_7).
- [38] D.G. Lowe, Distinctive image features from scale invariant keypoints, *Int'l J. Comput. Vis.* 60 (2) (2004) 91–110, <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [39] S.Y. Selmi, E. Promayon, J. Troccaz, 3D-2D ultrasound feature-based registration for navigated prostate biopsy: A feasibility study, in: *2016 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2016, pp. 4109–4112, <http://dx.doi.org/10.1109/EMBC.2016.7591630>.
- [40] M. Baumann, P. Mozer, V. Daanen, J. Troccaz, Towards 3D Ultrasound Image Based Soft Tissue Tracking: A Transrectal Ultrasound Prostate Image Alignment System, in: *Med. Image Comput. Comput. Assist. Interv.*, 2007, Part II., 2007, pp. 26–33, [http://dx.doi.org/10.1007/978-3-540-75759-7\\_4](http://dx.doi.org/10.1007/978-3-540-75759-7_4).
- [41] Z. Wang, G. Slabaugh, G. Unal, Tong Fang, Registration of ultrasound images using an information-theoretic feature detector, *Zhe Wang New Jersey Institute of Technology Department of Electrical and Computer Engineering Newark, NJ 07102S Corporate Research Intelligent Vision and Reasoning D.* (n.d.).
- [42] Z. Wang, S. Member, G. Slabaugh, G. Unal, A.R. Case, An Information-Theoretic Detector Based Scheme for Registration of Speckled Medical Images, vol. 2 (n.d.).
- [43] S. Golemati, A. Gastounioti, Ultrasound-image-based cardiovascular tissue motion estimation, *IEEE Rev. Biomed. Eng.* 9 (2016) 208–218, <http://dx.doi.org/10.1109/RBME.2016.2558147>.
- [44] M. Alessandrini, B. Heyde, S. Queirós, S. Cygan, M. Zontak, O. Somphone, O. Bernard, M. Serresant, H. Delingette, D. Barbosa, M. De Craene, M.O. Donnell, J. D'hooge, Detailed evaluation of five 3D speckle tracking algorithms using synthetic echocardiographic recordings, *IEEE Trans. Med. Imaging* 35 (2016) 1915–1926, <http://dx.doi.org/10.1109/TMI.2016.2537848>.

- [45] P. Viola, W.M.I. Wells, Alignment by maximization of mutual information, *Proc. IEEE Int. Conf. Comput. Vis.* 24 (1995) 16–23, <http://dx.doi.org/10.1109/ICCV.1995.466930>.
- [46] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens, Multimodality image registration by maximization of mutual information, *IEEE Trans. Med. Imaging* 16 (1997) 187–198, <http://dx.doi.org/10.1109/42.563664>.
- [47] R. Shekhar, V. Zagrodsky, Mutual information-based rigid and nonrigid registration of ultrasound volumes, *IEEE Trans. Med. Imaging* 21 (2002) 9–22, <http://dx.doi.org/10.1109/42.981230>.
- [48] C.R. Meyer, J.L. Boes, B. Kim, P.H. Bland, G.L. Lecarpentier, J.B. Fowlkes, M.A. Roubidoux, P.L. Carson, Semiautomatic registration of volumetric ultrasound scans, *Ultrasound Med. Biol.* 25 (1999) 339–347, [http://dx.doi.org/10.1016/S0301-5629\(98\)00148-3](http://dx.doi.org/10.1016/S0301-5629(98)00148-3).
- [49] M.M.J. Letteboer, P.W.A. Willems, M.A. Viergever, W.J. Niessen, Non-rigid registration of 3D ultrasound images of brain tumours acquired during neurosurgery, in: *Med. Image Comput. Comput. Assist. Interv.* 1, 2003, pp. 408–415.
- [50] G. Narayanasamy, G.L. LeCarpentier, S. Zabuawala, J.B. Fowlkes, M. Roubidoux, S. Sinha, P.L. Carson, Non-rigid registration of three-dimensional (3D) grayscale and Doppler ultrasound breast images, in: *Annu. Int. Conf. IEEE Eng. Med. Biol. – Proc.*, 2007, pp. 91–94. <http://dx.doi.org/10.1109/IEMBS.2007.4352230>.
- [51] V. V. Karnik, A. Fenster, J. Bax, L. Gardi, I. Gyacskov, J. Montreuil, C. Romagnoli, A.D. Ward, Evaluation of inter-session 3D-TRUS to 3D-TRUS image registration for repeat prostate biopsies, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 6362 LNCS, 2010, pp. 17–25. [http://dx.doi.org/10.1007/978-3-642-15745-5\\_3](http://dx.doi.org/10.1007/978-3-642-15745-5_3).
