Computer Vision for the Analysis of Cellular Activity

D.Phil Thesis

Visual Geometry Group
Department of Engineering Science
University of Oxford

Supervisor:
Prof. Andrew Zisserman

Amr El-Labban
Magdalen College

2013

Amr El-Labban
Magdalen College

Doctor of Philosophy
Michaelmas Term 2013

Computer Vision for the Analysis of Cellular Activity

Abstract

In the field of cell biology, there is an increasing use of time-lapse data to understand cellular function. Using automated microscopes, large numbers of images can be acquired, delivering videos of cell samples over time. Analysing the images manually is extremely time consuming as there are typically thousands of individual images in any given sequence. Additionally, decisions made by those analysing the images, e.g. labelling a mitotic phase (one of a set of distinct sequential stages of cell division) can be subjective, especially around transition boundaries between phases, leading to inconsistencies in the annotation. There is therefore a need for tools which facilitate automated high-throughput analysis.

In this thesis we develop systems to automatically detect, track and analyse subcellular structures in image sequences to address biological research needs in three areas: (i) Mitotic phase labelling, (ii) Mitotic defect detection, and (iii) Cell volume estimation.

We begin by presenting a system for automated segmentation and mitotic phase labelling using temporal models. This work takes the novel approach of using temporal features evaluated over the whole of the mitotic phases rather than over single frames, thereby capturing the distinctive behaviour over the phases. We compare and contrast three different temporal models: Dynamic Time Warping, Hidden Markov Models, and Semi Markov Models. A new loss function is proposed for the Semi Markov model to make it more robust to inconsistencies in data annotation near transition boundaries.

We then present an approach for detecting subtle chromosome segregation errors in mitosis in embryonic stem cells, targeting two cases: misaligned chromosomes in a metaphase cell, and lagging chromosomes between anaphase cells. We additionally explore an unsupervised approach to detect unusual mitotic occurrences and test its applicability to detecting misaligned metaphase chromosomes.

Finally, we describe a fully automated method, suited to high-throughput analysis, for estimating the volume of spherical mitotic cells based on a learned membrane classifier and a circular Hough transform. We also describe how it is being used further in biological research.

This thesis is submitted to the Department of Engineering Science, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. This thesis is entirely my own work, and except where otherwise stated, describes my own research.

Amr El-Labban, Magdalen College

Acknowledgements

I am eternally grateful to my supervisor, Professor Andrew Zisserman. This thesis would not have been possible without his guidance, expertise and endless enthusiasm and optimism. It has been a pleasure working with him for the past five years.

I would also like to thank my collaborators, Yusuke, Alex, Özlem, and Tony. The countless hours spent generating data (and explaining the biology to me!) were hugely appreciated.

Thanks are due to all members of VGG, past and present, for making it such a fun place to work. Special thanks go to Relja for being a good friend, great sounding board for ideas and problems, and absolutely ruthless pool opponent! Thanks to all the others, including, but not limited to: Patrick, Arpit, Varun, Karen, Yusuf, Max, Carlos, Eric, Victor, Alonso, Mircea, Meelis, Chai, Tomas, Alonso, Elliot, Lubor, Nataraj, Omkar, and Minh. I wish them all the best and hope we stay in touch.

Most importantly, I would like to thank my family for their years of support and understanding.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Motivation

In the field of Cell Biology, there is an increasing use of time-lapse data to understand cellular function. In a standard experiment, a protein is tagged with a fluorescent protein called GFP (green fluorescence protein). This protein can then be followed by time lapse fluorescence microscopy, described in Section 2.2, creating an image sequence. This sequence can be in three dimensions or four dimensions (2 or 3 dimensions of space and one dimension of time). An extra dimension can also be added by tagging multiple proteins and imaging them at the same time. Examples of these kinds of images can be seen in Figure 1.1.

The aim of one area of biological research is to determine the roles of various genes in the process of cell division. This can be achieved using RNA interference

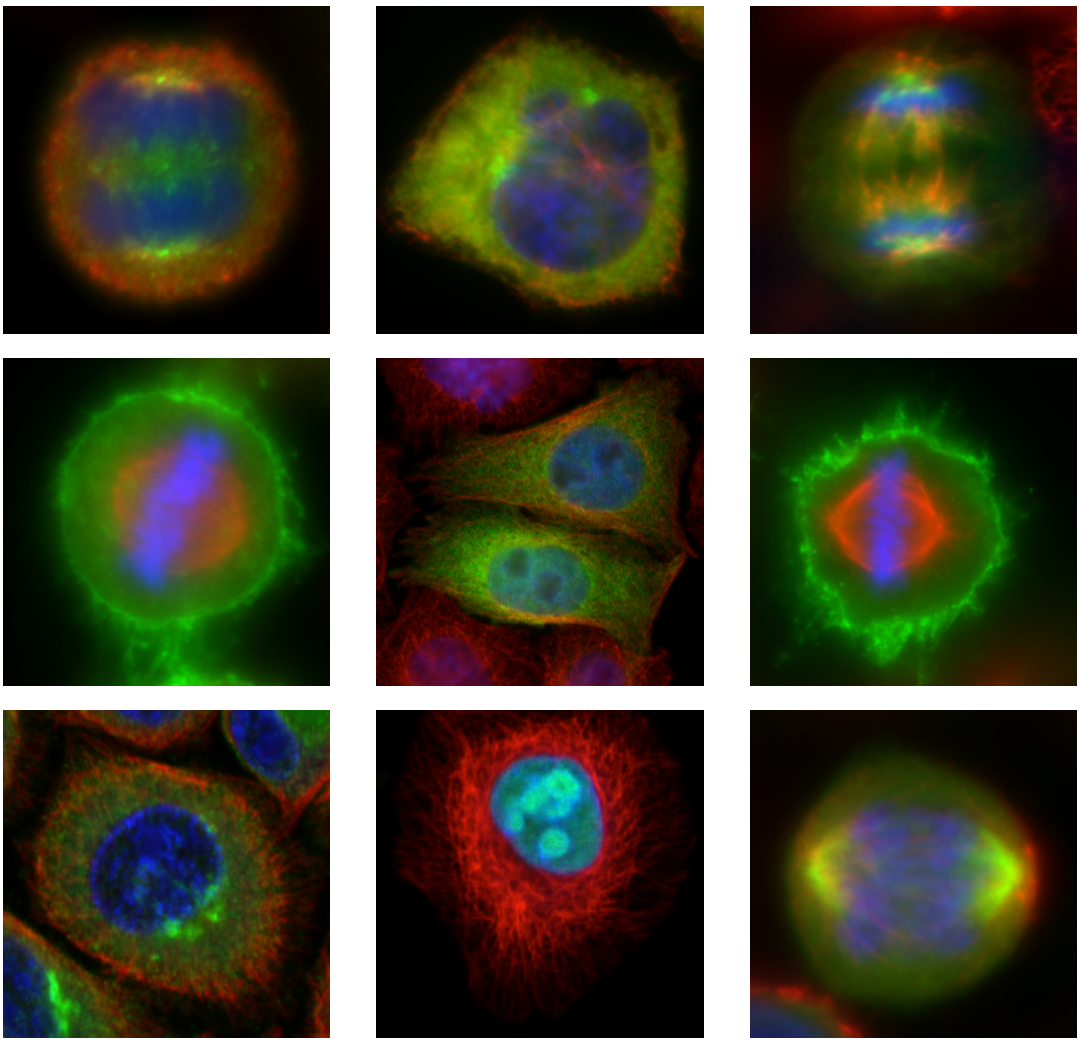
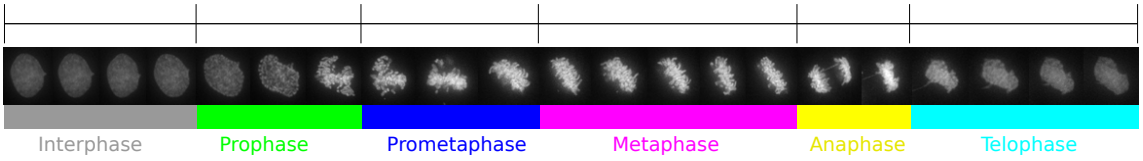
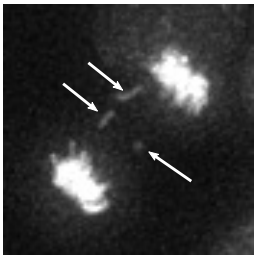


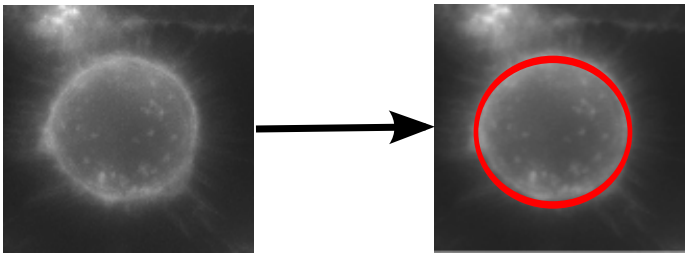
Figure 1.1: Sample fluorescence microscopy images. Images of cells with histone, a protein found in the nucleus, in blue, and tubulin, a protein present in filament-like structures in the cell, in red. Various other proteins are shown in green.



(a) Mitotic Phase Labelling



(b) Mitotic Defect detection



(c) Cell volume estimation

Figure 1.2: Biological image analysis objectives. The challenges adressed in this thesis. (a) Mitotic phase labelling. The objective is to track a cell undergoing mitosis and determine which of the illustrated phases it is in at any given point in time. This is addressed in Chapters 4 and 5. (b) Mitotic defect detection. The objective is to detect small aberrations occurring when cells divide, such as the lagging chromosomes indicated with white arrows. This is addressed in Chapters 6 and 7. (c) Cell volume estimation. The objective is to estimate the cell volume, based on area as illustrated by the red circle, in images of the cell membrane. This is addressed in Chapter 8.

(RNAi) which inhibits the expression of a given gene. By comparing images of RNAi GFP-tagged cells with a control set, the impact of inhibiting the gene on the cells' behaviour, and hence its function, can be determined.

Analysing the images manually is extremely time consuming as there are typically thousands of images in any such experiment. Decisions made by those analysing the images, e.g. choosing a phase of mitosis (which will be described in Chapter 4), can be subjective which can lead to inconsistencies in the annotation.

1.2 Objectives

Our overall objective is to develop a system to automatically detect, track and analyse sub-cellular structures in image sequences to address biological research needs. Specifically, we focus on three distinct problems in biology:

1. **Mitotic phase labelling.** This is useful in itself, detecting delays in cells' progression through mitosis in RNAi experiments, and also as a basis for further analysis, as will be seen in Chapters 6 and 7. Determining mitotic phase is a highly subjective manual task due the gradual change in cell appearance through mitosis necessitating the use of more consistent automated methods.

2. **Mitotic defect detection.** Defects during mitosis can lead to cells gaining or losing chromosomes. This can lead to problems in development, and is a characteristic frequently displayed in cancer cells. Detecting these defects reliably can help the determine the conditions which cause them and potentially lead to ways to prevent their detrimental long-term effects. Defects are typically small, 1-2 pixels in radius, and appear for a short time in huge volumes of data. Methods which can automatically detect them can significantly reduce the time needed for analysis.

3. **Cell volume estimation.** Observing cells' response in regulating their volume after osmotic shock can help determine which genes control cell volume regulation. In microscopy images, the change in a cell's radius after osmotic shock can be as small as 2-3 pixels. Changes of this order of magnitude are difficult to accurately label manually. An automated method for measuring this would provide accuracy and consistency of labelling.

The three objectives are illustrated in Figure 1.2. The biological motivation for these is addressed in greater detail in the relevant chapters of this thesis.

1.3 Thesis Overview

Chapter 2 introduces the basics of cell biology, including the structure of the cell, a description of different stages of mitosis, and an explanation of fluorescence microscopy. This is to aid the reader in understanding the applications in this thesis and the various biological terms used throughout.

Chapter 3 describes existing computer vision approaches to challenges in microscopy images. These include the basic tasks of cell detection, segmentation and tracking, as well as more complex analysis such as cell morphology and cell cycle phase analysis.

In Chapter 4 we compare and contrast three different temporal models for mitotic phase labelling: Dynamic Time Warping, Hidden Markov Models, and Semi Markov Models. A new loss function, robust to human annotation inconsistency, for the Semi Markov model is proposed. We also present a novel method for segmenting images of cell nuclei. The phase labelling methods are tested in Chapter 5 under two different experimental conditions. We also compare performance to one of the methods described in Chapter 3.

Chapter 6 gives a novel approach for detecting subtle chromosome segregation defects in mitosis in embryonic stem cells. This method uses a set of filters to reduce a large number of candidate detections. One of the stages used to eliminate candidates is based on mitotic phase, as classified by the methods of Chapter 4.

An unsupervised approach to defect detection is explored in Chapter 7. The approach taken is to learn a probabilistic appearance model of certain regions of the cell, then retrieve defects as image regions which have low probability under the appearance model.

In Chapter 8 we describe a fully automated method for estimating the volume of spherical mitotic cells, within given bounds, based on a learned membrane classifier and a circular Hough transform. The method is suited to high-throughput analysis, and its performance is evaluated on a set of manually annotated image sequences.

Finally, Chapter 9 summarises the contributions of this thesis and suggests further applications and extensions of the work presented here.

1.4 Contributions

This work was carried out in collaboration with the Hyman Lab, at the Max Planck Institute of Molecular Cell Biology & Genetics. Members of the lab provided all image data used throughout the thesis, as well as ground truth labelling.

Specifically:

- Chapters 4 and 5: Image acquisition and ground truth labelling done by Y. Toyoda. Labelling verified by A. W. Bird.
- Chapters 6 and 7: Image acquisition done by A. W. Bird and Ö. Demir. Labelling done by A. W. Bird.
- Chapter 8: Image acquisition and ground truth labelling done by Y. Toyoda.

Additionally, C. Arteta provided the software implementation of Arteta et al. (2012) used in Chapters 6 and 7, which at the time of publication of El-Labban et al. (2013) was not publicly available.

1.5 Publications

The following publications have resulted from this thesis:

- A. El-Labban, A. Zisserman, Y. Toyoda, A. W. Bird, and A. Hyman. Dynamic time warping for automated cell cycle labelling. In *Microscopic Image Analysis with Applications in Biology*, 2011
- A. El-Labban, A. Zisserman, Y. Toyoda, A. W. Bird, and A. Hyman. Discriminative semi-markov models for automated mitotic phase labelling. In *IEEE International Symposium on Biomedical Imaging*, 2012
- A. El-Labban, C. Arteta, A. Zisserman, A. W. Bird, and A. Hyman. Mitotic phase based detection of chromosome segregation errors in embryonic stem cells. In *IEEE International Symposium on Biomedical Imaging*, 2013

CHAPTER 2

BIOLOGY BACKGROUND

In this chapter we describe the basic background knowledge of cell biology and microscopy. A full description is beyond the scope of this thesis, but a basic understanding is necessary to adequately analyse image sequences of cell mitosis.

2.1 Cell Structure & Cell Cycle

Cells are the fundamental building units of all living organisms. In this section we give a brief description of cell structure and the cell cycle. Two cell types are considered in this thesis: HeLa (human epithelial adenocarcinoma) cells and mouse embryonic stem cells. These share essentially the same sub-cellular structure, undergo mitosis and are imaged in the same way as described in Section 2.2. For more background on the origins of HeLa cells, refer to Skloot (2010).

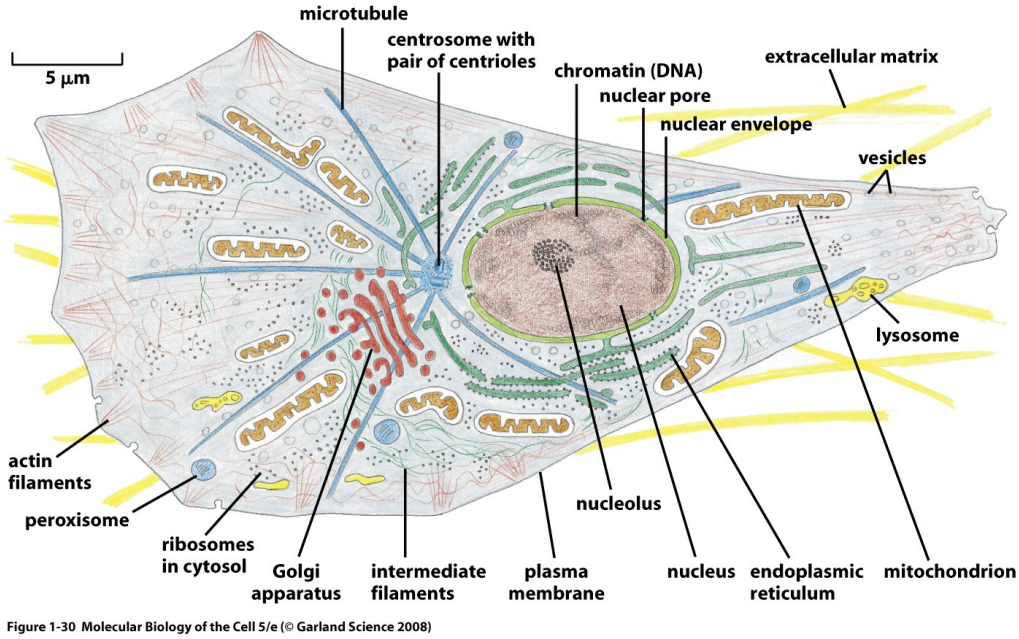


Figure 2.1: Structure of the cell. Relevant structures for this thesis are the nucleus, and plasma membrane. The remainder are beyond the scope. Figure reproduced from Alberts et al. (2009).

Unless otherwise stated, the information in this section is as described by Alberts et al. (2009).

2.1.1 Cell Structure

In this section we describe the structures within cells, starting with those which are relevant to this thesis. These are illustrated in Figure 2.1.

Nucleus. This is the most prominent structure in the cell. The nucleus contains chromatin, which consists of long polymer molecules of DNA which encode genetic information of the organism as well as many binding proteins including histones.

As cells prepare to divide, the chromatin molecules become more compact and are visible under a microscope as individual *chromosomes*.

Plasma membrane. This outer membrane is what contains all structures in the cell. It is permeable to water, and a number of channels facilitate cellular activities by allowing import and export of ions, molecules, and water. The membrane changes shape as the cell undergoes mitosis.

Cytoplasm. Contents of the cell within the plasma membrane, but outside the nucleus.

Actin filament. Protein filament formed from a chain of actin molecules. These are a major constituent of the *cytoskeleton*, a system of protein filaments which gives cells shape and capacity for movement.

Centrosome. Central organelle, containing two *centrioles*, that is the main organising centre for microtubules. Separates to form the poles of the mitotic spindle as described in the following section.

Golgi apparatus. Large organelle where proteins and lipids are modified and sorted for transport to other sites.

2.1.2 Cell Division

For the majority of its life, a cell exists in a state known as *interphase* shown in Figure 2.2(a). In this stage of the cell's life cycle, the cell increases in size and towards the end, the chromosomes are duplicated to prepare for mitosis.

After the interphase, the cell enters *M-phase*, consisting of mitosis followed by cytokinesis, where the cell divides into two daughter cells. Mitosis itself is divided into 5 stages which are all illustrated in Figure 2.2:

1. **Prophase:** Microtubules begin to appear, originating from the centrosomes, the organising centres of the microtubules which are not tagged with GFP in this data. The microtubules join together to form the mitotic spindle. The chromatin starts to condense, forming the mitotic chromosomes. This phase typically lasts for about 30 mins.

2. **Prometaphase:** In this phase, the mitotic spindle has fully formed and chromosomes can attach to the microtubules and are moved by them. This phase typically lasts for about 20 minutes but can be much longer.

3. **Metaphase:** At this point, the chromosomes line up with the equator of the mitotic spindle, midway between the centrosomes at the poles of the mitotic spindle. This phase typically lasts for 10-20 minutes but can be much longer when chromosome alignment is defective.

4. **Anaphase:** The chromosomes now begin to separate, being pulled apart by shortening microtubules and the centrosomes at the poles moving apart. This phase typically lasts for about 10 mins.

5. **Telophase:** During this phase, the chromosomes arrive at the spindle poles and a nuclear envelope forms around them. The chromosomes decondense to form the interphase chromatin. This marks the end of mitosis. This phase typically lasts for 60-120 mins.

After mitosis ends, the cells undergo cytokinesis, whereby the cells separate fully. For this to occur, a contractile ring forms between the two new nuclei. This gradually contracts, pinching the cell into two separate ones. The region in the middle of this ring is called the cleavage furrow.

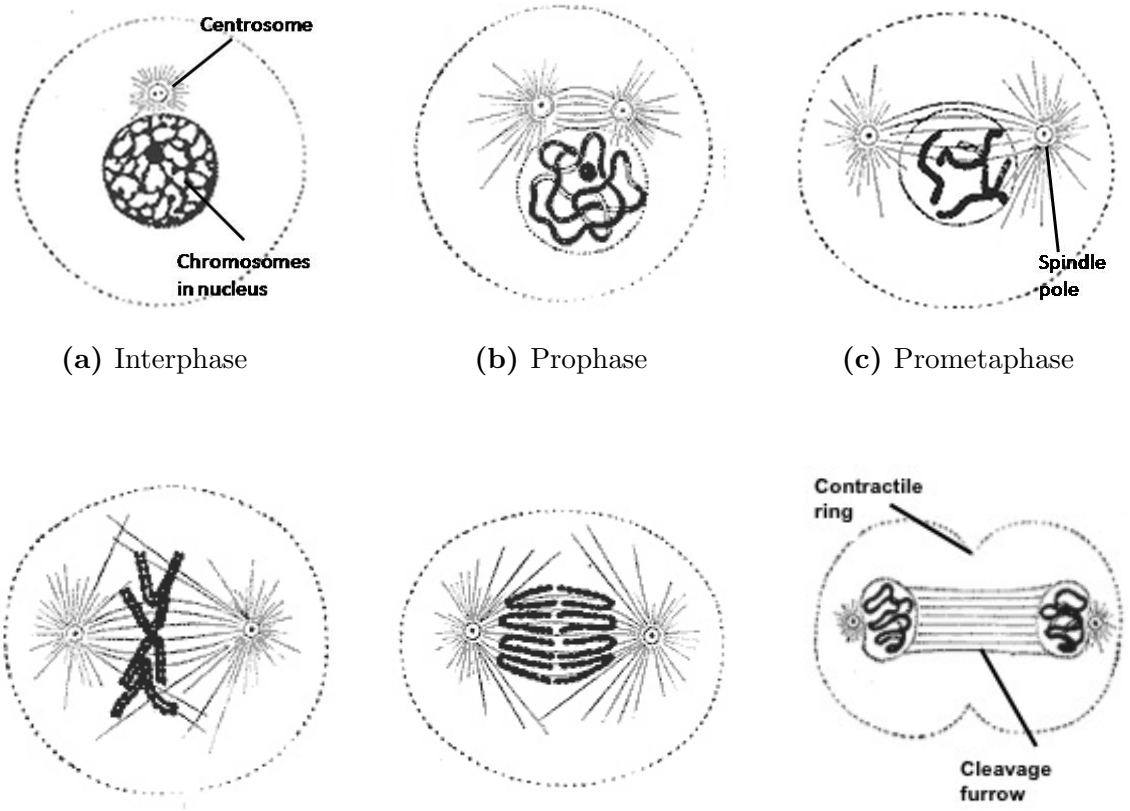
Examples of the appearance of cells in the image sequences analysed in this project are shown in Figure 2.3. Further examples of manually (Figure 4.5) and automatically (Figures 5.1 and 5.2) labelled images of mitotic phases can be seen in later chapters.

2.2 Microscopy Image Acquisition

The images used throughout this thesis are acquired using fluorescence microscopy. The first stage of this approach is to make cells fluorescent. This is either done by staining specific proteins with a fluorescent dye or by genetically modifying cells so that proteins of interest also carry a fluorescent protein. Once this is done, the proteins absorb light at one wavelength and emit it at another.

Cells can then be imaged using an apparatus as seen in Figure 2.4. Light from the source is filtered (excitation filter), with only the wavelength which causes fluorescence passing through to the sample. The emission filter prevents any light reaching the detector (a camera) apart from wavelengths emitted by the specimen. The combination of objective lens and ocular focus the image of the specimen.