- [52] N.D. Nanayakkara, B. Chiu, A. Samani, J.D. Spence, G. Parraga, J. Samarabandu, A. Fenster, Nonrigid registration of carotid ultrasound and MR images using a “twisting and bending” model, *Proc. SPIE* 6914 (2008), <http://dx.doi.org/10.1117/12.770337>. 691411–691411-8.
- [53] Z. Sun, H. Bai, B. Liu, Rigid and elastic registration for coronary artery IVUS images, *Technol. Health Care* 24 (2016) S455–S463, <http://dx.doi.org/10.3233/THC-161168>.
- [54] G. Xiao, J.M. Brady, J.A. Noble, M. Burcher, R. English, Nonrigid registration of 3-D free-hand ultrasound images of the breast, *IEEE Trans. Med. Imaging* 21 (2002) 405–412, <http://dx.doi.org/10.1109/TMI.2002.1000264>.
- [55] O.K. Øye, W. Wein, D.M. Ulvang, K. Matre, I. Viola, Real time image-based tracking of 4D ultrasound data, *Med. Image Comput. Comput. Assist. Interv.* 15 (2012) 447–454. <http://www.ncbi.nlm.nih.gov/pubmed/23285582>.
- [56] B. Heyde, D. Barbosa, P. Claus, F. Maes, J. D’hooge, Influence of the grid topology of free-form deformation models on the performance of 3D strain estimation in echocardiography, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 7945 LNCS, 2013, pp. 308–315. [http://dx.doi.org/10.1007/978-3-642-38899-6\\_37](http://dx.doi.org/10.1007/978-3-642-38899-6_37).
- [57] B. Heyde, R. Jasaityte, D. Barbosa, V. Robesyn, S. Bouchez, P. Wouters, F. Maes, P. Claus, J. D’hooge, Elastic image registration versus speckle tracking for 2-D myocardial motion estimation: a direct comparison in vivo, *IEEE Trans. Med. Imaging* 32 (2013) 449–459, <http://dx.doi.org/10.1109/TMI.2012.2230114>.
- [58] B. Heyde, D. Barbosa, P. Claus, F. Maes, J. D’hooge, Three-dimensional cardiac motion estimation based on non-rigid image registration using a novel transformation model adapted to the heart, *Stat. Atlases Comput. Model. Heart* 7746 (2013) 142–150.
- [59] M. Alessandrini, B. Heyde, S. Cygan, M. Sermesant, H. Delingette, O. Bernard, M. De Craene, J. D’Hooge, Elastic registration vs. block matching for quantification of cardiac function with 3D ultrasound: Initial results of a direct comparison in silico based on a new evaluation pipeline, in: *IEEE Int. Ultrason. Symp. IUS*, 2014, pp. 608–611. <http://dx.doi.org/10.1109/ULTSYM.2014.0149>.
- [60] H. De Veene, P.B. Bertrand, N. Popovic, P.M. Vandervoort, P. Claus, M. De Beule, B. Heyde, Automatic mitral annulus tracking in volumetric ultrasound using non-rigid image registration, in: *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*. 2015–Novem 2015, 1985–1988. <http://dx.doi.org/10.1109/EMBC.2015.7318774>.
- [61] D. Zikic, W. Wein, A. Khamene, D.-A. Clevert, N. Navab, Fast deformable registration of 3D-ultrasound data using a variational approach, *Med. Image Comput. Comput. Assist. Interv.* 9 (2006) 915–923. <http://far.in.tum.de/pub/zikic2006fast/zikic2006fast.pdf>.
- [62] R.E.E.S. Hefny, Wavelet-Based Variational Deformable Registration for Ultrasound Mohamed, 2010, pp. 1017–1020.
- [63] M. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Suhling, P. Hunziker, M. Unser, Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation, *IEEE Trans. Med. Imaging* 24 (2005) 1113–1126.
- [64] T. Klein, M. Hansson, A. Karamalis, N. Navab, Registration of RF Ultrasound Data Using Hybrid Local Binary Patterns Computer Aided Medical Procedures (CAMP), Technische Universität at Malmö University, School of Technology, Sweden, 2012, pp. 1072–1075.
- [65] P. Foroughi, P. Abolmaesumi, A modified HAMMER algorithm for deformable registration of ultrasound images, *Int. Congr. Ser.* 1281 (2005) 236–241, <http://dx.doi.org/10.1016/j.ics.2005.03.201>.
- [66] D. Shen, C. Davatzikos, HAMMER: hierarchical attribute matching mechanism for elastic registration, *IEEE Trans. Med. Imaging* 21 (2002) 1421–1439, <http://dx.doi.org/10.1109/TMI.2002.803111>.
- [67] S. Vijayan, S. Klein, E. Fagertun, F. Lindseth, B. Ystgaard, Validation of A Non-rigid Registration Method for Motion Compensation in 4D Ultrasound of the Liver, Norwegian University of Science and Technology, Trondheim, Norway SINTEF, Norway Biomedical Imaging Group Rotter, Dept. Medical Technology, 2013, pp. 780–783.
- [68] K.K. Bhatia, J. Hajnal, A. Hammers, D. Rueckert, Similarity metrics for groupwise non-rigid registration, *Med. Image Comput. Comput. Assist. Interv.* 10 (2007) 544–552, [http://dx.doi.org/10.1007/978-3-540-75759-7\\_66](http://dx.doi.org/10.1007/978-3-540-75759-7_66).
- [69] C.T. Metz, S. Klein, M. Schaap, T. van Walsum, W.J. Niessen, Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach, *Med. Image Anal.* 15 (2011) 238–249, <http://dx.doi.org/10.1016/j.media.2010.10.003>.
- [70] P. Foroughi, P. Abolmaesumi, Elastic registration of 3D ultrasound images, *Med. Image Comput. Comput. Assist. Interv.* (2005) 83–90.
- [71] B.K.P. Horn, Closed-form solution of absolute orientation using unit quaternions, *J. Opt. Soc. Am. A* 4 (1987) 629, <http://dx.doi.org/10.1364/JOSA.A.4.000629>.
- [72] B.J.A. Nelder, R. Mead, A simplex method for function minimization, *Comput. J.* 7 (1964) 308–313, <http://dx.doi.org/10.1093/comjnl/7.4.308>.
- [73] V. Lempitsky, C. Rother, S. Roth, A. Blake, Fusion Moves for Markov Random Field Optimization, vol. 32, 2009, pp. 1–15.
- [74] D. Rueckert, L.I. Sonoda, C. Hayes, D.L. Hill, M.O. Leach, D.J. Hawkes, Nonrigid registration using free-form deformations: application to breast MR images, *IEEE Trans. Med. Imaging* 18 (1999) 712–721, <http://dx.doi.org/10.1109/42.796284>.
- [75] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math.* 11 (1963) 431–441, <http://dx.doi.org/10.1137/0111030>.
- [76] C. Zhu, R.H. Byrd, P. Lu, J. Nocedal, Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization, *ACM Trans. Math. Softw.* 23 (1997) 550–560, <http://dx.doi.org/10.1145/279232.279236>.
- [77] M.J.D. Powell, On search directions for minimization algorithms, *Math. Program.* 4 (1973) 193–201, <http://dx.doi.org/10.1007/BF01584660>.
- [78] J. Frednck Cornhill Shekhar, V. Zagrodsky, R. Shekhar, Mutual Information Based Registration of Cardiac Ultrasound Volumes, vol. XX, Ponsen & Looijen, 2000, pp. 1–21.
- [79] M.J.D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comput. J.* 7 (1964) 155–162, <http://dx.doi.org/10.1093/comjnl/7.2.155>.
- [80] H.W. Sorenson, Least-squares estimation: from Gauss to Kalman, *IEEE Spectr.* 7 (1970) 63–68, <http://dx.doi.org/10.1109/MSPEC.1970.5213471>.
- [81] M.C. Yip, D.G. Lowe, S.E. Salcudean, R.N. Rohling, C.Y. Nguan, Tissue tracking and registration for image-guided surgery, *IEEE Trans. Med. Imaging* 31 (2012) 2169–2182, <http://dx.doi.org/10.1109/TMI.2012.2212718>.
- [82] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* 24 (1981) 381–395.
- [83] E. Haber, J. Modersitzki, Numerical methods for volume preserving image registration, *Inverse Probl.* 20 (2004) 1621–1638, <http://dx.doi.org/10.1088/0266-5611/20/5/018>.
- [84] C.R. Meyer, J.L. Boes, B. Kim, P.H. Bland, K.R. Zasadny, P.V. Kison, K. Koral, K.A. Frey, R.L. Wahl, Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations, *Med Image Anal.* 1 (1997) 195–206, [http://dx.doi.org/10.1016/S1361-8415\(97\)85010-4](http://dx.doi.org/10.1016/S1361-8415(97)85010-4).
- [85] Y. Liu, H.D. Cheng, J. Huang, Y. Zhang, X. Tang, J. Tian, An effective non-rigid registration approach for ultrasound image based on “demons” algorithm, *J. Digit. Imaging* 26 (2013) 521–529, <http://dx.doi.org/10.1007/s10278-012-9532-0>.
- [86] S. Klein, M. Staring, J.P.W. Pluim, Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines, *IEEE Trans. Image Process.* 16 (2007) 2879–2890, <http://dx.doi.org/10.1109/TIP.2007.909412>.