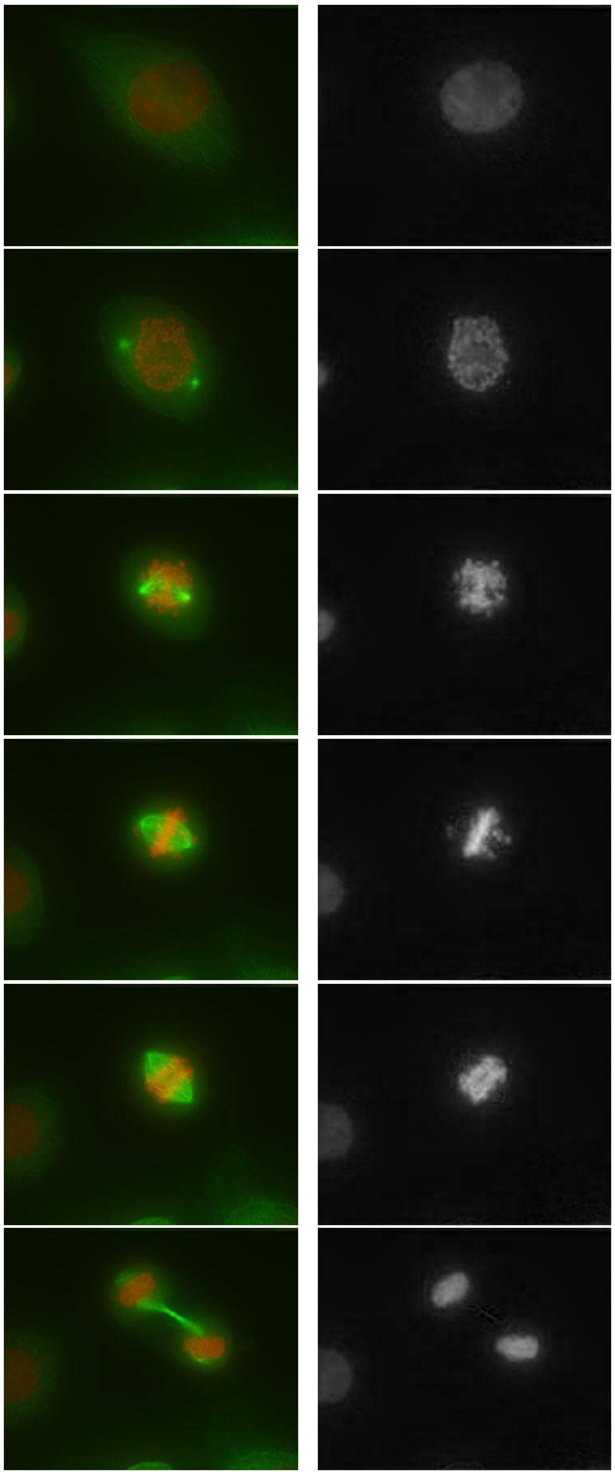
By moving the focal plane through the specimen, and acquiring a 2D image at fixed intervals, a 3D representation of the object can be built. This process results



(a) Interphase **(b)** Prophase **(c)** Prometaphase

(d) Metaphase **(e)** Anaphase **(f)** Telophase/Cytokinesis

Figure 2.2: The Phases of the Cell Cycle (from Gray (1918)). (a) Interphase is the state in which a cell spends the majority of its life. The dark region in the centre represents the chromosomes in the nucleus of the cell. (b)-(f) Progression of a cell through the stages of mitosis.



(a) Interphase. The phase in which cells spend the majority of their time. Nucleus texture is quite smooth (red). Microtubules are spread evenly throughout cell (green).

(b) Prophase. Cell starts to enter mitosis. Chromatin appears more textured as chromosomes become more densely packed (red). Bright spots appear at centrosomes as microtubule ends become denser there (green).

(c) Prometaphase. Chromatin now much more densely packed and textured (red). Mitotic spindle starts to form (green).

(d) Metaphase. Chromosomes organised into a disk-like structure, viewed from the thin edge (red). Mitotic spindle fully formed (green).

(e) Anaphase. Chromosomes being pulled apart look less bright than metaphase as they become less dense (red). Spindle appearance similar to before (green).

(f) Telophase. Chromatin, now separated into two daughter cells, appears very similar to interphase (red). Tubulin now concentrated in cleavage furrow between cells (green).

Figure 2.3: The Phases of the Cell Cycle (as in data used throughout this thesis). Left: Chromatin visible in red and tubulin in green. Right: Chromatin channel (red on the left) shown in grey.

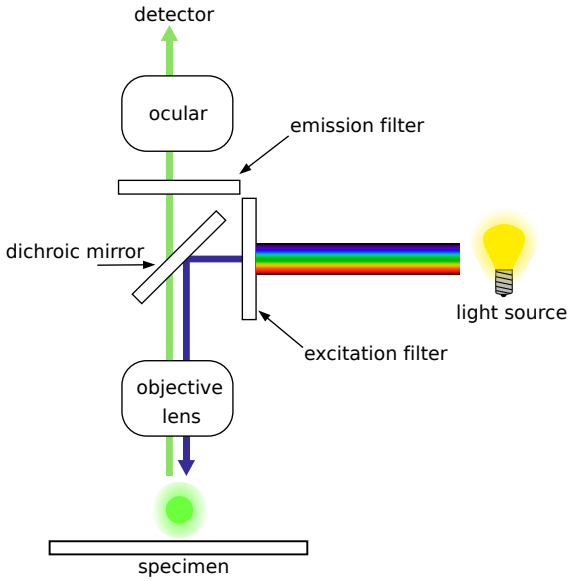
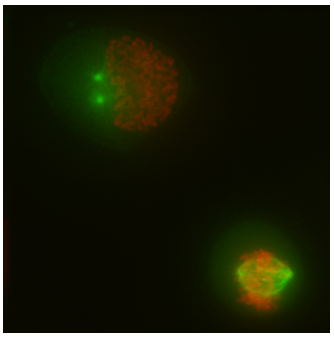


Figure 2.4: Schematic of a fluorescence microscope. Figure modified from Wikipedia (2013).

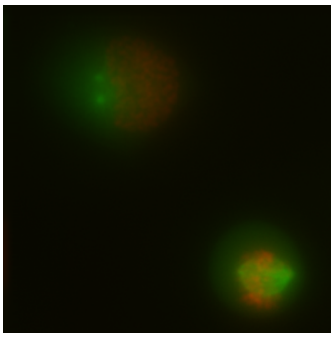
in a large amount of out-of-focus light being captured in each image. In fact, the total integrated fluorescence is the same in each optical section (Shaw, 2006).

The out-of-focus light can be corrected using a process called *deconvolution* (Biggs, 2001). This process assumes that every point in the image has been blurred/convolved with a 3D point spread function (PSF), and attempts to reverse it. The PSF used for deconvolution is either the theoretical one (based on microscope parameters), or is measured by imaging calibration beads of known size and shape. In principle, the convolution with the PSF is a simple, reversible, linear operation. However, given the nature of the specimens being imaged, the PSF can vary through the cells themselves and the medium they are in. This means that even after deconvolution there is still some level of out-of-focus light and noise in images.

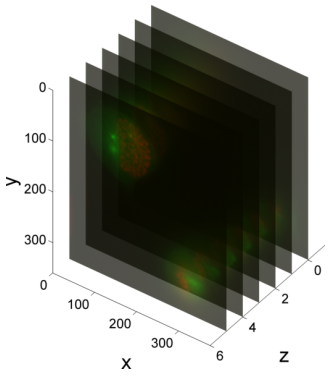
All images used in this thesis have been deconvolved and maximally projected along the z-axis.



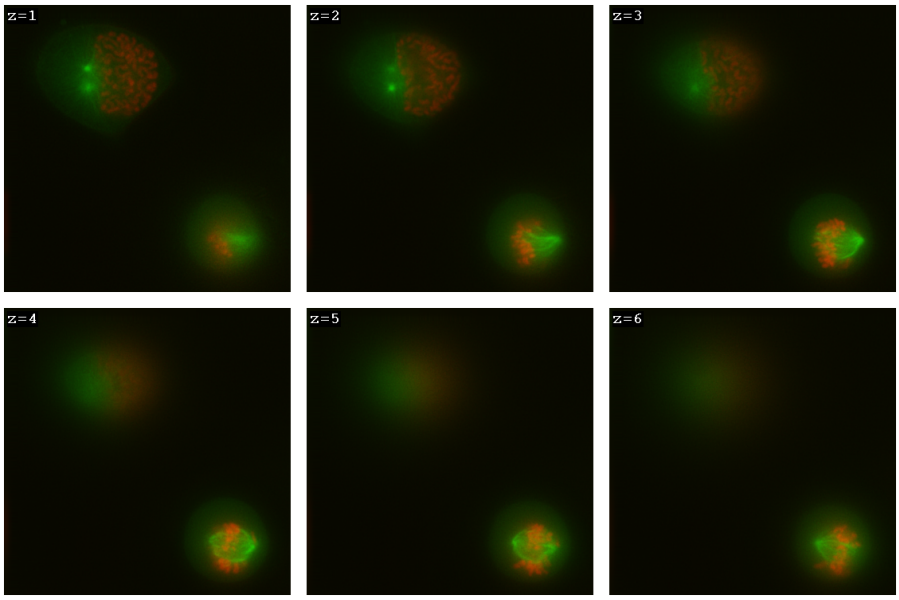
(a) Maximum intensity projection



(b) Mean intensity projection



(c) 3D stack



(d) Individual image montage

Figure 2.5: Different ways of displaying a 3D image stack. (a) A maximum intensity projection along the z-axis of the 3D image. This gives a 2D top-down view of the frame contents by using the maximum pixel intensity out of all z-planes at every (x, y) position in the image. This allows easier visualisation of the image contents than showing the 3D image. (b) Mean intensity projection obtained in a similar way to the max-intensity projection, but using a mean instead. The blurry appearance is due to the contribution of z-planes where the cells are out of focus. (c) The individual 2D images making up the stack. (d) The individual image slices in a montage.

CHAPTER 3

LITERATURE REVIEW

In this chapter we look at three key challenges in microscopy image analysis and review existing literature addressing them: (i) Cell detection and segmentation, (ii) Cell tracking, and (iii) Mitotic phase labelling.

There is a very large number of works in these areas, especially in the areas of detection and tracking. Many of these are developed to cope with challenges specific to a certain combination of cell type, imaging method, and other parameters. In this chapter, rather than focus on the specific adaptations, we aim to look at the broad categories of approaches, illustrated by some representative examples, and assess their relative strengths.

For a more in-depth review of the field, we refer the reader to more comprehensive reviews, such as those of Rittscher (2010); Rittscher et al. (2008).

3.1 Cell Detection & Segmentation

Cell detection is a crucial stage in the analysis of microscopy image analysis as it the foundation for subsequent tasks such as counting, segmentation, tracking or other more complex analyses. However, detection is often a byproduct of segmentation and so both classes of algorithm are grouped here. The segmentation yields additional size, shape, and texture information which can be used for further analysis.

The challenges in detection and segmentation systems are comparable: low signal-to-noise ratios; non-uniform intra-cellular and inter-cellular appearance; presence of image artefacts, either modality induced or due to presence of non-cell particles; cell density and occlusion. Methods for segmentation and detection can be divided into the following categories: (i) Region or blob detector methods, (ii) Thresholding and watershed methods, (iii) Active contour methods, (iv) Machine learning based methods.

3.1.1 Blob detector methods

The simplest approach for detection is to model cells as regions of local maxima or minima of image intensity. A number of works use the popular Laplacian-of-Gaussian blob detector (Al-Kofahi et al., 2010; Peng et al., 2009).

These methods are not flexible enough to capture complex cell shape models, but can be used as a starting point to generate sets of candidates for evaluation by more complex methods.

3.1.2 Threshold and watershed based methods

There is a great deal of work which uses thresholding based segmentation, especially in the context of fluorescence images of cell nuclei (Chen et al., 2006; Conrad and Gerlich, 2010; Harder et al., 2006a,b, 2009; Held et al., 2010; Neumann et al., 2006).

Harder et al. (2006a,b, 2009) use a locally adaptive thresholding approach, where a local threshold is used if intensity variance exceeds a given threshold within a spatial window. Otherwise, a global threshold is used. This approach overcomes image noise and slight variation in intensity between cells. However, this kind of approach cannot segment touching cells.

Wahlby et al. (2004) use a watershed split and merge approach, to better handle cases where cells touch. However, the approach requires a large number of steps. The watershed is seeded using the h-maxima from the intensity image. The watershed is then applied to the gradient image and objects with weak borders between them are then merged to reduce oversegmentation. The next step, to reduce undersegmentation, is to perform a watershed on the distance transform of the merging step. Finally, the merging operation is performed on the output of the previous step. This process is shown in Figure 3.1.

Global thresholding methods such as Otsu (Otsu, 1979) and Ridler-Calvard (Ridler and Calvard, 1978) are compared to adaptive threshold and watershed methods by Coelho et al. (2009) and found to perform poorly. However, while these approaches can be very effective, they make use of a number of heuristic post-processing steps and are generally fine-tuned to the specific application.

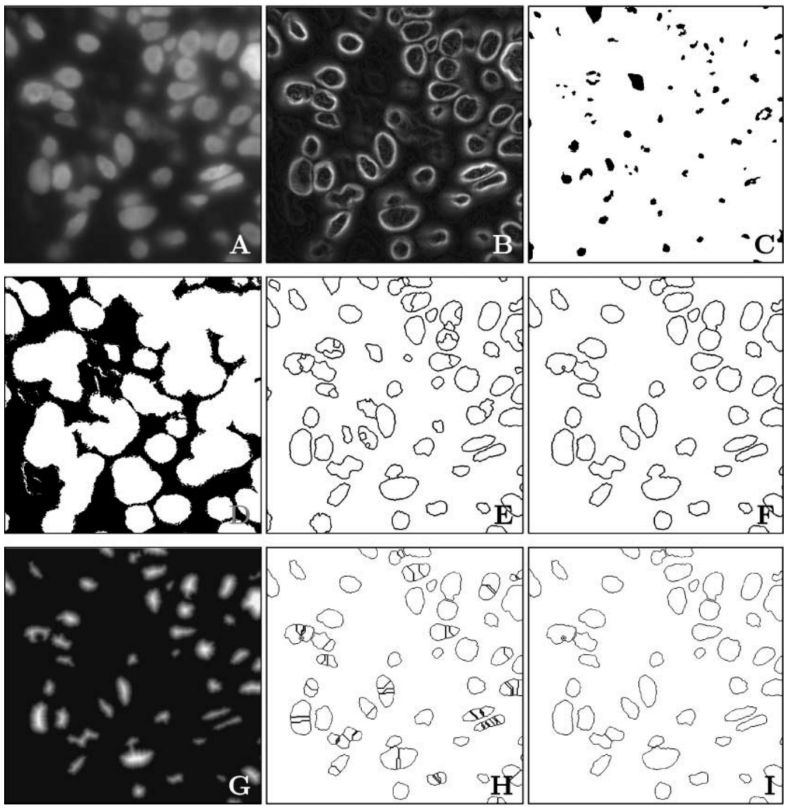


Figure 3.1: Segmentation process of Wahlby et al. (2004). Watershed based segmentation pipeline, also used by Held et al. (2010). (a) Original fluoresence image. (b) Gradient magnitude of (a). (c) Foreground seeds found using h-maxima algorithm on (a). (d) Background seeds from h-maxima of (b). (e) Result of first watershed step. (f) Result after first merging step. (g) Distance transform of (f). (h) Watershed segmentation of (g). (i) Final segmentation from merging step applied to (h).

Padfield et al. (2011) use thresholding and watersheds in wavelet space (Mallat, 1989), which allows them to handle cells at multiple scales and are more robust to image noise.

3.1.3 Active contour based methods

Active contours (Kass et al., 1987) and level-sets (Sethian, 1998) are widely used for segmentation in biomedical imaging in general and have been applied to cellular image segmentation. These methods are advantageous as they can handle images with low-contrast cell boundaries (Ali et al., 2012; Srinivasa et al., 2009) and can incorporate models of cell shape (Chen et al., 2011). They can also naturally extend to a tracking framework as seen in Section 3.2.2. However, contour initialisation can be an issue for active contour-based cell segmentation algorithms, therefore a separate detection algorithm can be necessary.

To handle multiple cells, a number of works use one contour per cell and add constraints to prevent merging e.g. Zimmer and Olivo-Marin (2005). To reduce the computational cost of having a contour for every cell, Nath et al. (2006) introduce a method which can cope with an arbitrary number of cells using only four level-sets.

Bergeest and Rohr (2011) introduce a globally optimal three stage method combining the popular energy of Chan and Vese (2001) with the Bayesian energy of Rousson and Deriche (2002). They extend this (Bergeest and Rohr, 2012) with a two stage method based on the regional-scaling energy of Li et al. (2008a). This approach is effective at separating nearly touching cells. However, the approach relies on all cells in the image having a similar colour distribution and so would not cope with image containing brighter mitotic cells.

3.1.4 Learning based methods

There are a number of methods which use models learned from annotated training data, both for cell detection and segmentation. One way in which supervised learning methods can be used is as a post-processing step, classifying interest regions generated by other methods. Examples include classification of interest regions generated by blob detectors (Pan et al., 2009), thresholding (Mao et al., 2006), watersheds (Marcuzzo et al., 2009), and level-sets (Cheng et al., 2010).

Arteta et al. (2012) propose a three stage method: first proposing a set of candidate cell regions using maximally stable extremal regions (MSER, Matas et al., 2002); second scoring the candidates using a structured output SVM, trained using simple dot annotations; finally dynamic programming is used to select the optimal set of non-overlapping regions from a tree of nested regions, as illustrated in Figure 3.2. This approach is used for mitotic cell detection in Chapter 6, where the detected MSERs are also used as segmentations of the cell nuclei.

Alternatively, supervised methods can directly be used to detect and segment cells, e.g. by classifying every image pixel individually using Bayesian classifiers (Yin et al., 2010), or using logistic regression as we do in Chapter 4. This kind of approach is especially suitable when pixel intensity alone is not sufficient to differentiate cells from the background.

Similarly, segmentation can also be carried out using graphical models, for example Wu and Shah (2010) optimise over a conditional random field (CRF) where the unary term is based on assigning each pixel to a latent topic generated using probabilistic latent semantic analysis (pLSA, Hofmann, 1999).

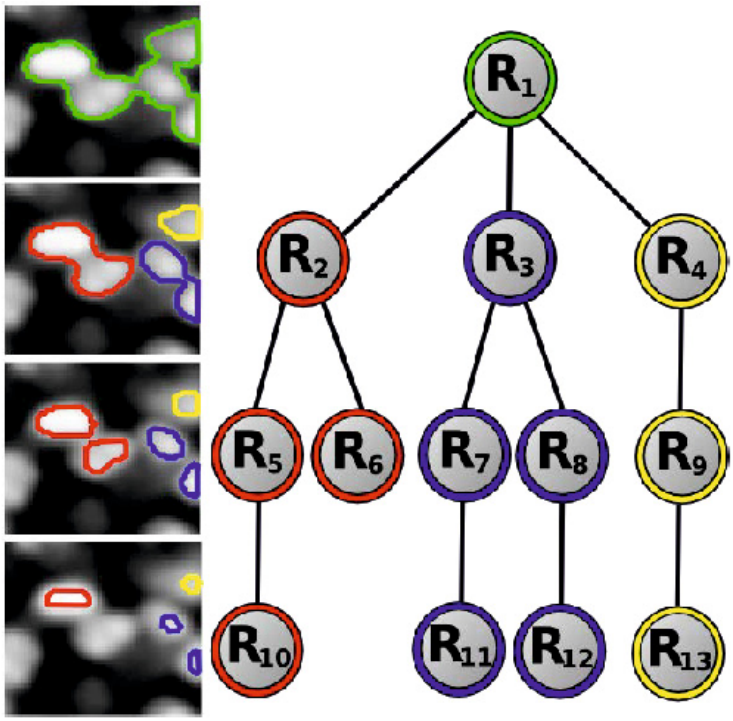


Figure 3.2: Detection approach of Arteta et al. (2012). A set of nested MSERs is found and arranged into a tree. Regions are scored using structured output SVM and the optimal set selected with dynamic programming.

Lou et al. (2012) use a graph cut based approach (Boykov and Jolly, 2001) for nucleus segmentation in fluorescence images. They use a random forest (Breiman, 2001) to generate a probability map for the unary term and use a standard contrast dependent pairwise term. Lou et al. extend the standard model by adding a shape prior which penalises graph cuts which are not orthogonal to a vector field (the Euclidean distance map) from the cell centroids, as illustrated in Figure 3.3. A pre-processing stage is required to obtain the seeds. Finally, structured output learning is used to learn the weighting between the terms in the graph cut formulation. This approach effectively models the shape of interphase nuclei and enforces that segmentations result in appropriately blob-like detections. However, using a more

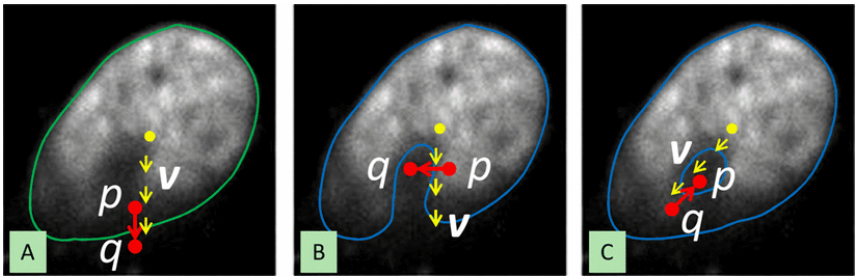


Figure 3.3: Segmentation approach of Lou et al. (2012). A blob-like shape prior is incorporated by aligning cuts (red) with a vector field (yellow). (a) A good cut. (b)-(c) Bad cuts.

simple method without the shape prior could be just as effective with simple post-processing, e.g. the case in Figure 3.3(c) could be solved by adding a hole-filling post-processing step.

3.2 Cell Tracking

Tracking in cell microscopy image sequences presents a number of challenges including multiple object tracking, objects entering and leaving the field of view, occlusion as well as needing to handle splitting objects as cells go through mitosis. Tracking cells becomes a much more challenging problem as cell density increases, e.g. in Figure 6.5. Cell tracking is used in a variety of ways to study cell dynamics, e.g. constructing lineage trees or quantifying cell motility, but the exact methods used vary by application. Approaches to tracking can be broadly divided into three categories: (i) Stochastic filters, (ii) Contour evolution, and (iii) Segmentation/detection and association.

3.2.1 Stochastic Filters

Stochastic filtering approaches, such as Kalman filters (Kalman, 1960), particle filters (Arulampalam et al., 2002; Isard and Blake, 1998) and mean-shift tracking (Comaniciu and Meer, 2002), rely on a motion model which is used to propagate position estimates. These approaches do not need a precise segmentation of the tracked objects, which can be advantageous in datasets which are challenging to segment. However, depending on the complexity of the motion model, they can be very computationally expensive. Additionally, their performance is dependent on being able to determine a reliable model of cell motion.

Debeir et al. (2005) use a basic cell appearance model to locate cell centroids and then update the position of the centroids using mean-shift. Kachouie et al. (2006) use a maximum a posteriori based probabilistic framework to associate hypothesis cells (peaks in a cell detector probability map). Due to the large number of hypotheses (candidate cells and tracks) this method is not very scalable, and so is not suited to datasets with large numbers of cells.

Wang et al. (2007) use a particle filter to track cells segmented with active contours. The observation model is based on the similarity of texture features between the contours in consecutive frames, and the 2D locations of the cells are used as the particles.

Li et al. (2006) use a Kalman filter in conjunction with a level set contour propagation approach. The level set adds robustness by handling discrepancies in the tracks, i.e. allowing for cells appearing/disappearing, mitosis etc. This approach is extended (Li et al., 2008b) using an interacting multiple model (IMM, Blom, 1984)

which runs multiple Kalman filters in parallel, allowing for multiple models of cell motion. IMMs are also utilised by Genovesio et al. (2006).

3.2.2 Contour Evolution Approaches

Contour evolution based tracking methods, e.g. level-set tracking (Kass et al., 1987; Shi and Karl, 2005), are natural extensions of the segmentation approaches described in Section 3.1.3. These use the object boundary curve from the previous frame as the initial estimate for the boundary in the current frame (Padfield et al., 2009; Yang et al., 2005). These methods can struggle with densely packed cells, and Yang et al. (2005) implement a watershed based post-processing step to separate touching cells.

Other methods use a level-set for each cell (Dufour et al., 2005; Dzyubachyk et al., 2010; Zimmer and Olivo-Marin, 2005). The level-sets are coupled with an overlap constraint to prevent cells from being merged. Due to the number of level-sets, this can be computationally expensive.

Although contour-based methods can handle topology changes effectively, they rely on some degree of cell overlap between frames to initialise the contours. Additional heuristics are necessary to reinitialise them in cases of fast cell motion, or cell appearance and disappearance.

3.2.3 Data association methods

Segmentation and association approaches, typically take the output of one of the segmentation or detection methods described in Section 3.1 and find frame-to-frame correspondences between them. This can be done using multiple hypothesis tracking (Blackman, 2004). This algorithm propagates a set of hypotheses combining every possible track up to the current time with every new detection and the most likely ones are used. Fast, globally optimal solutions exist (Amberg and Vetter, 2011; Pirsiavash et al., 2011), but they are still computationally prohibitive for large numbers of cells. Therefore, simple methods which find correspondences between frames are often used (Harder et al., 2009; Held et al., 2010).

These approaches can be very effective, given reliable segmentations/detections. However, as with other methods, additional heuristics are necessary to cope with mitoses, or cells appearing and disappearing.

Padfield et al. (2011) formulate the association problem as finding the minimum-cost flow in a graph. Extra nodes are added to represent cells moving in and out of the field of view or appearing due to a mitotic event. The cost of the graph edges is determined by the euclidean distance and size of cells. Magnusson and Jalden (2012) use a similar approach, but use the Viterbi algorithm to iteratively select the best path through the graph.

Lou and Hamprecht (2011) use a structured learning based approach to score the association hypotheses (see Figure 3.4). This approach is very accurate, but requires annotated tracks to train, and training can be computationally expensive for large datasets.

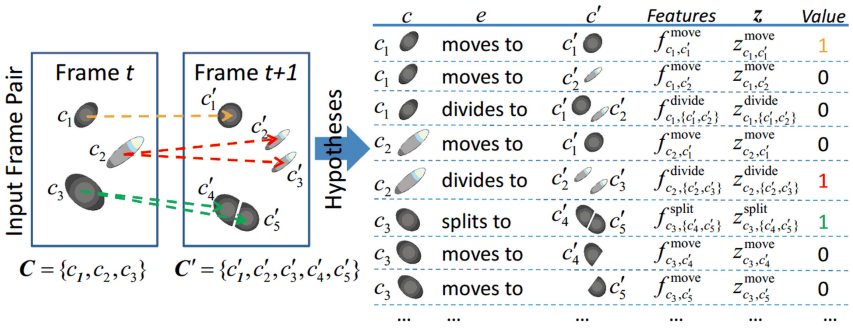


Figure 3.4: Tracking approach of Lou and Hamprecht (2011). Two sets of object candidates, and a small subset of the possible association hypotheses. One particular interpretation of the scene is indicated by coloured arrows (left) or equivalently by a configuration of binary indicator variables z (rightmost column in table).

3.3 Mitotic Phase Labelling

Automatic mitotic phase labelling of tracked cells is carried out in a number of works (Chen et al., 2006; Harder et al., 2006a, 2009; Held et al., 2010). This is restricted to a very general *interphase/mitosis/death/clustered nuclei* classification in Harder et al. (2006b) and extended to classify individual mitotic phases in Harder et al. (2006a, 2009), and Held et al. (2010).

All works follow the same overall framework for classification. After segmentation, shape, size and texture features are computed for each detected cell and classified into phases using a nearest neighbour classifier (Chen et al., 2006) or a support vector machine (SVM) with a Radial Basis Function kernel (Harder et al., 2006a, 2009; Held et al., 2010). The texture features used are again common to all of these works and are a combination of Haralick texture features (Haralick, 1979) and statistical geometric features (Walker and Jackway, 1996) along with some simple measures of gray-level intensity (max, min, mean, variance).

In Harder et al. (2009) these appearance features are complemented with dynamic measures computed as the difference of six features (object size, mean intensity, standard deviation of intensity, circularity, minimum and maximum diameters) between the current frame and the previous and successive ones, thus incorporating temporal information to the feature vector.

The total number of features used is quite large, with Held et al. (2010) using 186 features and Harder et al. (2009) using 376. The large number of features is necessary, as these methods focus on classifying frames independently, rather than using information from the whole track. Even the temporal information included by Harder et al. only gives a local temporal measure of feature changes.

In Chen et al. (2006) and Harder et al. (2009) classification output is corrected based on prior biological knowledge. Chen et al. (2006) use a set of pre-defined rules, only allowing certain phase transitions and durations. Harder et al. (2009) use a finite state machine which models plausible progression through mitosis, and enters error states, which are retrospectively corrected, when illegal phase transitions are encountered. Both approaches are heuristic to some extent, relying on a model which must be defined by the user, rather than being learnt from the data. Held et al. (2010) use a Hidden Markov Model (HMM, Rabiner (1989)) to correct the classification output of the SVM. In this model, the true phases of the cells are the hidden states, and the SVM output labels are the discrete observations. The observation emission probabilities are derived from the SVM training confusion matrix. Transition probabilities and initial state probabilities are learned at test time based on SVM classification results on the test tracks.

Despite the use of temporal information in these methods, the focus is on making very accurate classifications for individual frames, then correcting errors using temporal information. We compare our phase labelling approaches, introduced in Chapter 4, which make much greater use of temporal information,to that of Held et al. (2010) in Chapter 5.

A more recent method by Zhong et al. (2012) performs the task in an unsupervised fashion. This is done by first performing a binary clustering of detected cells in a track to obtain a mitotic/non-mitotic classification. These sub tracks are then individually clustered, using a method called temporally constrained combinatorial clustering (TC3). This clustering is then used as an initialisation for a Gaussian mixture model (GMM) and finally, an HMM is fit to the resulting assignments. This is an interesting approach as it requires no human annotation, and can potentially discover unusual cell appearances that would not have been annotated separately by humans, e.g. variations of appearance within mitotic phases.

3.4 Software packages

Many software packages exist for biological image analysis, here we list a few key ones. *CellProfiler*[1] (Carpenter et al., 2006) is a popular open-source software package, providing tools to carry out cell segmentation, feature extraction and classification, as used by Jones et al. (2009), to differentiate between different cell populations, for example. *Fiji*[2] (Schindelin et al., 2012) is another open-source platform,

[1] http://www.cellprofiler.org/
[2] http://fiji.sc/

which uses user-contributed plugins to perform various image analysis tasks such as segmentation and tracking.

The two commercial products we are aware of are *Metamorph*[3] and *Image-Pro Plus*[4]. Both provide tools for segmentation, tracking and classification, as well as various other functionality such as controlling automated image acquisition.

3.5 Summary

In this chapter we review the general approaches to cell segmentation, tracking and phase labelling. We compare the general categories of approach to these, illustrated by representative examples.

The most simple approach to detection is the blob detector approach, which is limited to detecting cells as intensity peaks in an image and does give precise segmentations. Threshold based approaches are simple, but need additional heuristic steps to cope with touching cells and intra-cell variation. Contour based methods can be effective in low contrast images, but are dependent on initialisation and can be expensive for large numbers of cells. Learning based methods are potentially the most powerful, but do not necessarily add sufficient improvement to justify the increased computational cost.

We described three types of approach to tracking, illustrated by representative examples. Contour evolution approaches are the best equipped to handle mitosis events (as a topology change), but struggle when cells do not overlap from frame

[3] http://www.moleculardevices.com/Products/Software/Meta-Imaging-Series/MetaMorph.html

[4] http://www.mediacy.com/index.aspx?page=Image_Pro_Software

to frame, which may occur with fast moving cells or image sequences with low temporal sampling. Stochastic filter based approaches require a motion model and can be very computationally intensive. Detection and association based methods can be the simplest, but require additional heuristics to cope with mitoses.

The key works carrying out mitotic phase labelling follow the broadly the same approach of making individual frame based classification, followed by error correction using temporal information. Many approaches use a user-defined model of mitotic progression, involving heuristic rules, rather than a model learned from data.

CHAPTER 4

TEMPORAL MODELS FOR MITOTIC PHASE LABELLING

In this chapter we present a system for automated segmentation and mitotic phase labelling using temporal models. This work takes the novel approach of using temporal features evaluated over the whole of the mitotic phases rather than over single frames, thereby capturing the distinctive behaviour over the phases. We compare and contrast three different temporal models: Dynamic Time Warping, Hidden Markov Models, and Semi Markov Models. A new loss function is proposed for the Semi Markov model to make it more robust to inconsistencies in data annotation near transition boundaries. The models are tested in the next chapter. We also introduce a novel method for segmentation of images of cells tagged with fluorescent chromatin marker. Finally we give simple methods to track cells and classify

tracks as either mitotic or not to give the data used for phase labelling experiments in Chapter 5.

4.1 Introduction

4.1.1 Biological Motivation

The objective of this chapter is to automatically identify and track the individual cells throughout a time-lapse sequence, and to label the mitotic phase for each cell at every time point. The image sequences we use are of HeLa cells, fluorescently tagged to show chromatin. A detailed description of the dataset and the acquistion method can be found in Section 5.1. Some example frames from such a sequence can be seen in Figure 4.1. Over large volumes of data, such labellings can be used to derive statistics on cellular function e.g. how depleting certain proteins from cells affects their progression through mitosis. As will be seen in Chapter 6, knowledge of mitotic phase can also be used as a basis for further analysis.

Related works, described in greater detail in Section 3.3, follow a three stage approach of first segmenting and tracking cells, then labelling the mitotic phase at each time point independently, and finally smoothing the result of the labelling using a Hidden Markov Model (HMM) for example. We follow a two stage approach. The first is similar to previous methods, segmenting and tracking cells. However, the key novelty in this work is that, instead of attempting to label the mitotic phase for each frame independently and then smoothing the result, the phases at every time point are jointly labelled using temporal signals of features.

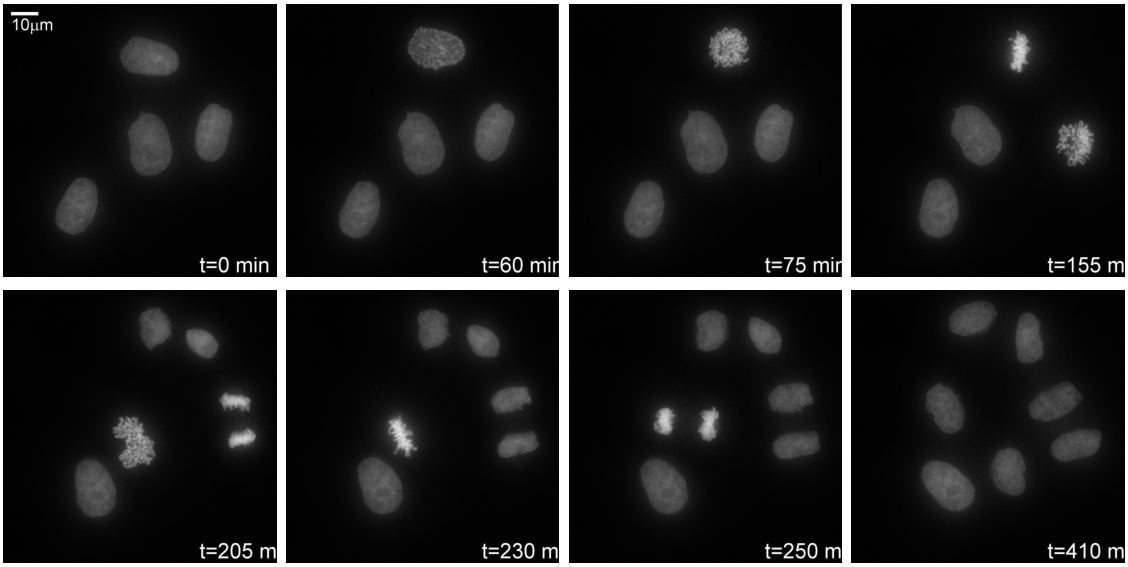


Figure 4.1: Typical time frames from one sequence in the MitoPhase dataset. Each frame is a maximum projection across the z-stack of 3D planes. Frame order: left to right, top to bottom. Three cells can be seen undergoing mitosis over the course of the sequence.

This is possible due to the distinctive behaviour of signals such as the maximum pixel intensity of cells over time, as seen in Figure 4.2. This arises from the change in chromatin density as cells go through mitosis.

4.1.2 Related work

While other phase labelling works are reviewed in Section 3.3, our approaches to solving the phase labelling problem are inspired by approaches used in other areas of signal processing and computer vision, specifically, in the fields of speech and human action recognition. While a full review of these areas is far beyond the scope of this chapter, we briefly mention selected works which have used methods similar to those used here. For a more detailed information we refer the reader to reviews

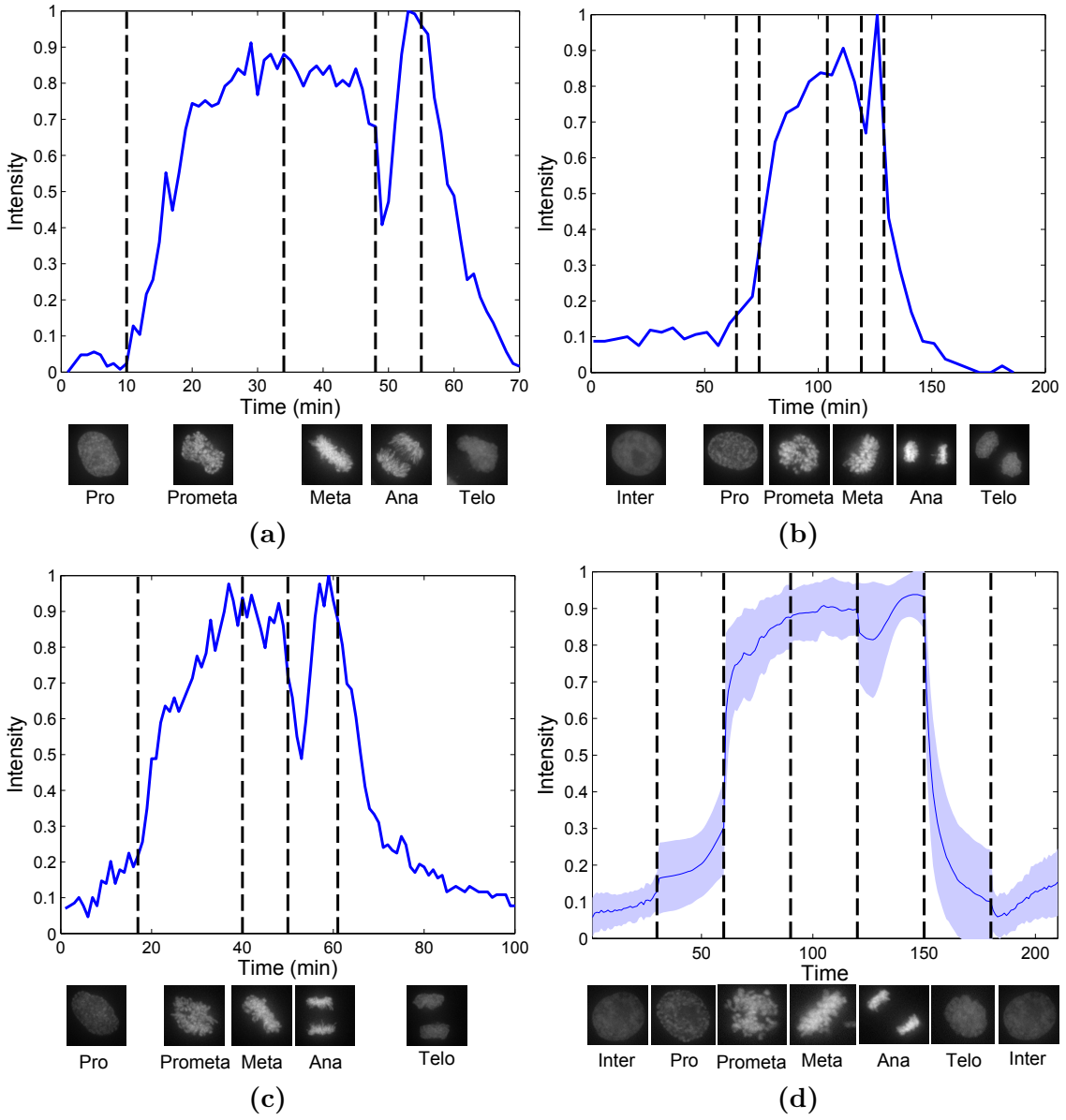


Figure 4.2: Sample temporal signals. (a)-(c) Maximum pixel intensity over time for three example cells from the dataset. Vertical black lines indicate transitions through mitosis. Corresponding phases of mitosis are shown below each image. (d) Maximum intensity signal averaged over the training data set. Signals are resampled such that interphase and each of the five mitotic phases has the same duration before averaging. The shaded region represents ±1 standard deviation.

such as those of Moe (2006); Poppe (2010) (action recognition) or Benzeghiba et al. (2007) (speech recognition). The inspiration comes from the similarity of the simple temporal signals we use here, as seen in Figure 4.2, to audio signals used in the speech recognition domain.

Dynamic time warping (DTW, Sakoe (1978)) has been used extensively in speech recognition, giving a distance between sequences, taking into account a temporal warping which best aligns the sequences. The approach has also been applied to visual features for human action recognition, e.g. by Veeraraghavan et al. (2005) who use DTW to compare sequences of shape features for gait recognition.

Similarly, the early applications of Hidden Markov Models (HMM, Rabiner (1989)) were in speech recognition, although they have since been applied in various ways to human action recognition. Approaches range from the very simple, for example Brand and Kettnaker (2000) simply fit an HMM to simple set of shape features, to the much more complex, e.g. İkizler and Forsyth (2008) fit an HMM per-limb and combine them to label activities. Our approach is more similar to the simple method, as the features we use are sufficiently simple that there is no clear need to use multiple HMMs.

More recent works have explored a discriminative approach to modelling temporal sequences, making use of the idea introduced by Altun et al. (2003) of Hidden Markov SVMs, and extended by Rätsch and Sonnenburg (2006) to Hidden Semi-Markov SVM. Shi et al. (2011) apply these methods to human action recognition, comparing them to HMMs and SVMs, with the Semi-Markov model shown to be the best. Shi et al. also apply the model to text segmentation (Shi et al., 2007).

Chapter Outline

In Section 4.2 we describe methods for labelling mitotic phases given cell tracks. We explore three approaches to doing this: first, by aligning temporal signals of simple features directly to a reference signal (learned from a training set of annotated cells) using Dynamic Time Warping (DTW); second, using an HMM directly on the raw feature signals; finally, using a discriminative Semi-Markov Model (SMM) with features evaluated over temporal segments of cell tracks. These approaches are all able to capture the distinctive dynamic behaviour of features during mitotic phases, as illustrated in Figure 4.2. They also immediately produce segments with contiguous labellings without the need for the two-stage approach of individual frame labelling smoothed with an HMM.

We describe the segmentation method in Section 4.3 and qualitatively compare it to related methods. Section 4.4 describes the tracking method used, and Section 4.5 describes the classifier used to detect candidate mitotic tracks for the phase labelling. Performance of the phase labelling methods is addressed in Chapter 5.

Parts of this work are published in El-Labban et al. (2011, 2012). In (El-Labban et al., 2011) we describe the DTW and HMM approaches, and extend them to the SMM approach in (El-Labban et al., 2012). In this chapter, in addition to more detailed explanation of the methods used, we include more extensive experiments comparing the performance of our methods to that of Held et al. (2010).

4.2 Temporal Models

In this section we describe three methods to perform mitotic phase labelling. All methods assume a given set of temporal feature signals for a tracked cell, such as maximum pixel intensity as shown in Figure 4.2. Methods for cell segmentation and tracking, and feature signal extraction are given in Sections 4.3, 4.4 and 4.6 respectively. The label set is shown in Figure 4.2 and covers the phases: *interphase, prophase, prometaphase, metaphase, anaphase,* and *telophase.*

4.2.1 Dynamic Time Warping

Our first approach to labelling frames in a novel track into stages of the cell cycle consists of aligning a track signal to a reference signal with known transition times using dynamic time warping (DTW) (Sakoe, 1978). DTW optimally warps (stretches) one sequence onto another to minimise an alignment cost under a set of allowed deformations which consist of repeating or removing samples in either the novel signal or the reference signal.

For a training set of feature signals, the reference signal is given by a mean signal, $\mu[j]$, of length J and a corresponding variance signal, $\sigma^2[j]$, where j indicates the temporal frame in the sequence. Before taking the mean, the individual signals are resampled such that each cell cycle phase has a fixed duration, thereby aligning the stages and fixing the transition boundaries, as shown in Figure 4.2(d).

The first step of DTW is to compute a cost matrix. Given a reference signal for a single feature f, a test signal $t[i]$ of length I and a reference signal, this is a $I \times J$

matrix with elements chosen as:

$$C(i,j)_f = \frac{(t[i] - \mu[j])^2}{\sigma^2[j]} \tag{4.1}$$

This cost is based on the assumption that feature values are Gaussian distributed, and that there is a sufficient difference between these distributions between phases. As can be seen in Figure 4.2(d), the maximum pixel intensity does seem to follow this distribution, and the standard deviations are sufficiently small such that this feature alone would be enough to distinguish between some phases.

For multiple features, the cost matrix is computed as:

$$C = \sum_f C_f \tag{4.2}$$

Dynamic programming is then used to find the minimum cost path across the matrix which passes through every point in the test signal, but can start and end anywhere on the reference. An example of aligning the sequence of Figure 4.2 is shown in Figure 4.3. The time complexity of the DTW alignment is $O(IJn)$, where I and J are the lengths of the signals and n is the number of allowed moves in the alignment. In practice n is at least an order of magnitude smaller than I or J.

For DTW, a reference mean signal and standard deviation similar to the one in Figure 4.2(d) is computed for each feature from signals in the training set. To account for the different durations of phases, before averaging, each signal is resampled so that each phase has a fixed, known duration (the mean duration of that phase in the training set).

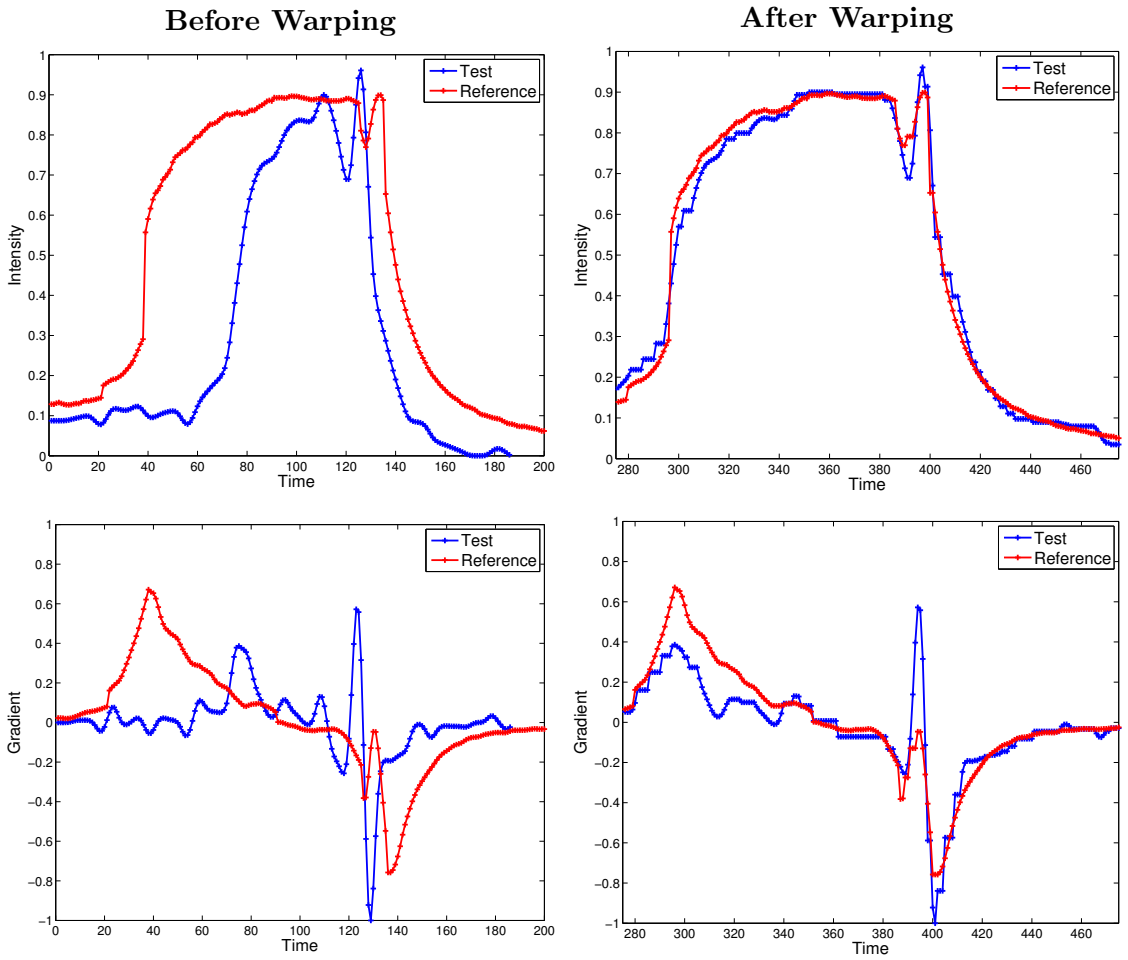


Figure 4.3: Before and after DTW. Samples from the test sequence (in blue) are repeated to obtain a signal which matches the reference (in red). Top row: maximum pixel intensity. Bottom row: maximum intensity gradient.