- [87] S. Klein, J.P.W. Pluim, M. Staring, M.A. Viergever, Adaptive stochastic gradient descent optimisation for image registration, *Int. J. Comput. Vis.* 81 (2009) 227–239, <http://dx.doi.org/10.1007/s11263-008-0168-y>.
- [88] S. Klein, M. Staring, M.A. Viergever, J. Pluim, Elastix: a toolbox for intensity-based medical image registration, *IEEE Trans. Med. Imaging* 29 (2010) 196–205, <http://dx.doi.org/10.1109/TMI.2009.2035616>.
- [89] M.B.-N. and C. Gramkow, Fast Fluid Registration of Medical Images, in: *Statew. Agric. L. Use Baseline* 2015, vol. 1, 2015. <http://dx.doi.org/10.1017/CBO9781107415324.004>.
- [90] F.L. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 567–585, <http://dx.doi.org/10.1109/34.24792>.
- [91] S. Ourselin, A. Roche, G. Subsol, X. Pennec, N. Ayache, Reconstructing a 3D structure from serial histological sections, *Image Vis. Comput.* 19 (2001) 25–31, [http://dx.doi.org/10.1016/S0262-8856\(00\)00052-4](http://dx.doi.org/10.1016/S0262-8856(00)00052-4).
- [92] X. Yang, H. Akbari, L. Halig, B. Fei, 3D non-rigid registration using surface and local salient features for transrectal ultrasound image-guided prostate biopsy, *Proc. SPIE* 7964 (2011) 79642V, <http://dx.doi.org/10.1117/12.878153>.
- [93] X. Yang, B. Fei, 3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning, *Proc. SPIE* 8316 (2012) 83162O, <http://dx.doi.org/10.1117/12.912188>.
- [94] J. Kybic, M. Unser, Fast parametric elastic image registration, *IEEE Trans. Image Process.* 12 (2003) 1427–1442, <http://dx.doi.org/10.1109/TIP.2003.813139>.

- [95] P. Besl, N. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (1992) 239–256, <http://dx.doi.org/10.1109/34.121791>.
- [96] J. Banerjee, C. Klink, E.D. Peters, W.J. Niessen, A. Moelker, T. Van Walsum, 4D liver ultrasound registration, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 8545 LNCS, 2014, pp. 194–202. [http://dx.doi.org/10.1007/978-3-319-08554-8\\_20](http://dx.doi.org/10.1007/978-3-319-08554-8_20).
- [97] H. Chui, A. Rangarajan, A new point matching algorithm for non-rigid registration, *Comput. Vis. Image Underst.* 89 (2003) 114–141, [http://dx.doi.org/10.1016/S1077-3142\(03\)00009-2](http://dx.doi.org/10.1016/S1077-3142(03)00009-2).
- [98] C. Wachinger, R. Shams, N. Navab, Estimation of acoustic impedance from multiple ultrasound images with application to spatial compounding, in: *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work.* 2008. <http://dx.doi.org/10.1109/CVPRW.2008.4563028>.
- [99] T.S. Yoo, The insight toolkit: An open-source initiative in data segmentation and registration, in: *Vis. Handb.*, 2005, pp. 733–748. <http://dx.doi.org/10.1016/B978-012387582-2/50039-3>.
- [100] S.Y. Sun, M. Gilbertson, B.W. Anthony, Probe localization for freehand 3D ultrasound by tracking skin features, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 8674 LNCS, 2014, pp. 365–372. [http://dx.doi.org/10.1007/978-3-319-10470-6\\_46](http://dx.doi.org/10.1007/978-3-319-10470-6_46).
- [101] H. Rafii-Tari, P. Abolmaesumi, R. Rohling, Panorama ultrasound for guiding epidural anesthesia: A feasibility study, in: *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. 6689 LNCS, 2011, pp. 179–189. [http://dx.doi.org/10.1007/978-3-642-21504-9\\_17](http://dx.doi.org/10.1007/978-3-642-21504-9_17).
- [102] H.A.D. Ashab, V.A. Lessoway, S. Khallaghi, A. Cheng, R. Rohling, P. Abolmaesumi, AREA: An augmented reality system for epidural anaesthesia, *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, 2012, pp. 2659–2663. <http://dx.doi.org/10.1109/EMBC.2012.6346511>.
- [103] K. Rajpoot, V. Grau, J. Alison Noble, H. Becher, C. Szmigielski, The evaluation of single-view and multi-view fusion 3D echocardiography using image-driven segmentation and tracking, *Med. Image Anal.* 15 (2011) 514–528, <http://dx.doi.org/10.1016/j.media.2011.02.007>.