4.2.2 Hidden Markov Model

Dynamic Time Warping can be thought of as a special case of an HMM, with each point along the reference signal representing a hidden state and transition probabilities restricted to prevent backward moves in time. With this in mind we extend our DTW approach to an HMM.

To do this, the reference signal is discretised into a number of states, representing the mitotic phases. However, due to the simplicity of the features used, an extra step is needed to account for intra-phase variation. This is done by discretising each phase into a number of shorter sub-phases and computing the mean and standard deviation for each feature within these.

All phases apart from interphase and anaphase are divided into two sub-phases to capture the early and late behaviour of the signals over the phase. Interphase is not divided as there is little variation in the features, while anaphase is divided into four as there is significant fluctuation in the features, as seen in Figures 4.2 and 4.4. In total, thirteen sub-phases are used. This is illustrated in Figure 4.4.

Emission probabilities can then be given as:

$$P(i,k) = \exp\left(-\frac{(t[i]-\mu[k])^2}{2\sigma^2[k]}\right) \tag{4.3}$$

where, as in Equation (4.1), $t[i]$ represents the value of the feature signal at point i, and $\mu[k]$ and $\sigma[k]$ represent the mean and standard deviation of sub-phase k respectively.

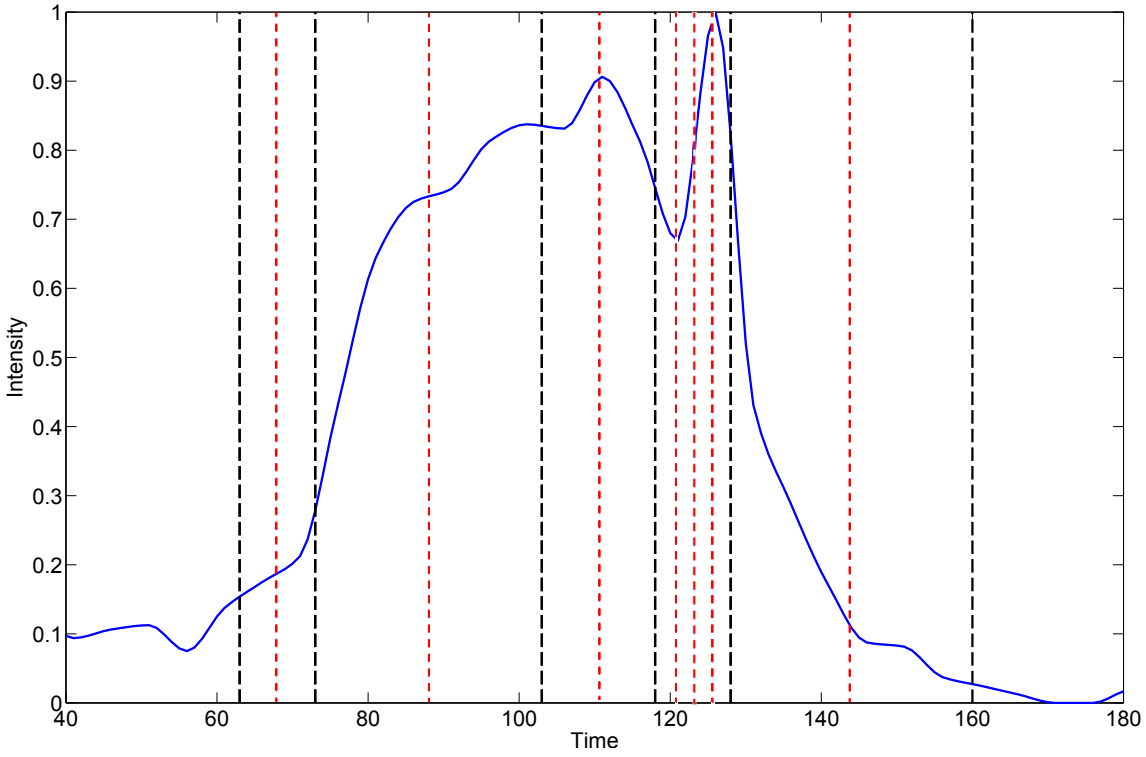


Figure 4.4: HMM sub-phases. Black vertical lines represent phase transitions. To account for intra-phase variation, phases are sub-divided for use in the HMM (red vertical lines represent sub-phase divisions). Interphase (first and last phase shown in figure) has little variation and so is not divided. Anaphase exhibits the greatest variation, and so is divided into four sub-phases. Other phases are divided into two, represent early and late behaviour in the phase.

This approach has a time complexity of $O(IY^2)$ where Y is the number of possible states (the sub-phases). As this is significantly smaller than the length of the reference signal in DTW, the computational cost of the HMM is significantly lower.

4.2.3 Semi-Markov Model

A Semi-Markov model (SMM) (Ostendorf et al., 1996; Shi et al., 2011) is an extension of a HMM, allowing the underlying process to be a semi-Markov chain with each state having a variable duration. By exploiting segment-level features, rather than just frame-level, this model overcomes the duration modelling limitation of HMMs, where the probability of remaining in a given state decays exponentially with time.

For our purposes, we use the wide margin discriminative SMM model of Shi et al. (2011) which, for a given sequence $\mathbf{x}$, aims to find an optimal segmentation, into temporal segments of frames, $\mathbf{s} = [s_1 \dots s_n]$, given by: $s_j = [a_j, b_j, y_j]$ where a_j and b_j are the start and end positions of the segment respectively and $y_j \in [1 \dots 6]$ is the label within the segment, interphase or one of the 5 mitotic phases.

Given the model parameter $\mathbf{w}$ and feature mapping $\Phi(\mathbf{x}, \mathbf{s})$, the maximum likelihood optimal segmentation is then the one given by:

$$\mathbf{s}^* = \arg\max_{\mathbf{s}} \langle \mathbf{w}, \Phi(\mathbf{x}, \mathbf{s}) \rangle \tag{4.4}$$

where $\langle \mathbf{w}, \Phi(\mathbf{x}, \mathbf{s}) \rangle$ represents the inner product between the model parameter and the feature mapping, which can be solved efficiently using a Viterbi-like algorithm.

Learning the model parameter $\mathbf{w}$ from a training set of manually labelled sequences uses the same formulation as for learning a structured output SVM (Tsochantaridis et al., 2005). This is done by solving the regularised optimisation problem:

$$\min_{\mathbf{w}} \frac{||\mathbf{w}||^2}{2} + C \sum \xi_i \tag{4.5}$$

$$\text{s.t. } \forall \mathbf{s}, \langle \mathbf{w}, \Delta\Phi(\mathbf{x}, \mathbf{s}_i) \rangle \geq \Delta(\mathbf{s}_i, \mathbf{s}) - \xi_i \tag{4.6}$$

where the i subscript indicates ground truth segmentation, ξ_i are slack variables allowing some violations of the constraints, C is the regularisation parameter controlling the degree of constraint violation, set using cross-validation, $\Delta(\mathbf{s}_i, \mathbf{s})$ is the label loss between ground truth and predicted sequences, discussed below, and

$$\Delta\Phi(\mathbf{x}, \mathbf{s}_i) = \Phi(\mathbf{x}, \mathbf{s}_i) - \Phi(\mathbf{x}, \mathbf{s}) \tag{4.7}$$

The set of constraints is over all possible segmentations $\mathbf{s}$.

4.2.3.1 Loss functions

The label loss of Shi et al. (2011) is given by:

$$\Delta(\mathbf{s}_i, \mathbf{s}) = |\mathbf{s}| + |\mathbf{s}_i| - 2|\mathbf{s} \cap \mathbf{s}_i| \tag{4.8}$$

where $|\mathbf{s}|$ represents the number of segments in the predited segmentation, $|\mathbf{s}_i|$ is the number of segments in the ground truth segmentation, and $|\mathbf{s} \cap \mathbf{s}_i|$ is the number of segments which exactly match, based on start and end frames, and phase label.

Therefore, the loss, $\Delta(\mathbf{s}_i, \mathbf{s})$, is the total number of segments in $\mathbf{s}_i$ which do not have an exact match in $\mathbf{s}$ and vice versa.

We modify this to be a frame-based loss rather than a segment based one. Each segmentation, $\mathbf{s}$, is expanded into a label vector, $\mathbf{y}$, with length equal to the number of frames in the sequence containing the label at every frame. The mapping $\mathbf{s} \mapsto \mathbf{y}$ is defined as:

$$\begin{bmatrix} \overbrace{a_1}^{s_1} & \overbrace{a_2}^{s_2} & \cdots & \overbrace{a_n}^{s_n} \\ b_1 & b_2 & \dots & b_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix} \mapsto \left[\begin{array}{ccccc} \overbrace{y_1 \dots y_1}^{s_1} & \overbrace{y_2 \dots y_2}^{s_2} & \cdots & \overbrace{y_n \dots y_n}^{s_n} \\ \longleftrightarrow & \longleftrightarrow & & \longleftrightarrow \\ b_1 - a_1 & b_2 - a_2 & & b_n - a_n \end{array} \right] \tag{4.9}$$

As a loss, we use the number of misclassified frames, ignoring errors which are within 2 frames of the ground truth, effectively a slackened Hamming loss:

$$\Delta(\mathbf{s}_i, \mathbf{s}) = \sum_t \Delta(\mathbf{y}_i[t], \mathbf{y}[t]) \tag{4.10}$$

where:

$$\Delta(\mathbf{y}_i[t], \mathbf{y}[t]) = \begin{cases} 0 & \mathbf{y}[t] \in \mathbf{y}_i[t-2:t+2] \\ 1 & \mathbf{y}[t] \notin \mathbf{y}_i[t-2:t+2] \end{cases} \tag{4.11}$$

where $\mathbf{y}$ and $\mathbf{y}_i$ denote the predicted and ground truth labellings respectvely and the t index denotes the frame number.

Allowing this slack helps to overcome inconsistencies in the ground truth annotation which can arise due to the very gradual changes in appearance of the chromatin

at some transitions. We train and test models using both loss functions and the performance improvement our loss function gives is shown in Section 5.3.

The time complexity of the inference algorithm is $O(IY^2D)$ where I and Y are the same as in the HMM case and D is the maximum allowed duration of any phase. In practice this is allowed to be the duration of the signal, giving a time complexity of $O(I^2Y^2)$.

4.2.3.2 Feature mapping

The feature mapping for a segment s_j, $\Phi(\mathbf{x}, s_j)$, is formed from the concatenation of the following three feature vectors, which are illustrated in Figure 4.5:

(i) **Segment level feature vector.** This captures the characteristics of the signal over the segment as a whole. $\phi_1(\mathbf{x}, a_j, b_j, y_j) = f_1(\mathbf{x}, a_j, b_j) \otimes y_j$, where $\otimes y_j$ represents the tensor product with a vector of length equal to the number of classes and with the y_jth element set to one and the rest to zero (as used in Tsochantaridis et al. (2005)). $f_1(\mathbf{x}, a_j, b_j)$ consists of the signal vector at 5 evenly spaced points taken from the segment of $\mathbf{x}$ from a_j to b_j, linearly interpolating where necessary.

(ii) **Segment boundary feature vector.** This captures information about the start of a segment, obtained from a five frame local window centered on the boundary frame a_j. The feature vector is given by the mean of the signal vectors in the window. As with the segment level feature vector, the tensor product with the label vector is taken.

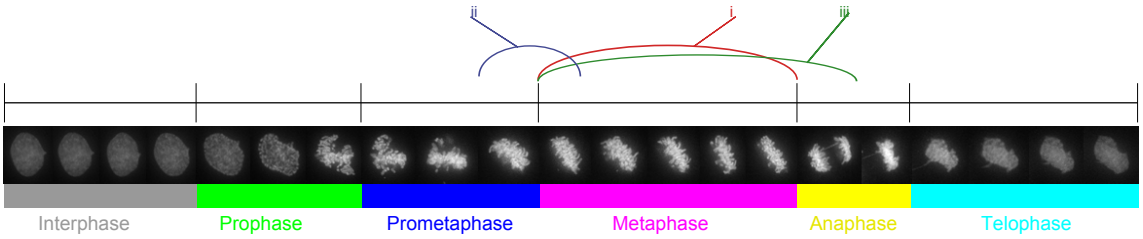


Figure 4.5: SMM feature mapping. Shown here for metaphase: (i) Segment level feature captures characteristics over the phase. (ii) Boundary feature captures behaviour at transitions. (iii) Neighbouring segment feature captures information about subsequent segment, allowing the model to learn possible transitions out of the phase.

(iii) **Neighbouring segment level feature vector.** This captures the interaction between neighbouring segments therefore allowing the model to learn which transitions between labels are feasible. The feature vector is the concatenation of the segment level feature vector evaluated over the current segment from a_j to b_j and over a subsequent segment from b_j to $b_j + d$, where d is a defined minimum duration. In this case the tensor product is taken with both the current and subsequent label vectors.

Using component (i) of the feature vector alone would be equivalent to learning a multi-class SVM over segments. Component (ii) adds additional information about the starting position of the segment to the feature vector. Adding component (iii), is what allows the model to learn relationships between phases, due to the tensor product with two label vectors.

4.3 Segmentation

The approach taken here is to consider the segmentation as a two-class classification task where cell nuclei are one class (foreground) and background is the other. We do this in a two stage process which is illustrated in Figures 4.6 and 4.7. First, a logistic regression classifier (Bishop, 2006) is used to generate a probability map of pixels being foreground:

$$P\left(foreground|\mathbf{x},\mathbf{w}\right)=\sigma\left(\mathbf{w}^{\top}\mathbf{x}\right)=\frac{1}{1+e^{-\mathbf{w}^{\top}\mathbf{x}}} \tag{4.12}$$

where $\mathbf{x}$ is a feature vector evaluated at every pixel in the image, and $\mathbf{w}$ a weight vector learned by training on a manually segmented subset of the training data.

In order to learn the weight vector, we minimise the negative log likelihood function given by:

$$\mathcal{L}\left(\mathbf{w}\right)=-\sum_{i}^{N}y_{i}\log\sigma\left(\mathbf{w}^{\top}\mathbf{x}_{i}\right)+\left(1-y_{i}\right)\left(1-\log\sigma\left(\mathbf{w}^{\top}\mathbf{x}_{i}\right)\right) \tag{4.13}$$

where N is the number of training data points and y_i are their labels. This can be minimised using gradient descent.

Before segmentation, each image sequence is individually normalised such that all pixel intensities lie in the range $[0,1]$. The feature vector $\mathbf{x}$ is a pixel intensity histogram computed over a 5×5 window centred on the test pixel. This gives an approximately rotationally invariant measure which is also robust to noise in the images.

We then use graph cuts (Boykov and Jolly, 2001) to obtain the final, globally optimum, binary segmentation, L, by minimising the energy:

$$E(L) = \sum_{p} D_p + \lambda \sum_{\{p,q\}} V_{pq} \tag{4.14}$$

where $\{p, q\}$ is a local pixel neighbourhood, D_p is the unary term, V_{pq} is a pairwise cost, which penalises differing labels in neighbouring pixels unless there is an intensity gradient between them, and λ is a weighting parameter between the two terms (set by cross-validation on the training data).

We use the output of the logistic regression classifier from Equation (4.12) as our unary term, D_p. For the smoothness term V_{pq}, we set the cost to 0 for $L_p = L_q$, and to the local image gradient for $L_p \neq L_q$. The image gradient is used as it is expected to have high values at discontinuities (cell boundaries) in the image, therefore V_{pq} will have small values where there should be discontinuities in the labelling and will not penalise them.

We use an 8-connected neighbourhood for the pairwise term. The image gradient is computed by filtering with the derivative of a 7×7, $\sigma = 1$ Gaussian rotated to 4 evenly spaced orientations, 0°,90° and ±45° degrees from horizontal to correspond to the directions of connectivity in the neighbourhood. Only 4 gradients are computed as we only use the gradient magnitude in our formulation.

We use the derivative of a Gaussian for the gradient computation as it is better for detecting smooth edges than the typically used difference of adjacent pixels. The Gaussian standard deviation is set to 1 pixel to minimise the smoothing effect. The

Gaussian window size is selected to be $6\sigma + 1$ (three standard deviations in each direction).

4.4 Tracking

Detected cells in each new frame are associated with existing tracks using a nearest neighbour match on their centroids and areas. This simple method assumes that the distance a cell is likely to move from one frame to the next is sufficiently small compared to the distance to other cells and that any change in size from one frame to the next is sufficiently small. In practice this is the case, and there is almost no track confusion or track loss.

The performance of this approach to tracking is highly dependent on both segmentation accuracy, and cell density. The dataset used contains a low density of cells, and the main focus of this work is not tracking, so we use the most simple approach which gives sufficiently accurate results. More challenging datasets would require the use of more complex tracking methods.

A simple heuristic is added to account for cells splitting, detected as two cells with similar appearance being closest the same existing track. Figure 5.1 shows sample frames of correctly tracked cells.

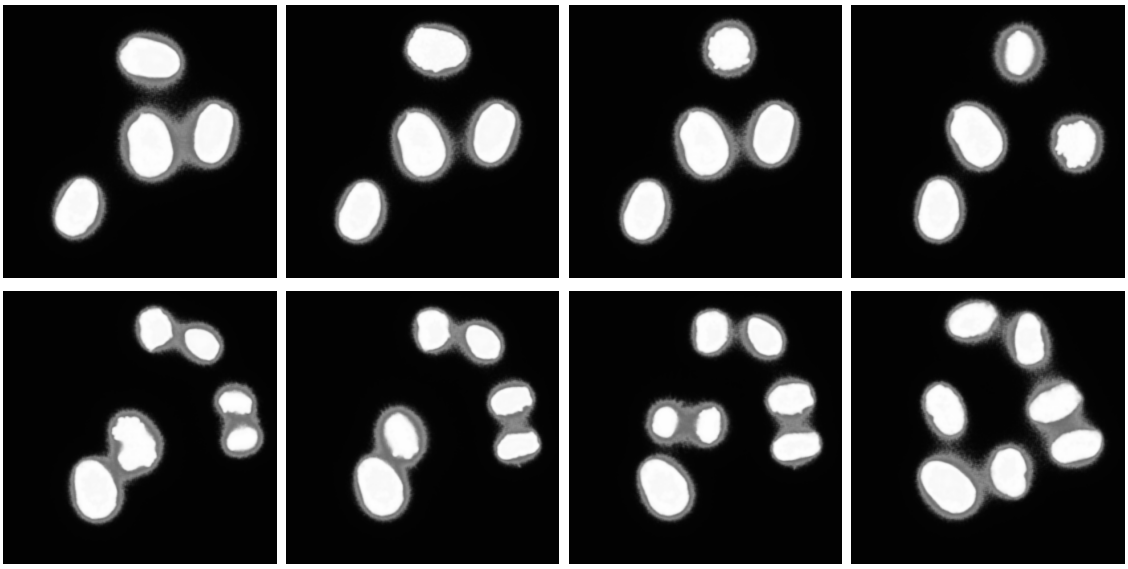


Figure 4.6: Segmentation process (1). Probability map output of the logistic regression classifier for the frames shown in Figure 4.1. White means high probability of being part of a cell.

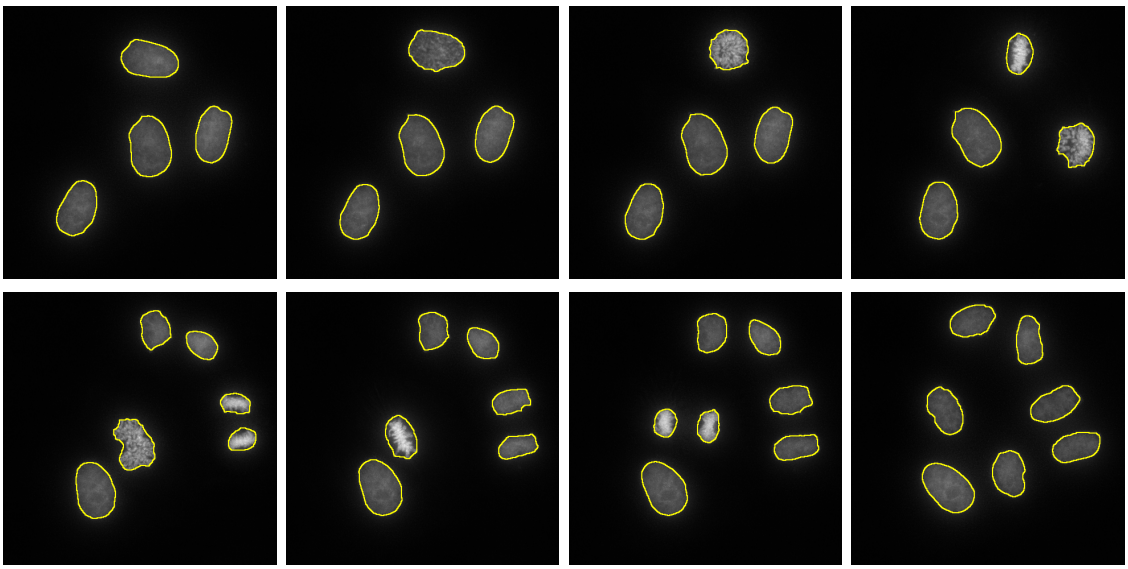


Figure 4.7: Segmentation process (2). Final segmentation results on the sample frames of Figure 4.1. Boundary shown in yellow, overlaid on the original image.

4.5 Mitosis Detection & Phase Labelling

To reduce the computational cost of the overall framework, a simple method is used to distinguish between mitotic and interphase-only tracks. Using the fact that interphase cells remain largely unchanged over the course of a sequence, whereas mitotic cells change significantly in size and brightness, a linear SVM is learned on the mean and variance of a number of pixel intensity features (max, mean, variance), as well as area and eccentricity of cells over the course of an entire track. Mitotic tracks are then labelled using the methods described in Section 4.2.

4.6 Temporal Features

The maximum pixel intensity signal is extracted for each tracked mitotic cell at every time point. These signals are then normalised on a per-track basis to lie in the range $[0, 1]$. Gradients are computed at two scales by filtering with the derivative of a 1×7, $\sigma = 1$ and a 1×13, $\sigma = 2$ Gaussian (window size $1 + 3\sigma$ in each direction), then normalised on a per-track basis to the range $[-1, 1]$. These three signals are concatenated to give a 3×1 vector at each time frame, used for the temporal models. Gradients are computed at two different scales as the smaller scale gives a measure of local gradient and the larger scale gives a measure of the longer term evolution of the signal.

For SMM, each temporal sample of a feature signal is additionally discretized into evenly spaced bins using "soft assignment", where the weight assigned to each bin is a linear function of distance to the bin centre, and represented by an L1-normalized

signal vector. Bins of width 0.1 (a design choice based on the trade-off between small bins and high-dimensional vectors) are used for the both the intensity and gradient features. This feature encoding is better suited to the linear kernel of the discriminative SMM. A sequence $\mathbf{x}$, as described in Section 4.2.3, consists of a concatenation of these temporal signal vectors.

4.7 Summary

In this chapter we have described three temporal models for mitotic phase labelling of tracked cells: Dynamic Time Warping, Hidden Markov Models, and Semi Markov Models. We have taken the novel approach of using temporal features evaluated over the whole of the mitotic phases rather than over single frames, thereby capturing the distinctive behaviour over the phases. Additionally we proposed a new loss function for the Semi Markov model to make it more robust to inconsistencies in data annotation near transition boundaries. We concluded by describing a simple framework for segmentation, tracking and classifying mitotic tracks from images cells tagged with fluorescent chromatin marker. Performance of the phase labelling methods is assessed in Chapter 5.

CHAPTER 5

EXPERIMENTS FOR MITOTIC PHASE LABELLING

In this chapter we assess the performance of the three different temporal models: Dynamic Time Warping, Hidden Markov Models, and Semi Markov Models described in the previous chapter. The models are tested under two different experimental conditions to explore robustness of the methods to changes in biological conditions. We also compare performance to the *Cellcognition* approach (Held et al., 2010). This method, described in Section 3.3, follows the same overall, three-stage, framework of our approach (segmentation, tracking, labelling), but uses an independent frame-based classification which is subsequently smoothed with a Hidden Markov Model. Comparing to this approach allows us to to show the advantage of our methods' ability to capture the temporal behaviour of features.

For the purposes of this chapter, only tracks of mitotic cells, extracted as described in Section 4.5, are considered.

Chapter outline

Section 5.1 describes the dataset used for the experiments, including a description of the manual annotation necessary for quantitative evaluation and the extraction of features used for the three temporal models. Descriptions of the two experiments tested and the evaluation measure used to compare methods are described in Section 5.2. Section 5.3 shows the results of the experiments and some sample outputs. Finally, Section 5.4 discusses potential areas of improvement for the methods presented here.

5.1 Data

For the purposes of this chapter, we use the dataset introduced in El-Labban et al. (2011). This dataset consists of 54 3D time-lapse image sequences of HeLa (human epithelial adenocarcinoma). To visualize chromatin, HeLa cells stably expressing histone H2B-mCherry and beta tubulin (TUBB)-EGFP are used. The HeLa cells were maintained in Dulbecco's modified eagle medium (DMEM, LifeTechnologies) with 10% fetal bovine serum (FBS), penicillin/streptomycin, and antibiotics for transgenes (geneticin and puromycin), in a 5% CO2 incubator.

To film chromatin live, the HeLa cells were seeded onto a 8-well μ-slide chamber (ibidi GmbH). Before imaging, medium was replaced with a pre-warmed CO2-independent medium. For data acquisition, a DeltaVision system (Applied Precision) equipped with a 60x oil objective and a heat chamber was used. To image both interphase and mitotic cells, which exhibited the typical height of around 5 μm

Sequence	No. of tracked cells	No. of tracked mitoses	Depleted protein
1-28	255	49	control siRNA
29-40	124	33	TACC3
41-46	95	13	CLTC
47-54	96	23	GTSE1
Total	**579**	**118**	-

Table 5.1: Cell sequences in the dataset. Sequences 1–28 are of cells under control conditions. The remainder were obtained under conditions in which one of three proteins required for timely progression though mitosis (TACC3, CLTC, or GTSE1) were depleted from cells by RNAi.

and 20 μm respectively, the cells were filmed at 1 to 5 minute intervals with 5-14 z-sections at 1-3 μm steps. The images were deconvolved and maximally projected.

Of these sequences, 23 were obtained under conditions in which one of three proteins required for timely progression though mitosis (TACC3, CLTC, or GTSE1) were depleted from cells by RNA interference (RNAi). This information is summarised in Table 5.1. RNAi was performed as described by Bird and Hyman (2008).

Frames have around 8–10 fully visible cells, with more partially visible around image boundaries, and there are up to 4 mitotic cells per video. Some examples of typical frames from the dataset can be seen in Figure 4.1.

5.1.1 Ground Truth

For the phase labelling experiments, only tracks of mitotic cells, extracted as described in Section 4.5, are considered. These 118 detected mitotic tracks in all 54 sequences are manually annotated, assigning each frame to interphase or one of the mitotic phases illustrated in Figure 4.2 and these tracks are used for training and

evaluation of the mitotic phase labelling methods. The annotation was carried out by one of our biological collaborators (*Y. Toyoda*), and verified by a second (*A. W. Bird*). Manual annotation was based on the appearance of both chromatin, and tubulin in the sequences.

5.2 Evaluation Measures

Two training and testing data splits are considered: *Experiment 1*: using half the sequences, chosen randomly, for training and the other half for testing, giving 58 training tracks and 60 test tracks; *Experiment 2*: using the sequences of control siRNA transfected cells for training and the remaining sequences for testing, giving 49 training tracks and 69 test tracks.

Confusion matrices of frame classification into phases are generated for each test track after labelling all frames. The mean per-class accuracy, where the class is one of six phases, is then calculated as the mean of the diagonal of the confusion matrix. This measure is preferred to the per-frame classification accuracy which is dominated by correct classification of interphase frames which appear approximately 10× more than the next most frequent phase in the mitotic tracks.

To assess how accurate the labelling methods are near phase boundaries, classification accuracy is also measured for the cases where transition boundaries are allowed to be within ±1 and ±2 frames of their true positions. These values are selected as the smallest possible increments of slack, without completely disregarding whole phases from the evaluation.

5.3 Results

The three phase labelling approaches are trained and tested using the data splits described in Section 5.1, and in the case of the SMM [1], using both loss functions described in Section 4.2.3. Three feature combinations are tested: the maximum pixel intensity signal, the gradient of the maximum intensity signal (computed at two scales) and a combination of all three feature signals. For comparison, we also test the SVM based approach of *CellCognition*[2] software using the 186 features described in Held et al. (2010). Qualitative results are shown in Figures 5.1 and 5.2.

5.3.1 Experiment 1

Table 5.2 shows the results for DTW, HMM and both versions of SMM approaches using the various feature combinations. There is a clear improvement in classification accuracy when using the SMM, which achieves very high accuracy within two frames of the true boundaries. The table also shows the improvement given by our change to the original loss function of Shi et al. (2011) used in training the SMM.

The mean per-class evaluation is very sensitive to the number of frames in the shorter phases. For example, for the case of using all features with our loss function (final row of Table 5.2), the overall proportion of correctly classified frames only improves from 0.90 to 0.92 by allowing a 2 frame slack, whereas the mean per class score improves from 0.81 to 0.91. It can be seen from the confusion matrices for the SMM shown in Table 5.4 that the most significant improvements are in prophase, metaphase and anaphase. This is likely due to the fact that these phases have very

[1] Code available from: http://cs.adelaide.edu.au/~javen/code/smm_release.tgz
[2] http://cellcognition.org/

short durations on average in the data (7, 10, and 4 mins respectively on average, compared with an average of 206 mins for interphase) and can appear for as little as a single frame in the sequences with the lowest temporal resolutions.

The most significant labelling errors occur at the end of prometaphase and telophase where, as seen in Figure 4.2, the change in the intensity signal is quite gradual and frequently affected by signal noise. The error is particularly clear for the case of telophase where the model consistently predicts a transition to interphase too soon, suggesting that intensity alone is not sufficient to distinguish between the two phases.

The performance of all three temporal models is significantly better than the independent frame-by-frame based SVM of *CellCognition*, despite using a much smaller number of features in our case. In fact, the SMM achieves better performance using a single feature signal. Adding the HMM error correcting step to *CellCognition* improves performance, however scores are still lower than our HMM and SMM for the three feature signal case, and comparable to DTW.

5.3.2 Experiment 2

The phase labelling approaches are also trained using the 28 control sequences, as listed in Table 5.1, and tested on the remaining RNAi sequences. Results are shown in Table 5.3. Confusion matrices for the SMM, three feature case are shown in Table 5.4.

The performance of all methods decreases, as expected, due to the differences in the biological conditions between the training and testing sets. However, the decrease in SMM performance is significantly smaller than for any of the other methods.

The largest drop in performance occurs with the HMM which is consistent with the fact that it models phase duration through the transition probabilities and the RNAi sequences are expected to have delays in mitosis. The expected phase duration using the HMM is given by $d = \frac{1}{1-t}$, where t is the transition probability corresponding to staying in the current state. However, there is no learning of phase duration with the SMM. The segment based approach of the SMM therefore avoids the dependence on the unrealistic exponentially decaying duration model of the HMM or the duration model used to learn the reference signal for DTW.

The use of our modified loss function does not improve SMM performance using this train/test split, most likely due to greater annotation consistency within the control set than across different sets of sequences.

The drop in *CellCognition* SVM performance is most likely due to the significantly smaller number of training samples. However, due to the fact that there is no temporal modelling, it outperforms our DTW and HMM. The addition of the HMM stage still improves performance slightly as, unlike our HMM, it is learned from the test data and is unaffeced by the change in experimental conditions. However, the *CellCognition* approach is still outperformed by our SMM in this experiment.

5.4 Summary

In Chapters 4 and 5 we have described and compared three machine learned temporal models for mitotic phase labelling: (i) aligning temporal signals of simple features directly to a reference signal using Dynamic Time Warping; (ii) learning a Hidden Markov Model from feature signals; (iii) a discriminative Semi-Markov

Feature & method	Exact	±1 frame	±2 frames
CellCognition (186 features, no HMM)	0.71	0.81	0.86
CellCognition (186 features, HMM)	0.75	0.86	0.90
Max Intesity			
(1) DTW	0.61	0.71	0.76
(2) HMM	0.33	0.40	0.45
(3) SMM with loss function of Shi et al.	0.72	0.81	0.84
(4) SMM with our loss function	0.75	0.84	0.89
Max Intensity gradient			
(1) DTW	0.44	0.54	0.58
(2) HMM	0.71	0.83	0.89
(3) SMM with loss function of Shi et al.	0.74	0.82	0.85
(4) SMM with our loss function	0.76	0.84	0.87
Max Intensity w/gradient			
(1) DTW	0.76	0.84	0.87
(2) HMM	0.80	0.88	0.89
(3) SMM with loss function of Shi et al.	0.76	0.85	0.88
(4) SMM with our loss function	0.81	0.89	0.91

Table 5.2: Phase Labelling Results (1). Results for train and test sets with mixed biological experimental conditions. Scores shown are the mean per class accuracy i.e. the mean of the diagonal of the confusion matrix. Results are given for exact matching and allowing for 1 or 2 frames slack in detection of the transition boundary.

Model incorporating information over temporal segments. The SMM is shown to outperform an independent frame classification approach using a significantly larger feature set.

We also introduce an alternative SMM loss function to the original of Shi et al. (2011) which is robust to inconsistencies in training data annotation at phase transition boundaries. This new loss function outperforms the original in this context.

Additionally, the Semi-Markov's segment based approach is shown to be robust to differences between the training and testing datasets. Considering variations of features over whole mitotic phases allows the labelling to cope with changes

Feature & method	Exact	±1 frame	±2 frames
CellCognition (186 features, no HMM)	0.65	0.77	0.83
CellCognition (186 features, HMM)	0.66	0.80	0.85
Max Intesity			
(1) DTW	0.57	0.75	0.82
(2) HMM	0.17	0.30	0.42
(3) SMM with loss function of Shi et al.	0.59	0.67	0.73
(4) SMM with our loss function	0.63	0.75	0.80
Max Intensity gradient			
(1) DTW	0.42	0.59	0.70
(2) HMM	0.17	0.28	0.37
(3) SMM with loss function of Shi et al.	0.54	0.64	0.70
(4) SMM with our loss function	0.56	0.71	0.79
Max Intensity w/gradient			
(1) DTW	0.58	0.78	0.88
(2) HMM	0.21	0.37	0.50
(3) SMM with loss function of Shi et al.	0.51	0.63	0.70
(4) SMM with our loss function	0.74	0.84	0.88

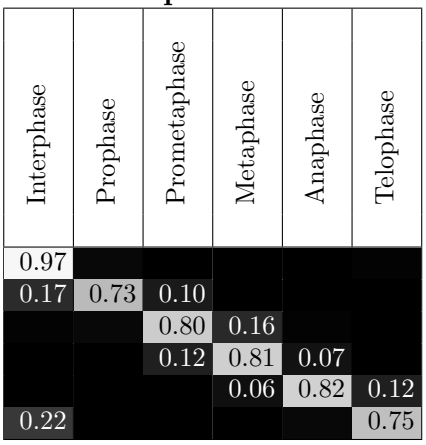
Table 5.3: Phase Labelling Results (2). Results using control cells for training and RNAi depleted cells for testing. Scores shown are the mean per class accuracy i.e. the mean of the diagonal of the confusion matrix. Results are given for exact matching and allowing for 1 or 2 frames slack in detection of the transition boundary.

in experimental and imaging conditions. Specifically, the performance drop when training the model on sequences of control RNAi cells and tested on cells expected to experience mitotic delays is much smaller than that of other labelling methods.

The SMM can therefore be used to detect delays in mitotic phases, and the robustness shown by the SMM shows that the methods can be used for new biological experiments. Additionally, the phase labelling results of the temporal models can be used as a basis for more detailed additional analysis, as will be seen in Chapter 6.

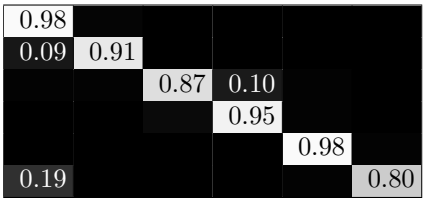
Further work, discussed in greater detail in Chapter 9, is necessary to investigate how the addition of more features impacts the performance of the SMM. Alterna-

Experiment 1



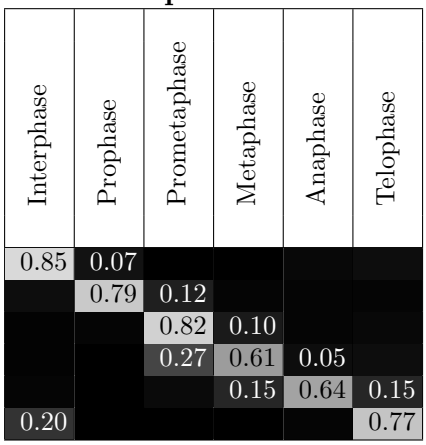
(1a) Exact boundary.
Mean per-class accuracy: 0.81

(1b) ±1 frame.
Mean per-class accuracy: 0.89



(1c) ±2 frames.
Mean per-class accuracy: 0.91

Experiment 2



(2a) Exact boundary.
Mean per-class accuracy: 0.74

(2b) ±1 frame.
Mean per-class accuracy: 0.84

(2c) ±2 frames.
Mean per-class accuracy: 0.88

Table 5.4: Confusion matrices. Confusion matrices of testing results for Experiments 1 (left) and 2 (right), using the SMM with three features, allowing up to 2 frames of slack in the classification. Rows represent the true classes and columns are the predictions. Values less than 0.05 have not been shown.

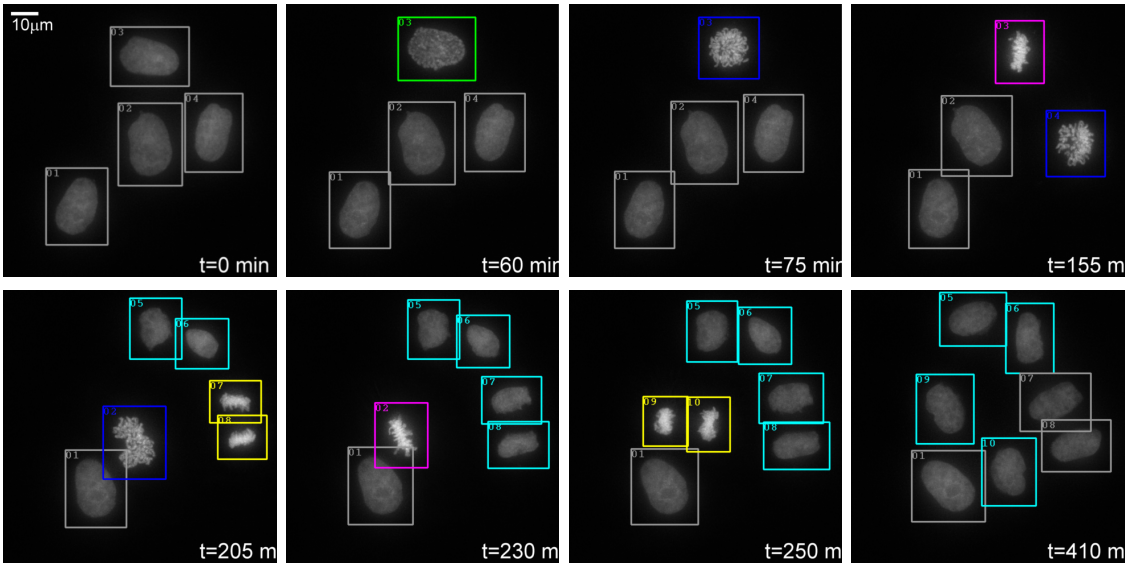


Figure 5.1: Labelling results (1). Final labelling results on sample frames from Figure 4.1. Cells are tracked correctly as shown by numbered bounding boxes, which are colour-coded by phase using the same colour-code as in Figure 4.5.

tive feature encodings could be explored, as the high-dimensionality of the current encoding makes the addition of more features very computationally expensive.

There is potential room for improvement in the overall framework through use of more robust cell detection methods, e.g. (Arteta et al., 2012), and globally consistent, though more complex, tracking methods e.g. (Lou and Hamprecht, 2011; Padfield et al., 2011).

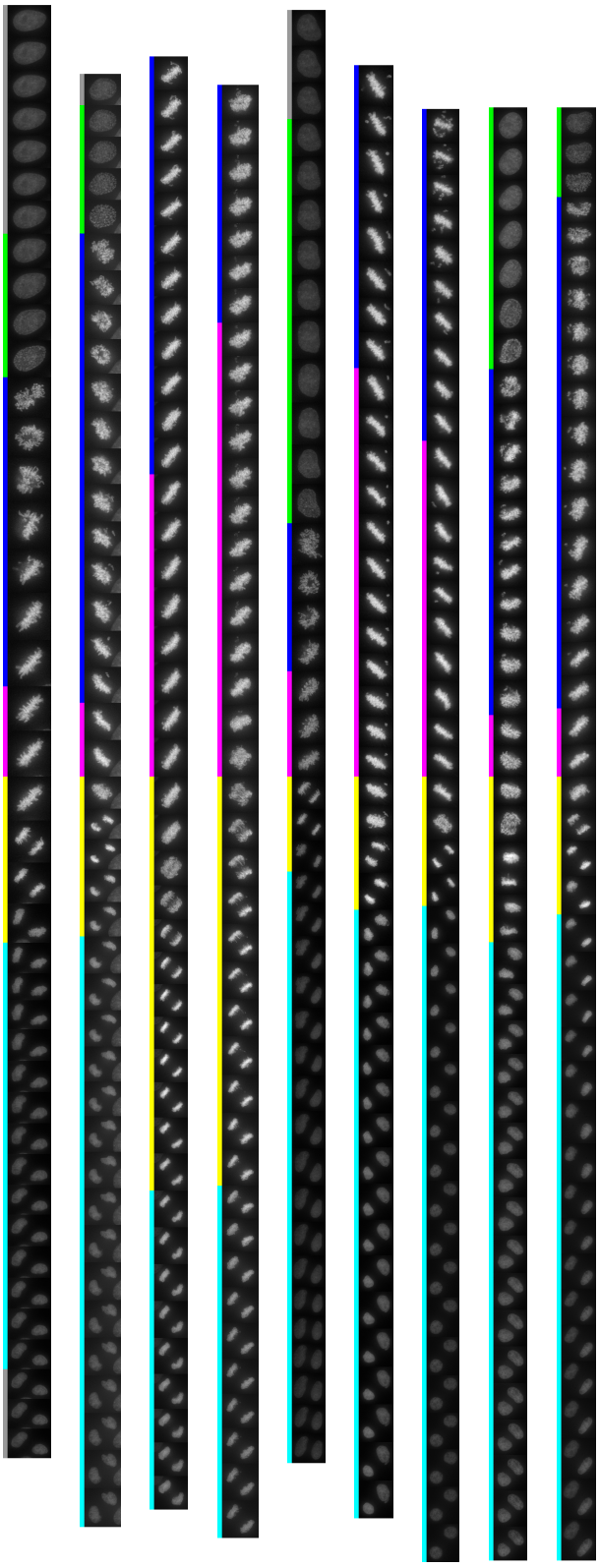


Figure 5.2: Labelling results (2). Final labelling results for a number of tracks. Tracks are aligned to the metaphase–anaphase transition, with frames before and after the transition shown. Tracks are colour-coded by phase using the same colour-code as in Figure 4.5.

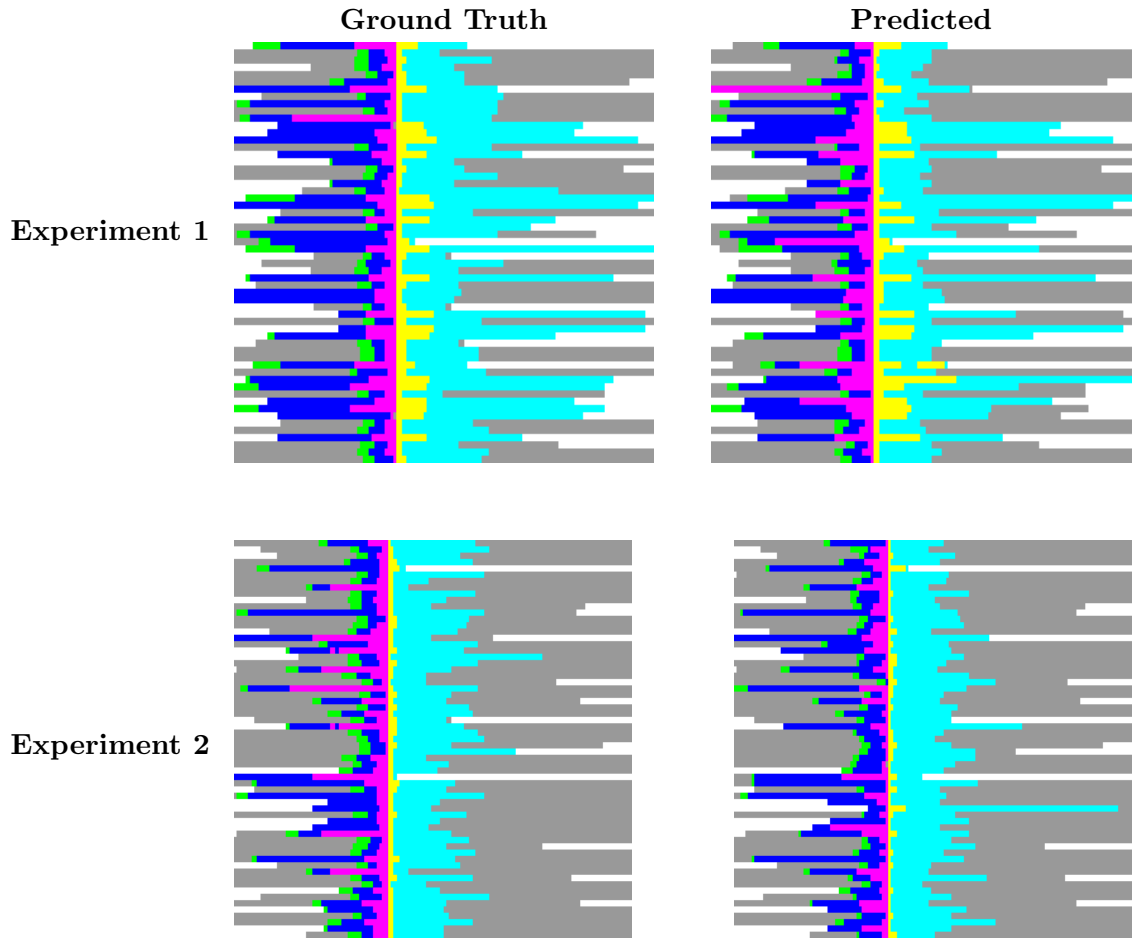


Figure 5.3: Result tracks. Colour-coded tracks comparing ground truth labellings (left) to those predicted by the SMM (right) for Experiment 1 (top) and Experiment 2 (bottom). Colour-code is as in Figure 4.5. Tracks are all aligned to the metaphase-anaphase transition.

CHAPTER 6

DETECTION OF CHROMOSOME SEGREGATION ERRORS IN EMBRYONIC STEM CELLS

The detection of chromosome segregation errors in mitosis is an important area of biological research. Due to the rarity and subtle nature of such errors in untreated cell lines, there is a need for automated, high-throughput systems for quantifying the rates at which such defects occur.

In this chapter we present a novel approach for detecting subtle chromosome segregation errors in mitosis in embryonic stem cells, targeting two cases: misaligned chromosomes in a metaphase cell, and lagging chromosomes between anaphase cells.

Our method builds on the approaches for analysis of HeLa cell lines from previous chapters which label mitotic phases through mitosis. We apply these to more challenging, denser, stem cell lines. Leveraging the mitotic phase labelling allows us to

detect smaller, more subtle defects within mitosis. This results in a very high recall rate, as necessary for detection of such rare events.

The work in this chapter was originally presented in El-Labban et al. (2013). In this chapter we expand on the implementation details and provide additional results figures.

6.1 Introduction

6.1.1 Biological Motivation

When a cell divides, it must segregate all of its chromosomes equally into two daughter cells. Errors in chromosome segregation can lead to the gain or loss of chromosomes in a cell, resulting in a state known as aneuploidy. Aneuploidy can lead to problems in development, and cancer cells frequently display aneuploidy resulting from a high rate of chromosome missegregation, a condition termed chromosomal instability (Compton, 2011).

Compared to many cancer cell lines routinely studied, such as HeLa (human epithelial adenocarcinoma) cells, embryonic stem cells have relatively low levels of chromosome instability (Saito et al., 2011), making them useful for identifying conditions that lead to small but biologically significant increases in chromosome segregation defects. It is also important to be able to determine accurately the degree of chromosome segregation errors and aneuploidy in stem cell lines, as they become candidates for medical treatments.

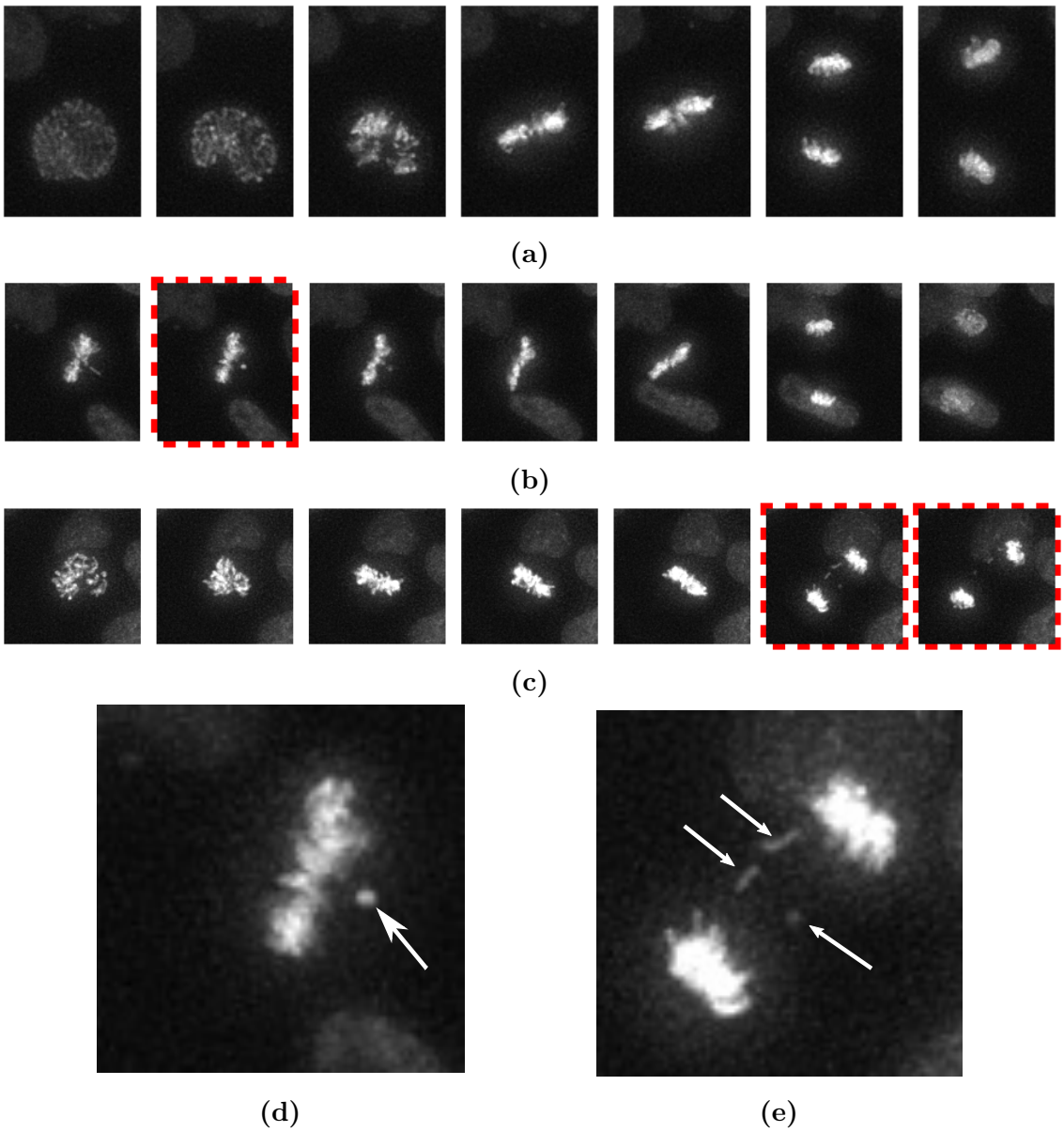


Figure 6.1: Examples of chromosome segregation defects. (a) Track of cell undergoing mitosis normally. (b) & (c) Tracks of two cells undergoing mitosis, frames with errors are highlighted in red (dashed). (d) & (e) Enlarged view of frames with defects, indicated with white arrows. (b) & (d) Misaligned chromosome, fails to align to metaphase plate during mitosis. (c) & (e) Lagging chromosomes, not segregated into either daughter cell during mitosis.

Due to the rarity of such errors in stem cell lines, large amounts of data need to be analysed to detect statistically significant changes in error frequency under different experimental conditions. Additionally, due to the subtlety of the defects, human annotation can be difficult and time-consuming. There is therefore a need for automated high-throughput methods for quantifying this data.

The novel aspect of this work is that it aims to detect much more subtle mitotic errors than previous methods, which focused on detection of substantial variations from normal mitoses (i.e. durations or morphologies). Our method leverages mitotic track phase labelling information to significantly reduce the search space for defects to just the phases in which the desired defects are expected to occur. This approach uses existing methods to detect and track mitotic cells and label the mitotic phases over time, though applied to a different cell type than previous methods – embryonic stem cells rather than HeLa cells. The detection of errors targets two cases: either around a metaphase cell for misaligned chromosomes or between anaphase cells for lagging chromosomes, as illustrated in Figure 6.1. This 'phase aware search' results in reliable detections, with a high recall rate, as required for this application given the rarity of defects.

Chapter Outline

In Section 6.2 we cover the full method for detecting chromosome segregation defects in mitosis, including: mitotic cell detection, phase labelling, candidate defect detection and candidate defect filtering. Section 6.3 describes the stem cell image dataset used for experiments. Experiments and results are discussed in Section 6.4.

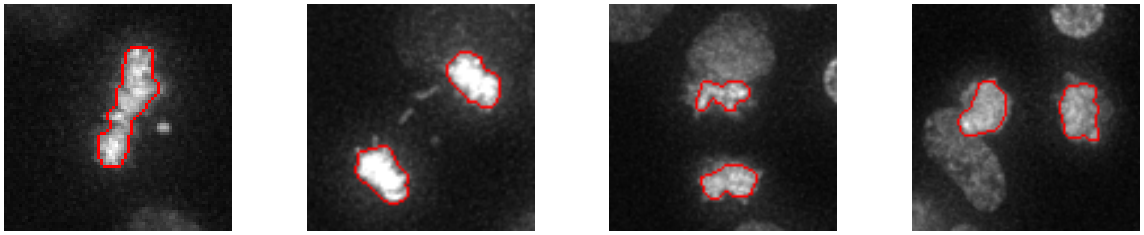
6.2 Method

We follow a three stage approach for detecting chromosome segregation defects: first detecting and tracking mitotic cells, second labelling the phase of mitosis in every frame of the tracks, and finally detecting defects in metaphase (misaligned chromosomes) and anaphase (lagging chromosomes).

6.2.1 Detection & Tracking

Due to the very high density of cells in the data, resulting in a large proportion of cells touching or overlapping, detecting and tracking all cells in a sequence would be an extremely challenging task; simple foreground segmentation techniques would not differentiate between cells. Also, it is unnecessary to detect and track all cells when the purpose is detecting defects in mitosis. Instead, we detect only mitotic cells (prometaphase, metaphase, anaphase and telophase) which can be distinguished from interphase cells as they are typically substantially brighter in appearance. The extremal region based approach of Arteta et al. (2012) is used for this purpose. This method involves first proposing hundreds of candidate extremal regions, and then selecting the regions of interest using a classifier together with a non-overlap constraint. It can robustly detect mitotic cells even when they overlap with interphase cells, as can be seen in Figure 6.2, and the resulting regions provide good chromatin segmentations for use in subsequent steps. For brevity, the chromatin segmentations will be referred to as cells in the remaining sections.

Cells are then associated from frame to frame by a nearest neighbour approach using a feature consisting of centroid position and size, and incorporating validation



(a) (b) (c) (d)

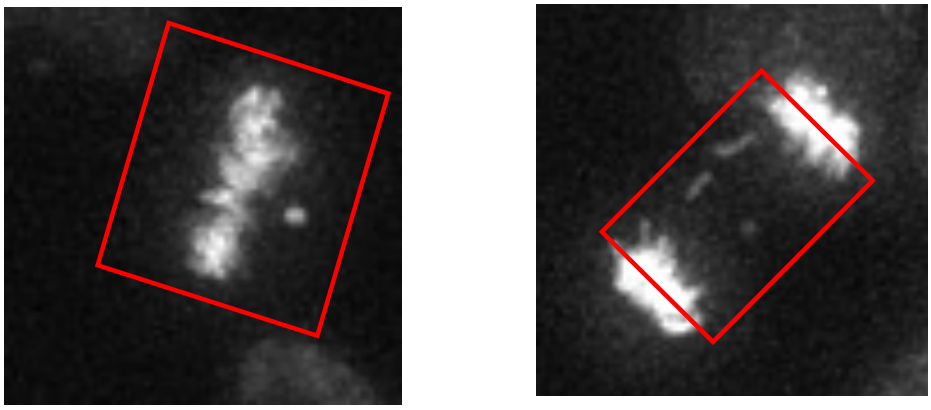
Figure 6.2: Example detection results. Detected extremal region boundaries overlaid on images. Detection gives good segmentation boundaries and copes with overlapping or touching cells e.g. in (c) and (d). View in colour to see overlaid boundaries.

gating in the form of an association threshold. This approach is sufficient to cope with the relatively well spaced out detections, both spatially and temporally.

6.2.1.1 Track post-processing

A post-processing stage is applied to remove false positive tracks: a classifier is used to remove tracks of artefacts from cell death, and non-mitotic cells which are bright enough to be detected. The features used for this consist of the mean, standard deviation, maximum and minimum values of cell maximum pixel intensity and size over a track. These features should vary substantially as a cell undergoes mitosis, as seen for pixel intensity in Figure 4.2, but remain reasonably constant for artefacts and interphase cells, whose appearance remains unchanged throughout the tracks. The classifier used is a linear SVM trained using LIBSVM[1] (Chang and Lin, 2011). The classifier is trained on a 100 track subset of the data, and the parameter is set with two-fold cross-validation.

[1] http://www.csie.ntu.edu.tw/~cjlin/libsvm



(a) (b)

Figure 6.3: Aligned regions of interest for defect detection. (a) Misaligned chromosomes occur along the major axis of metaphase cells. (b) Lagging chromosomes appear between anaphase cells. Detections within these regions are encoded in a 3×3 bin spatial histogram. Segmented cell regions, shown in Figure 6.2, are excluded from the region of interest.

6.2.2 Phase labelling

For phase labelling we use the HMM based method from Chapter 4. As the performance of this method improves when a track is sufficiently long to contain all mitotic phases, the detected mitotic tracks are continued three frames forward and backwards in time to obtain some frames with the cells in interphase, allowing for better normalisation of the features.

6.2.3 Defect detection

For detection of the defective chromosomes, a two stage approach is used: first, candidate structures are detected in a region of interest around the cell, and second detections are filtered based on a number of intensity and spatial characteristics.

6.2.3.1 Candidate Detection

The region of interest is defined where defects are expected to occur – orthogonal to the major axis of metaphase cells, or between dividing anaphase cells. Example regions are illustrated in Figure 6.3.

Within the interest regions, candidate misaligned/lagging chromosomes are initially detected as blobs as local maxima of Laplacian of Gaussians filter response. The filtering operation is carried out at two different scales ($\sigma = 1, 2$ pixels) to account for variation in defect sizes. Note, the chromatin segmented region determined in Section 6.2.1 is excluded when proposing candidates and in other subsequent processing.

6.2.3.2 Candidate Filtering Stages

The candidates contain a large number of false positives primarily due to interference from other nearby cells. The purpose of the second stage is a set of filters to remove these false positives.

Intensity Filter. For the intensity based filtering step, we observe that actual mitotic defect detections have pixel intensities closer to the chromatin in the segmented cell than the background portion of the interest region. Therefore, we assume that the probability of a blob corresponding to a mitotic defect is proportional to the ratio of intensity differences:

$$p(defect|i_{blob}) \propto \exp\left(\frac{-(i_{blob} - i_{cell})^2}{2\sigma_{cell}^2} - \frac{-(i_{blob} - i_{bg})^2}{2\sigma_{bg}^2}\right) \tag{6.1}$$

where i_{blob} is the pixel intensity at the blob detection, i_{cell} and σ_{cell} are the mean and standard deviation of chromatin pixel intensities within the segmented cell, and i_{bg} and σ_{bg} are the mean and standard deviation of pixel intensities in the interest region background. A threshold, τ, on this measure removes false positives.

Size Threshold. The image is then thresholded based on intensity at each detection ($\pm 50\%$ of detected blob intensity), with the resulting segmentation used to estimate the size of the detected blob. Segmentations larger than individual chromosomes are discarded.

Spatial Entropy. The remaining detections are then encoded in a 3×3 bin spatial histogram. Histogram bin counts are weighted by the magnitude of detected peaks in the filter response, reducing the contribution of any remaining peaks caused by image noise. Entropy is then used as a measure to define defects, based on the observation that true defects will appear as well localised detections, and consequently generate histograms with low entropy.

6.3 Data

The data consists of 40 sequences of mouse embryonic stem cells. To visualize chromatin, R1/E mouse ES cells are stably transfected with a BAC expressing H2A-GFP. The ES cells were maintained in Dulbecco's modified eagle medium

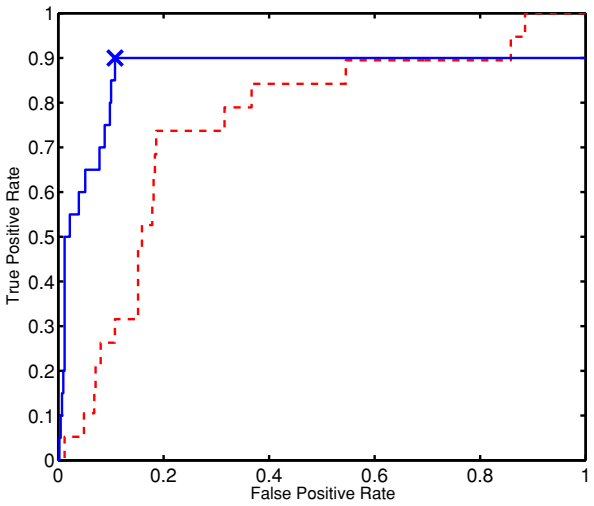
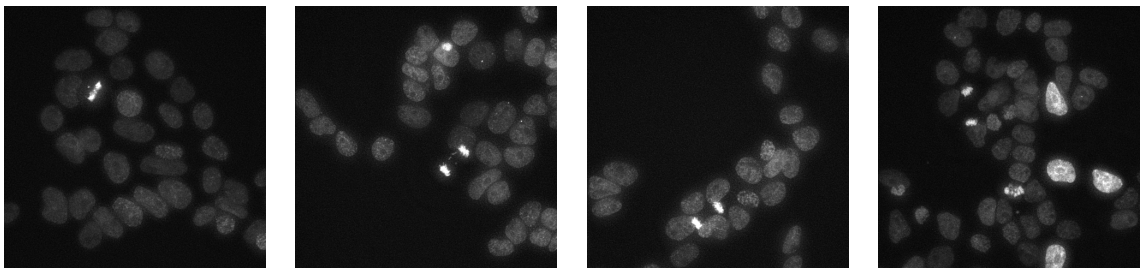


Figure 6.4: ROC curve of defective mitotic track detection. The red (dashed) curve uses only candidate defects (before filtering) as described in the first stage of Section 6.2.3. The blue curve incorporates mitotic phase and spatial information filters. The operating point used for results in Table 6.2 is indicated with a blue '×'.



(a) **(b)** **(c)** **(d)**

Figure 6.5: Four typical frames from the data. There is one visible mitosis in each frame. The challenges of this dataset include densely clustered cells, overlap of mitotic and interphase cells (e.g. in (b)-(c)), artifacts from dead cells (e.g. top of (b) and middle of (d)), and significant variations in interphase cell brightness (e.g. the four interphase cells on the right of (d)).

(DMEM, LifeTechnologies) with 20% ES-cell screened fetal bovine serum (FBS), penicillin/streptomycin, 1X NEAA, 1mM Glutamine, LIF, B-Mercaptoethanol and G418 for transgenesis, in a 5% CO2 incubator.

To film the chromatin live, the ES cells were seeded onto a 8-well μ-slide chamber (ibidi GmbH) coated with 1 μg/cm2 laminin-511. Before imaging, medium was replaced with a pre-warmed CO2-independent medium containing all the supplements in the culture medium. For data acquisition, a DeltaVision system (Applied Precision) equipped with a 40x oil objective and a heat chamber was used. To image both interphase and mitotic cells, the cells were filmed at 4 minutes intervals with 5 sections at 2 μm steps, in a total of 10μm height. The laser intensity was 10% during image acquisition and a 10 ms light exposure was applied onto the cells. Total imaging time was kept as 12 hours to prevent counting the same cell again. The images were deconvolved by enhanced-ratio method and maximally projected in Fiji.

The resolution is 480 $\times$ 480 pixels. Sequences initially contain 19 cells on average, approximately 760 in all sequences, a number which increases as cells divide or new ones move into focus, and last for an average duration of 170 frames. In total, over all frames in all sequences, there are approximately 13×10^4 possible individual cell detections. Sequences contain from 2 up to 41 mitoses. In total, 62% of all cells in the dataset undergo mitosis, giving 473 mitoses, of which 10 contain a misaligned chromosome error and 10 contain a lagging chromosome error (slightly over 4% of the 473 mitoses are defective). Mitoses typically last 10-12 frames, and defects are visible for at most 3 frames in mitotic tracks i.e. fewer than 0.2% of all possible

Sequences	
No. of sequences	40
Average frames per sequence	170
Average cells per sequence	19
Mitoses	
No. of mitoses	473
Average mitosis duration (frames)	11
Defects	
No. of mitoses with defects	20
Average defect duration (frames)	2

Table 6.1: Dataset statistics.

individual cell detections are mitotic and 0.03% contain chromosome segregation errors.

The data presents a number of challenges, illustrated in Figure 6.5, including densely clustered cells, overlap of mitotic and interphase cells, artifacts from dead cells, and significant variations in interphase cell brightness. The chromosomes that must be detected are extremely small: either blobs only 2-3 pixels in diameter, or as slightly elongated, strand-like objects 4-5 pixels long and 1-2 pixels wide.

6.4 Experiments & Results

Results are reported in Table 6.2 for the cell detection and mitotic defect detection components of this work. For each of these, performance values are given for two conditions: individual frame detections and whole track detections. The track performance is particularly important since detecting a defective mitotic track is still

	True Pos.	False Pos.	False Neg.	True Neg.	Precision	Recall
Mitosis detection (frame)	2853	2658	852	$\approx 13 \times 10^4$	0.52	0.77
Mitosis detection (track)	409	19	64	≈ 750	0.96	0.86
Defect detection (frame)	23	80	12	4491	0.22	0.66
Defect detection (track)	18	47	2	364	0.29	0.90

Table 6.2: Results for the two key stages of the method – mitotic cell detection, and defect detections. The performance of each stage is assessed on a frame-by-frame basis and on a whole track basis. Mitotic defect detection results are based on detected mitotic tracks from the previous stage. The true negative values for mitosis detection are estimates, as ground truth data is unavailable for non-mitotic cells and tracks.

useful for analysing large scale biological data even if one or two frames containing defects are missed.

For the mitotic cell detection part of this work, the 40 sequences in the dataset are divided into two equal partitions of 20 for training and testing. Within the training partition, two-fold cross-validation is used to set parameters of the detection system. Training is then carried out on the full 20 sequence partition and tested on the remaining half of the data. The partitions are then switched and the training and testing process is repeated to obtain mitotic cell detections for all 40 sequences. This results in very high detection accuracies, with overall precision and recall values of 0.96 and 0.86 for mitotic track detection. False positive track detections are mostly caused by artifacts remaining from cell death and oversegmentations of correctly detected cells.

The defect detection results are based only on detected mitotic tracks from the first stage. All of the 20 mitotic tracks in the dataset containing defects are detected.

For the mitotic defect detection, the threshold, τ, described in Section 6.2.3 is varied to produce ROC curves, as shown in Figure 6.4. This results in a very high rate of recall, with only two tracks with defects missed. Precision is lower due to the significant imbalance in the data. The two false negative tracks which occur, shown in Figure 6.10, are both lagging chromosomes in anaphase cells. All misaligned chromosomes in metaphase are correctly detected as seen in Figure 6.6. Correctly detected lagging anaphase chromosomes are shown in Figure 6.8.

6.5 Summary

In this chapter we have presented a novel method for detecting very rare and subtle mitotic errors in large volumes of microscopy images. The system reduces the search space for defects by detecting and tracking only the subset of cells in the data which undergo mitosis to reduce the total number of cells to analyse. Candidate defects are proposed using a simple blob detector, and subsequently filtered, based on size and appearance, as well as by mitotic phase as labelled by the system described in Chapter 4.

Our method is useful as a filter for further biological analysis, as it reduces the amount of data to be analysed manually by several orders of magnitude, although further work can improve the precision. Potential improvements are discussed in Chapter 9.

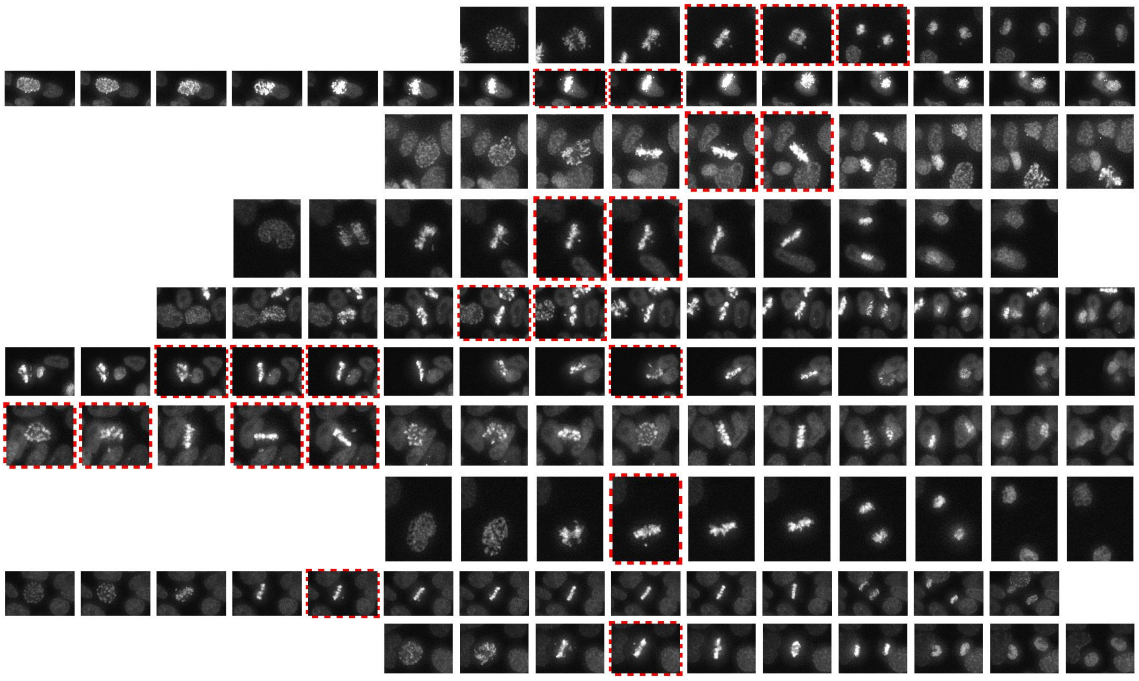


Figure 6.6: Correctly detected tracks containing misaligned metaphase chromosomes. Frames where defects are detected are highlighted in red (dashed). Tracks are aligned to start of anaphase.

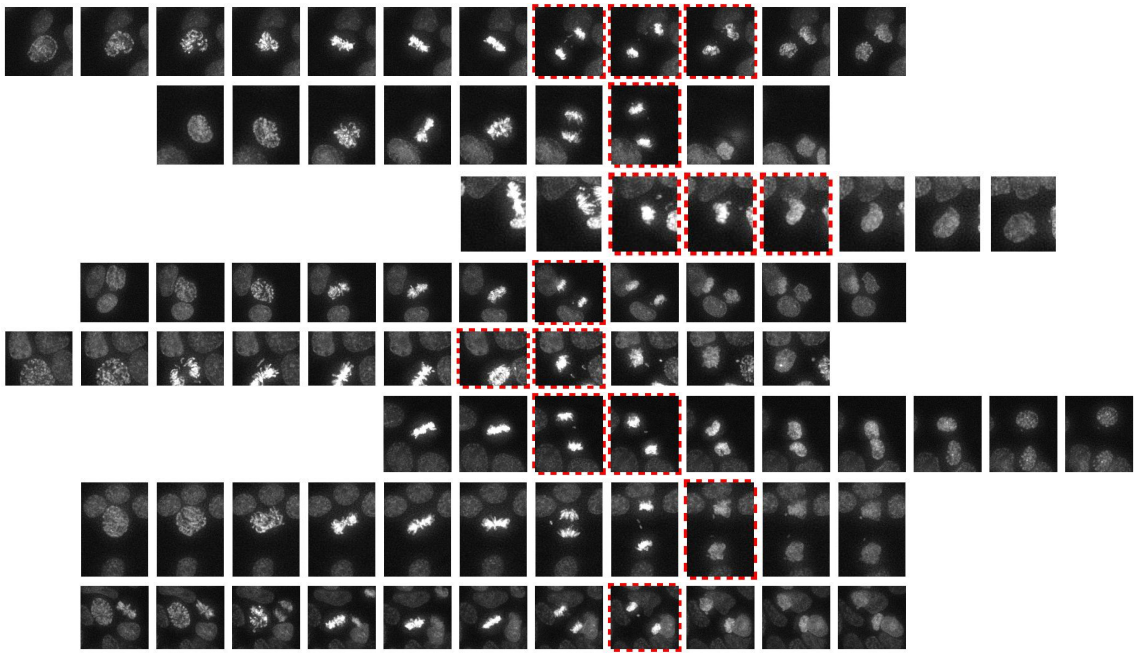


Figure 6.8: Correctly detected tracks containing lagging anaphase chromosomes. Frames where defects are detected are highlighted in red (dashed). Tracks are aligned to start of anaphase.

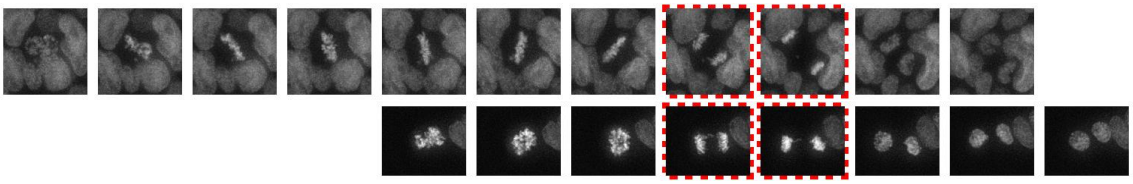


Figure 6.10: Missed tracks containing lagging anaphase chromosomes. Frames containing defects are highlighted in red (dashed). Tracks are aligned to start of anaphase.

CHAPTER 7

UNSUPERVISED DETECTION OF MITOTIC DEFECTS

In the previous chapter we described a discriminative method designed to detect specific, blob-like defects. In this chapter we investigate the alternative approach of using a generative method.

By modelling the appearance of regions surrounding large numbers of mitotic cells we are able to obtain a probabilistic model of the normal surroundings of cells. Defects, which are expected to be rare events, then have low probability under this model, and can then be discovered in a simple retrieval system.

We test this system for the case of misaligned metaphase chromosomes, using a variety of features and discuss their relative performance.

Chapter Outline

Section 7.1 outlines the problem and describes the underlying intuition behind our approach. The features used to represent image regions are described in Section 7.2. In Section 7.3 we describe the overall framework for the defect discovery system. Finally, in Section 7.4 results are shown.

7.1 Appearance model

As previously described in Chapter 6, chromosome segregation errors are extremely rare. In the dataset used here, as few as 0.03% of individual mitotic cell detections have a visible misaligned chromosome (Section 6.3). In contrast to the previous chapter, where we detected such defects by modelling them as blobs and finding likely candidates, here we attempt to model the 'normal' appearance of mitotic cell surroundings and then discover cells with unlikely surroundings, i.e. potential defects.

To learn the distribution of normal appearances we fit a Gaussian Mixture Model (GMM, Bishop 2006) to features evaluated at every pixel within an interest region around mitotic cells, as in Figure 6.3. Using a GMM allows the model to learn multiple types of normal appearance (e.g. empty background, and interphase cells). The intuition here is that the number of pixels representing the defects is sufficiently

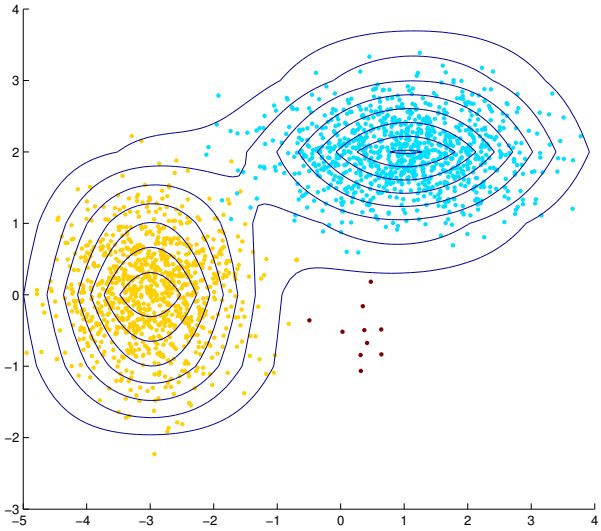


Figure 7.1: Synthetic example. Points sampled from three Gaussian distributions: 1000 points each from the yellow and pale blue distributions, representing normal features; 10 red points represent infrequent features from defects. Lines are contours of a three-component Gaussian mixture fitted to the full set of data. Red points are in the very low probability region and so would be retrieved by our system.

small, only 5-6 pixels for each of the 0.03% of mitotic detections, so as not to significantly influence the fitting of the GMM. Pixels containing defects would then have a low probability of occurring under the model. Figure 7.1 illustrates this intuition using simple synthetic data.

7.2 Features

As features we test the responses of three filter banks, combined with one of two representations of position relative to the cell. The features are evaluated in an interest region around the cells, as previously shown in Figure 6.3.

7.2.1 Filter banks

The three sets of filters we consider are:

1. **Laplacian of Gaussian (LoG).** As used for candidate defect detection in Section 6.2.3. Here we use an additional two scales ($\sigma = 1, 2, 5, 8$ pixels) as the objective is to model the appearance of other objects in the background, not just the defects.

2. **Maximum Response (MR).** Using an approach similar to Varma and Zisserman (2005), we use a set of bar and edge filters at four orientations and two scales ($\sigma_x = 1, 2$ pixels, $\sigma_y = 3\sigma_x$), and use the maximum response over orientations for each scale. These are then combined with the responses of Gaussian and Laplacian of Gaussian filters at the same scales.

3. **Schmid filters (S).** Using the filters of Schmid (2001), which are rotationally invariant. These filters combine frequency and scale, and so can capture some measure of texture. We use a total of ten scale/frequency combinations.

7.2.2 Position representation

Position is measured on axes with origin at the centre of the interest region. The first position representation we consider is a cartesian coordinate system, where (x, y) values are normalised to lie in the range $[-1, 1]$ to make the measure invariant to cell size.

We also consider a polar coordinate representation, (r, θ), with angle, θ measured relative to the major axis of the cell. However, rather than taking r to be the distance from the origin, the value of the distance transform of the cell segmentation is used.

7.3 Retrieval Framework

The retrieval framework used in this chapter, follows the following stages: (i) using the same data as in Chapter 6, the same detection, tracking, and phase labelling stages as described in Section 6.2 are applied; (ii) individual frames are selected based on phase; (iii) an extra background masking stage is applied; (iv) features, as described in Section 7.2, are evaluated; (v) a GMM is fit to the data and probability

of the images under the GMM is evaluated; (vi) images are ranked based on low probability.

7.3.1 Phase filtering

In theory, the misaligned metaphase chromosomes we aim to retrieve instances of should only occur in metaphase. However, due to errors in the automated phase labelling stage of the framework, a portion of ground truth defect-labelled frames are labelled as either prometaphase or anaphase (the phases preceding and following metaphase).

To account for this, we explore two approaches:

1. **Metaphase only.** Fit the appearance GMM to features evaluated only around cells labelled as metaphase, and subsequently consider retrieval results labelled as metaphase.

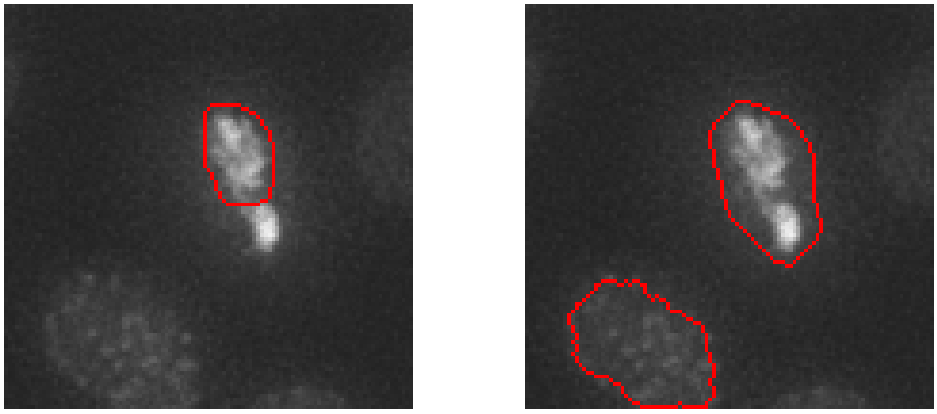
2. **Three phases.** Fit the appearance GMM to features evaluated around cells labelled as any one of prometaphase, metaphase, or anaphase and retrieve cells with any of these labels.

7.3.2 Cell masking

As in Chapter 6, detected mitotic cells are excluded from the interest regions used for feature evaluation. The approach we use for mitotic cell detection (Arteta et al., 2012) is not specifically designed to provide accurate segmentations. As seen in Figure 7.2(a), this sometimes results in undersegmentation of cells. If this occurs infrequently, it could result in false positive retrievals.

To explore the effect of this, we use a foreground segmentation method, as used for HeLa cells in Section 4.3. The same approach is used, generating a probability map, and finding a final segmentation using graph cuts. The only difference is that the probability map is obtained from a random forest classifier Breiman (2001). We use a forest of 5 trees, with a minimum of 500 observations per leaf to limit the depth of the tree and avoid overfitting.

The result of this new segmentation can be seen in Figure 7.2(b). This generates an overall foreground mask, not individual cell segmentations, as groups of cells which are in contact are segmented together. The segmented foreground is then excluded from the interest regions.



(a) (b)

Figure 7.2: Cell masking stage. (a) Cell segmentation from MSER based detection stage. Mitotic cell is undersegmented. (b) Segmentation after using Random Forest classifier. Mitotic cell is now correctly segmented, and additional interphase cell is segmented. Adding this step prevents this (false positive) image from being ranked as the least unlikely.

7.3.3 GMM fitting & image ranking

A feature vector consisting of one of the three sets of filter responses, one of the position representations, and the raw pixel intensity is evaluated for every pixel in the interest region (except those masked by the segmentation), in every image in the dataset. The interest regions are as described in Section 6.2.3, and Figure 6.3. Additional examples of these regions can be seen in Figure 7.4, where the effect of the cell masking can also be seen.

Before fitting, each feature dimension is normalised to have zero mean and unit standard deviation. A three-component Gaussian mixture is used. Both full and diagonal covariance Gaussians are tested.

For image ranking, features are evaluated in the interest region as before, and the probability of every pixel under the GMM is evaluated. Images are then ranked by the minimum probability within the interest region.

7.3.4 Evaluation measures

Evaluation of our method is difficult due to the nature of the problem. Whereas the approach in Chapter 6 is specifically engineered to detect a certain type of defect and can be assessed based on its performance, the method here simply aims to find unlikely events. These should include misaligned chromosome defects, but can also include other unusual events for which there is no ground truth labelling.

We use average precision (AP) as a performance measure, as well as precision of the top 5, 10 and 25 results. While AP gives a measure of overall performance, the other measures ensure that the desired detections are ranked highly.

7.4 Results

Results for various experiments are shown in Table 7.1. Overall, the metaphase-only experiments return higher precision results in the top 25 results, but average precision suffers as some true positives are not detected. As seen from the first two rows, the addition of the background cell masking gives a substantial performance

Filter			Position		Covariance		Cell	Metaphase Result				Three phase result			
LoG	MR	S	(x, y)	(r, θ)	Full	Diag.	Masking	AP	P@5	P@10	P@25	AP	P@5	P@10	P@25
■			■			■		0.044	0.000	0.000	0.080	0.032	0.000	0.000	0.000
■			■			■	■	0.159	0.600	0.500	0.200	0.127	0.400	0.200	0.200
■				■		■	■	0.159	0.600	0.500	0.240	0.131	0.400	0.200	0.160
	■			■		■	■	0.110	0.600	0.500	0.240	0.087	0.000	0.200	0.120
		■		■		■	■	0.022	0.000	0.000	0.040	0.029	0.000	0.000	0.000
■				■	■		■	0.137	0.600	0.400	0.280	0.088	0.200	0.200	0.160
	■			■	■		■	0.040	0.000	0.000	0.160	0.024	0.000	0.000	0.000
		■		■	■		■	0.103	0.400	0.500	0.200	0.112	0.400	0.300	0.240

Table 7.1: Retrieval results. Retrieval scores for various experimental combinations. The '■' symbol indicates which features are used. Average precision (AP) and precision of the top N retrieval results (P@N) are shown.

improvement. The image shown in Figure 7.2 is the top ranked result if the segmentation step is not used, as the unsegmented part is unlikely under the GMM. Changing from (x, y) to (r, θ) position representation (rows 2-3) gives a slight improvement in AP. Overall, the LoG filters give the best performance, but this is decreased by switching to a full covariance. The Schmid filter case is the most improved by using full covariances, possibly due to the higher dimensionality of the features.

Figure 7.4 shows the top ten retrieval results for the case where we use the Schmid filters, polar position representation and use only metaphase cells. The results are reasonable, with five true positives. False positives occur in regions where the cell masking is not effective, near the image boundary. They also occur in a case where

two metaphase cells are very close to each other. This only occurs once in the dataset, where metaphase cells are generally more spatially separate. The region between the cells is the cause of the two false positives (one for each cell). This is an example of an unlikely appearance which is detected, but not labelled in the ground truth.

We also compare the approach of this chapter to the one used in Chapter 6. The ROC curve is shown in Figure 7.3, N.B. this is an ROC curve for frame-level detections, not track-level as shown in Figure 6.4. As expected, the discriminative approach performs better at low values of false positive rate (FPR), as it is specifically designed to discover the mitotic defects. However, that approach is limited by the blob detector stage, and so true positive rate (TPR) does not exceed 0.65, as the remaining defects are not detected as blobs (however, as seen from Figure 6.4, at least one defect frame is detected for most tracks containing them). The generative approach retrieves false positives more quickly. However, as seen from Figure 7.4, this is caused by unlikely occurrences other than defects, which is to be expected in this context.

7.5 Summary

In this chapter we investigated an unsupervised approach to discovering mitotic defects. By fitting a Gaussian mixture to image features evaluated around large

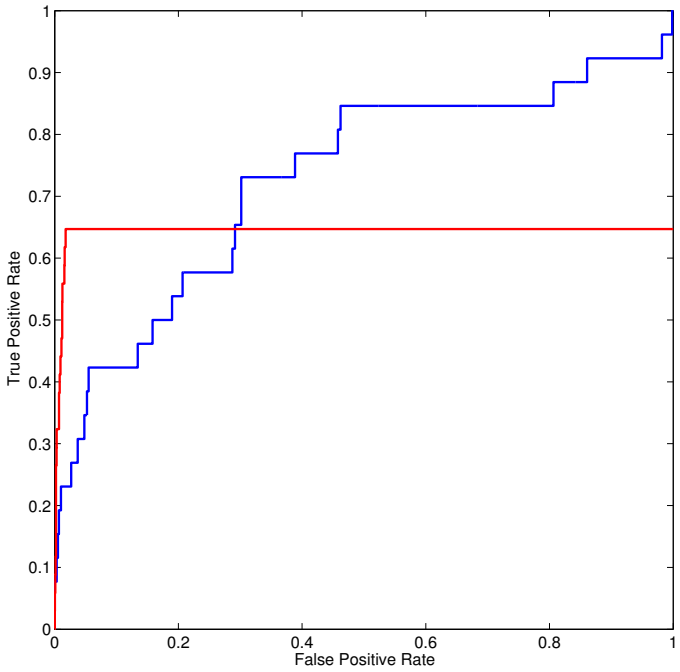


Figure 7.3: ROC curve of defective mitotic frame detection. The blue line shows results for the unsupervised approach of this chapter. The red line shows the performance of the discriminative approach of Chapter 6.

numbers of cells we learn the likely appearance of mitotic cell surroundings. Images which contain regions with low probability under this model can be identified as likely defects.

Within a retrieval framework, this approach does identify mitotic defects, but other events which occur infrequently in the data are also highly ranked. Further investigation is necessary to either improve features to improve the precision of retrieval for specific defects, or to better define what constitutes unlikely behaviour during mitosis.

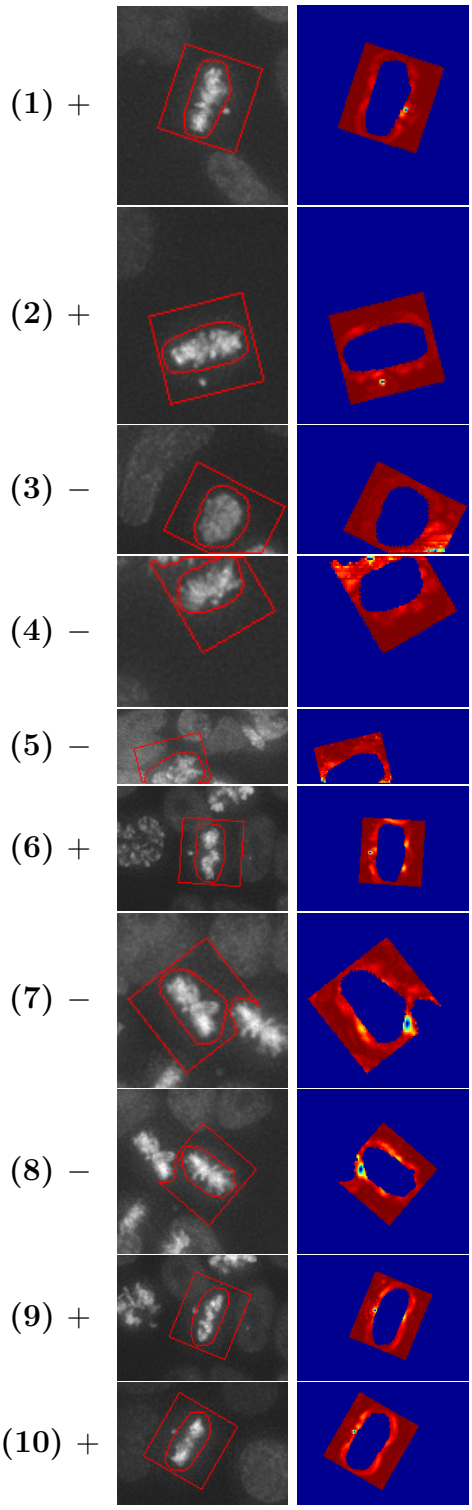


Figure 7.4: Top retrieval results. The top ten least likely metaphase cells according to our model. True positives are indicated with '+', and false positives are indicated with '−'. Left: Original images, with interest regions drawn in red. Right: log-probability map for the interest region; red indicates more likely regions. Yellow/green are lower probability. The majority blue is the ignored region. There are five true positives (1, 2, 6, 9, 10). False positives occur near image boundaries, where cell segmentation for masking is less reliable (3, 4, 5). False positives 7 and 8 are caused by the region between two nearby mitotic cells, of which this is the only instance in the dataset (mitotic cells are generally better separated in space and time).

CHAPTER 8

AUTOMATED CELL VOLUME ANALYSIS

Understanding the genes involved in cell volume regulation is an important area in biological research. Modern screens involving RNA interference often involve analysing many thousands of cells, and manual techniques are too slow.

In this chapter we describe a fully automated method for estimating the radius of spherical mitotic cells (metaphase), within given bounds, based on a learned membrane classifier and a circular Hough transform. The method is suited to high-throughput analysis, and its performance is evaluated on a set of manually annotated image sequences.

The method has been used in a large scale study to determine the effects of suppression of the CLCN2 gene, which encodes a chloride channel protein, on cell volume

response to osmotic shock, and has led to interesting preliminary results (unpublished work). It is also being used to test the effects of other genes on osmotic shock response.

8.1 Introduction

In this chapter we present a fully automated method for estimating the radius of spherical mitotic (metaphase) cells, within given bounds, based on a learned membrane classifier and a circular Hough transform. This method is used to track volume changes in mitotic cells over time, after osmotic shock. The approach is also applied in an experiment to assess the effect of depleting the CLCN2 gene on volume response after osmotic shock leading to the discovery of an unexpected phenotype.

The approach we take to estimating cell volume relies on the assumption that cells become spherical during mitosis, with a circular cross-section. Stewart et al. (2011) show that during mitosis, a loss of adhesion to the surface they are on leads to cells becoming more round. They show that in metaphase cells have a height-to-width ratio of 0.86 ± 0.04, and as seen in Figure 8.3 and other figures throughout the chapter, the cross-sections are circular. For the purposes of the biological goal of assessing response to osmotic shock the spherical assumption is sufficiently accurate.

8.1.1 Related work

To our knowledge, this is the only work attempting to solve this specific problem in biology. However, there are many works which aim to fit circles in noisy data for various applications, e.g. for quality inspection of mechanical parts (Landau, 1987), or fitting circular trajectories of particles (Karimäki, 1991). A full review is beyond the scope of this chapter, but here we briefly describe the two broad categories of existing methods: maximum likelihood estimation (MLE) based, and voting based.

MLE approaches fit circles to a set of points, typically xy coordinates, in a least squares sense e.g. Gander et al. (1994); Landau (1987). These kinds of approaches rely on the set of points being detected by some other algorithm. These approaches are very sensitive to noise, and can struggle with image occlusion (leading to missing points on the circle).

The most common voting based method is the Circular Hough Transform (Duda and Hart, 1972). For this approach, image pixels' votes are stored in a 3-dimensional accumulator array (xy coordinates and radius of the circle), and maxima in the accumulator represent circles in the image. Votes can be based on detected image edges, raw pixel intensities, image gradients, or any other appropriate measure, such as the membrane classifier we use here. This approach is more robust to noise and occlusion than MLE, as well as having an obvious way of detecting multiple circles (multiple peaks in accumulator space). However, the Circular Hough Transform

can be slow and have high memory requirements due to the size of the accumulator. Many other voting based work are typically variations of this which aim to reduce the computational requirements, for example, Xu et al. (1990) use a randomly selected subset of image points for voting.

8.1.2 Biological Motivation

Homeostasis of volume is essential for cell function and survival, making the study of the genes involved in cell volume regulation an important area in biological research. To study this problem comprehensively using modern RNA interference screens requires individually analysing the volume of many thousands of cells, for which current manual techniques are too slow. Therefore, automated techniques are required that allow high-throughput measurement of cell volume at the level of single cells.

Chapter Outline

In Section 8.2 we describe the method used to estimate cell volume, consisting of a learned membrane detector and circular Hough transform. Section 8.3 describes the post-processing steps used, false positive suppression and tracking. The datasets used for experiments are described in Section 8.4, along with associated train/test splits. Quantitative results of the method are shown in Section 8.5.

Finally, some example applications of the method are described in Section 8.6, first in a complete small scale experiment and then in an ongoing large scale experiment.

8.2 Volume Estimation

For volume estimation we assume that mitotic cells are spherical (Stewart et al., 2011), appearing circular in the maximum z-projection images being used. We use a two stage approach to estimate cell size, first using a learned classifier to generate a probability map of the cell boundary, then using a circular Hough transform to find circular regions in this probability map.

8.2.1 Membrane Detection

As can be seen from Figure 8.6, the appearance of cells in the dataset is defined by their boundaries, with image statistics inside and outside the boundaries being very similar. Therefore, methods which rely on such statistics are not suited to this task e.g. Boykov and Jolly (2001). The boundary is detected using a logistic regression classifier (Bishop, 2006), $f(\mathbf{x}) = \sigma(\mathbf{w}^\top \mathbf{x}) + b$, to give a probability of any pixel being on the boundary. The feature vector, $\mathbf{x}$, used for this consists of the pixel intensity values and gradient magnitude values in a 31×31 window centered on the test pixel giving a 1922 dimensional feature. Before processing, pixel values are normalised on

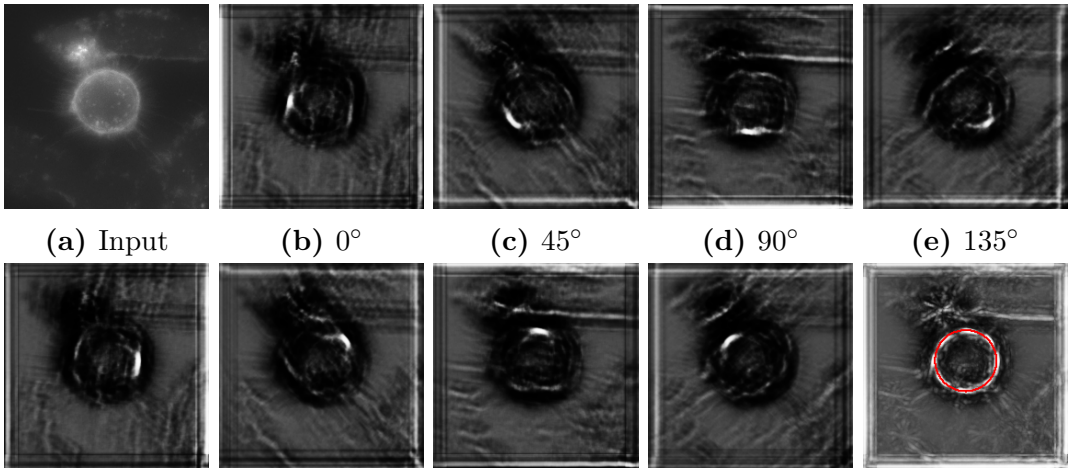
a per-sequence basis to lie in the range $[0, 1]$. Gradients are computed by convolving with the derivatives of a 7×7, $\sigma = 1$ gaussian. In training, each window is rotated to align all training examples such that the y-axis is tangential to the membrane as shown in Figure 8.3. During testing, we use the approach of Gonzalez et al. (2009), where every pixel is tested at 8 evenly spaced orientations, as illustrated in Figure 8.1, and the maximum response is taken such that the probability, P, of a pixel being being on the boundary is given by:

$$P = \max_{\theta} f(R^{\theta}\mathbf{x}) \tag{8.1}$$

where R^{θ} represents a rotation by an angle θ.

8.2.1.1 Classifier training

Positive training data points are selected from the pixels on the perimeters of the circular annotations described in Section 8.4. The region within 5 pixels of the perimeter is defined as the unknown region, which is withheld from the training data to account for the thickness of the membrane and small errors in the annotation, and negative training points are selected from the remaining pixels. Training data is selected iteratively in a "bootstrap" strategy (Sung and Poggio, 1998). This consists of training a classifier on a balanced subset of the pixels in the first frame in the



(a) Input **(b)** 0° **(c)** 45° **(d)** 90° **(e)** 135°

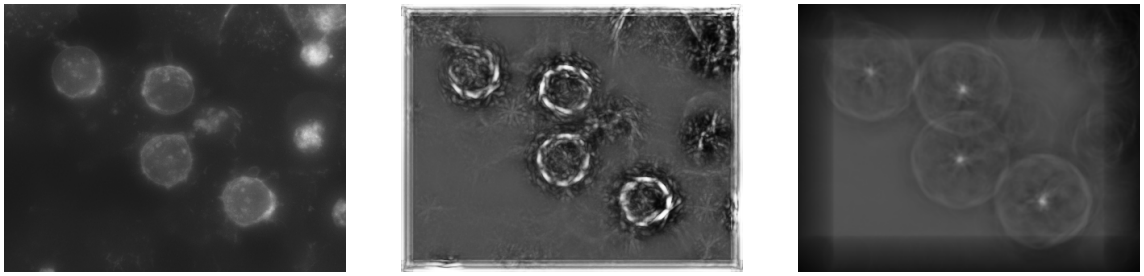
(f) 180° **(g)** 225° **(h)** 270° **(i)** 315° **(j)** Maximum

Figure 8.1: (a) Sample input image from the dataset. (b)-(i) Membrane classifier response at 8 evenly spaced orientations from the horizontal axis. (j) Maximum response over all orientations. The overlaid red curve represents the ground truth annotation.

data, testing on the next frame and adding pixels misclassified with a confidence greater than 75%, repeating this on subsequent frames.

8.2.2 Circular Hough Transform

To detect circles in the membrane detection probability map we use the Hough transform (Duda and Hart, 1972), a form of template matching. For a circle parametrized as $(x-a)^2+(y-b)^2 = r^2$, every point in the probability map, P, contributes "votes" to an accumulator array, A, at all points $A(a,b)$ at a distance r from $P(x,y)$, linearly interpolating where necessary. The votes are weighted by the value of $P(x,y)$ and the array is normalized to lie in the range $[0,1]$ by dividing by $2\pi r$. Peaks in the accumulator array then represent the centers of circles of radius r, as shown in



(a) (b) (c)

Figure 8.2: Hough transform. (a) Input image. (b) Maximum response from membrane classifier over eight orientations. (c) Hough accumulator array for $r = 30$. Local maxima (bright points) in the accumulator are detected as circle centres. The method is robust enough to cope with missing portions in the detected membrane e.g. the top-left cell.

Figure 8.2. Testing over a range of radii gives a three-dimensional accumulator array, $A(a, b, r)$, allowing detection of circles of different sizes by selecting the optimal radius for a given center:

$$r^*(a, b) = \arg\max_r A(a, b, r) \tag{8.2}$$

To detect peaks, each layer of the accumulator is first smoothed with a 7×7, $\sigma = 1$ gaussian filter. Local maxima above a threshold are extracted, suppressing lower maxima within a distance r to prevent overlapping circles.

The Hough transform is evaluated over a range of radii $[25, 45]$ in increments of 0.5, which is roughly 5 pixels beyond the expected true values. To obtain a sub-pixel estimate of the radii, a signal consisting of the values of the accumulator across all radii at a center is extracted. A parabola is then fitted to the values greater than

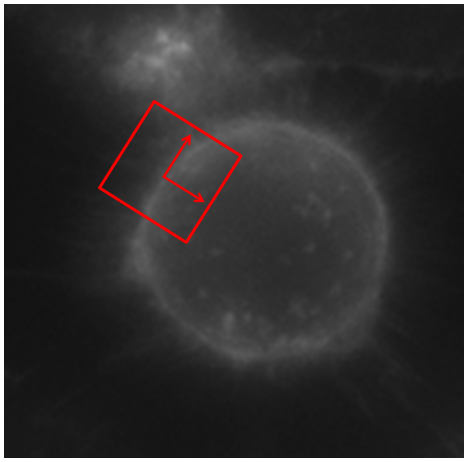


Figure 8.3: Membrane classifier feature. Features consist of pixel and gradient intensity values in a 31×31 pixel window around the test point. The window is aligned such that the y-axis is tangential to the cell membrane.

70% of the maximum of this signal, as illustrated in Figure 8.4, and the maximum of the parabola is used as the prediction of the radius. Upper and lower bounds of this estimate are given by the region in which the parabola is above 90% of the maximum value.

8.3 Post-processing

After detecting the circular regions in an image sequence, we implement two post-processing steps, first suppressing false positive detections, then tracking cells through the sequence.

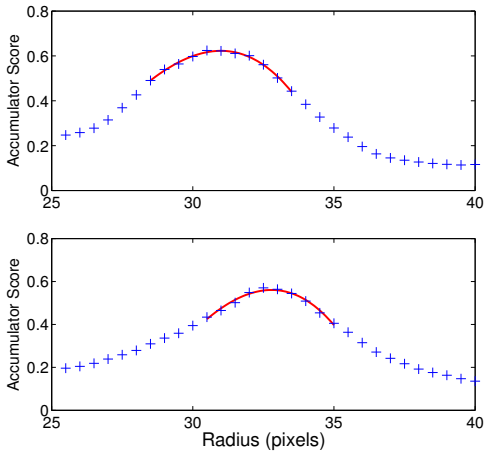
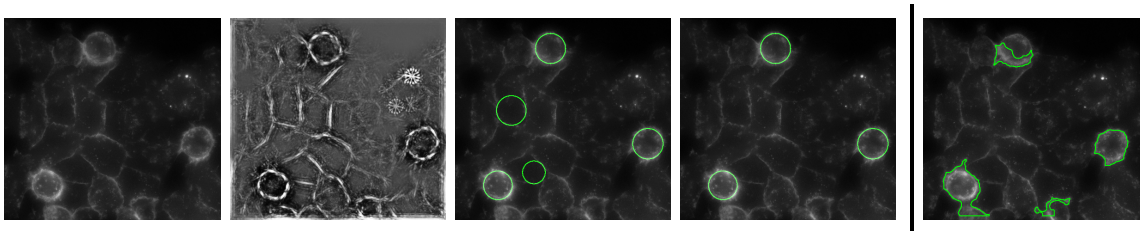


Figure 8.4: Estimating the radius of detected circles. The blue '+' symbols are the values of the accumulator at tested radii. The red curve is the parabola fitted to points above 70% of the maximum value. The parabola maximum is used as the final estimate of the radius.

8.3.1 Short arc removal

The magnitude of the modes in the Hough accumulator is given by mean value of the membrane detector response along the perimeter of the corresponding circle in the image. This means that peaks can be generated by non-circular cells, accidental alignments for empty regions between cells, or other noise in the image if the membrane response is sufficiently high. To suppress such false positives we enforce that all detected circles have at least one continuous arc spanning over 120° with all membrane detector responses above the threshold for local maxima. Figure 8.5 illustrates the benefit of this.



(a) (b) (c) (d) (e)

Figure 8.5: Short arc removal. (a) Input image. (b) Membrane detection response. False positives occur due to membranes of non-mitotic cells. (c) Detected circles from the Hough transform. (d) Detected circles after removing false detections caused by short arcs with strong response. (e) Segmentation using graph cuts delivers inferior results.

8.3.2 Tracking

After detection, cells are tracked using a nearest neighbour match on the circle centres. This simple method assumes that the distance a cell is likely to move from one frame to the next is sufficiently small compared to the distance to other cells. In practice this is the case due to the high temporal resolution under which the cells are imaged and their low motility. However, a simple check for a global shift of all cells is required due to motion caused by adding the medium to induce osmotic shock.

8.4 Data

The dataset consists of a number of image sequences of HeLa cells. To visualise the plasma membrane, HeLa cells stably expressing H2B-EGFP mCherry-CAAX

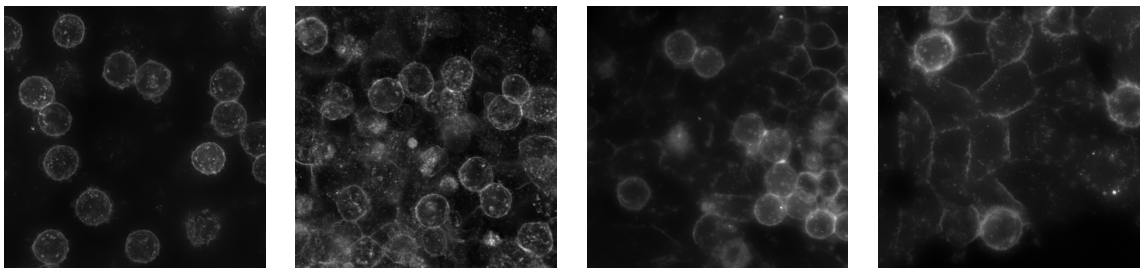
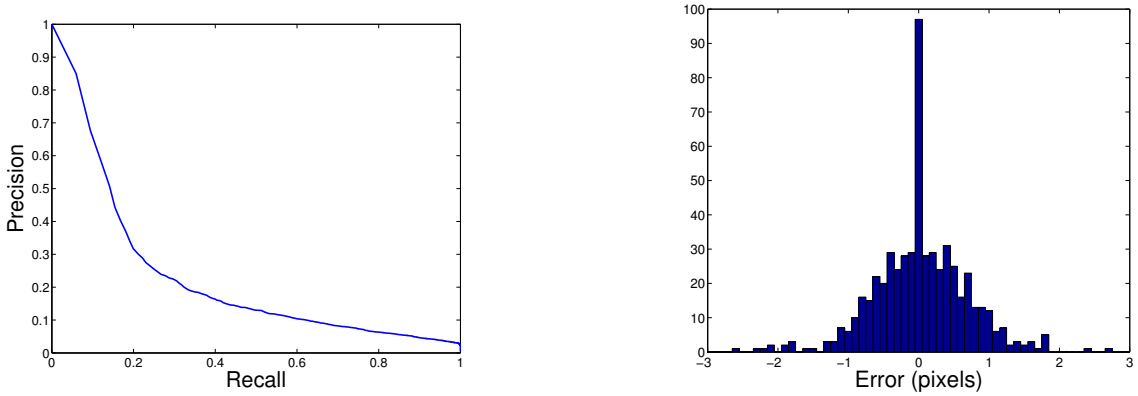


Figure 8.6: Sample images from the dataset. Images are typically about 512×512 pixels.

(clone 2B4) are used. Cells were synchronised in prometaphase with 2 μM S-trityl-L-cysteine (STC) overnight, and mitotic arrest was confirmed by condensed chromosomes. Cells were seeded onto a a 8-well μ-slide chamber (ibidi GmbH) in a CO2-independent medium (gibco) and filmed at 1-minute intervals with 10 z-sections at 1 μm steps so that the largest cross section area of the cells could be collected. For filming we used a DeltaVision system (Applied Precision) equipped with a 40x oil objective and a heat chamber to maintain the temperature at 37°C. The same volume (200 μl) of water and the CO2-independent medium supplemented with 2.87% Xylose was added for hypotonic (-150 mOsm [milliosmole, unit of osmotic pressure]) and hypertonic (+100 mOsm) treatment, respectively. Maximally projected images were used for training and analysis.

The portion of the data used for training and evaluation of the processing system contains 32 sequences of untreated cells, each 10 minutes long, which are manually annotated. The annotation consists of circular regions corresponding to the regions enclosed by the cell membranes, giving positions and sizes of mitotic cells. All



(a) (b)

Figure 8.7: Performance evaluation. (a) Precsion-recall curve for the intermediate membrane pixel classification stage. Precision drops quickly due to false detections of membranes in non spherical cells. (b) Histogram of absolute error in radius estimation, in pixels. Errors fit a zero-mean normal distribution with standard deviation 0.65.

mitotic cells are annotated. The sequences are divided into a 10 sequence training set, 6 sequence validation set, and a 16 sequence test set for, evaluation of the visual processing system.

The system is also applied to a much larger additional set of hour-long sequences of RNAi treated cells to examine the effects of suppressing the various genes on cell volume response after osmotic shock as described in Section 8.6.

8.5 Results

The membrane classifier is initially trained on the training set and tested on the validation set, as described in Section 8.4, to select the best performing value of the

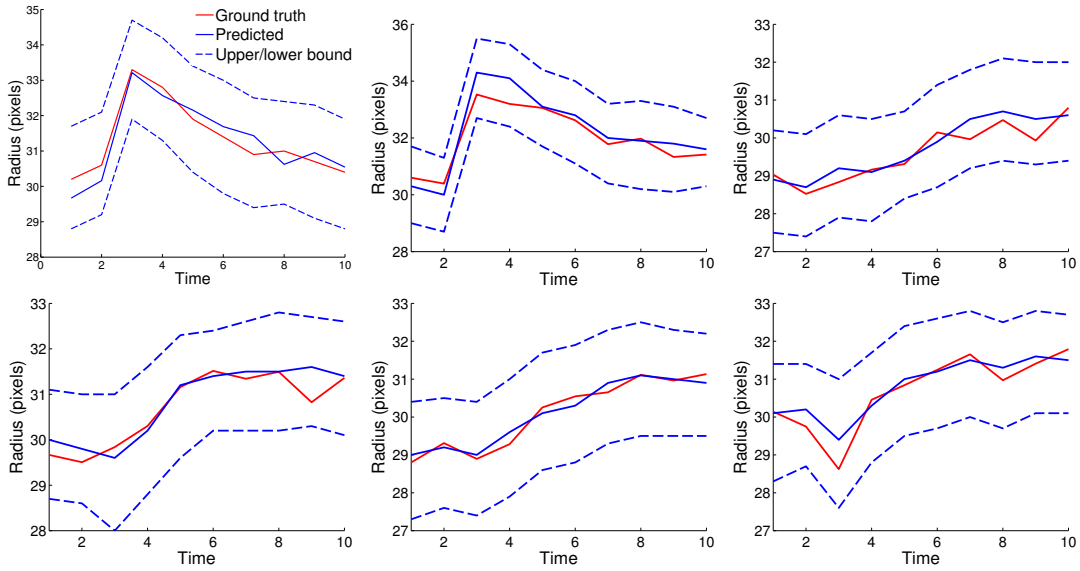


Figure 8.8: Results for individual cells over time. The red line is the manually annotated radius. The solid blue line is the prediction and the two dashed lines are the upper and lower bounds on the prediction. The ground truth is consistently within the bounds, which generally have a range of 3 pixels. The automated annotation generally produces a smoother estimate of radius over time than manual annotation.

regularisation parameter of the classifier. Figure 8.7(a) shows the precision recall curve for the classifier. It can be seen that precision drops quickly with increasing recall. This is due to false detections of membranes from non-spherical cells that are inevitable given the local processing of the classifier. However, this is not a problem as the subsequent Hough transform and short arc removal are sufficiently robust to reject these false positive pixels. Performance on the validation set is similarly used to select a threshold for maxima in the Hough accumulator. The membrane classifier is then trained on the combined training and validation sets and the whole system tested on the test set.

For evaluation of the full method on the test set, cells detected within 5 pixels of the ground truth are counted as true positives and any other detections are considered false positive. Using these criteria, we obtain a detection precision value of 0.998 and recall of 0.951 showing that our method reliably detects the mitotic cells with minimal false detections.

Figure 8.7(b) shows a histogram of errors, in pixels, in radius estimation relative to the ground truth for correctly detected cells. These fit a zero-mean normal distribution with standard deviation 0.65. This means that in the majority of cases the true value of the radius lies within the bounds on the prediction, as illustrated in Figure 8.8.

For comparison, a region based segmentation approach is also tested, using the output of a logistic regression classifier for foreground (cell interior) vs. background learned on the training data as the unary term for a graph cut (as used in Section 4.3). As seen in Figure 8.5(e), this performs poorly with a detection precision of 0.616 and recall of 0.623. The mean error in area estimation for correctly detected cells is 119%, due to frequent under-segmentation of adjacent cells, in contrast to the value of 0.32% using our learned membrane Hough based approach. This shows that a region based approach is not suitable for this task, and that a more specific solution such as our approach is required.

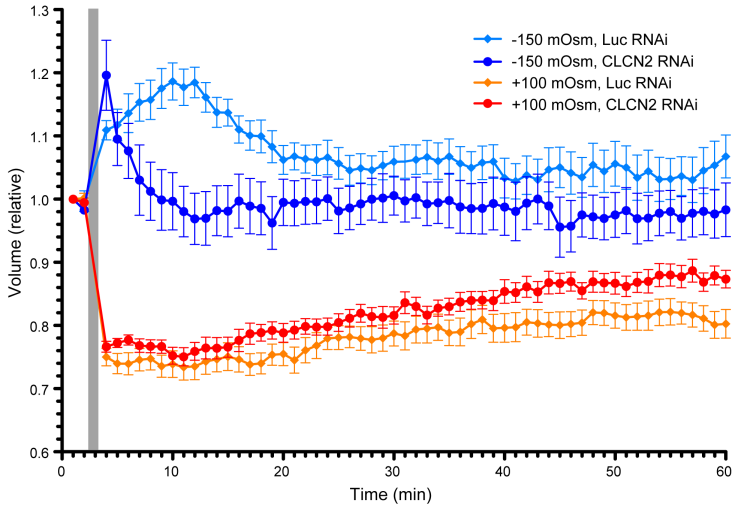


Figure 8.9: NavyBlue **Cell volume changes upon hypotonic and hypertonic shock.** Mean ± SEM is shown. Osmotic shock applied after $t = 2$ (grey). RNAi-treated cells (dark blue, red). Volume change magnitudes after shock are similar to the control (Luciferase RNAi; pale blue, orange) cells, but eventually reach equilibria closer to starting volume.

8.6 Applications

8.6.1 Response to Osmotic Shock after RNAi of a Chloride Channel

The system has been applied to a large set of sequences, containing approximately 100 mitotic cells. The purpose of this experiment is to examine the effects of suppressing CLCN2, a gene encoding a chloride ion channel, on cell volume response to osmotic shock. This experiment consists of testing a set of sequences where expression of the CLCN2 gene has been suppressed by RNAi, allowing a comparison of the cell volume response to the control.

Estimated volumes are normalised on a per-cell basis and plotted with standard error (SEM). The results of this experiment are shown Figure 8.9, showing volume response for both hypertonic .

CLCN2 is used as it is known to be a swelling-activated chloride channel that regulates cell volume (Grunder et al., 1992; Huang et al., 2009; Huber et al., 2004) and to be expressed in mitosis (Zheng et al., 2002), but its role in mitotic cell volume control is not fully understood. As seen in Figure 8.9, the extent of volume change upon osmotic challenge is similar, but the equilibrium volume (last five frames) of the CLCN2 RNAi cells is nearer to the original volume after either hypotonic (-150 mOsm, $p = 0.002$) or hypertonic (+100 mOsm, $p < 0.001$) treatment. Both volume increase and recovery upon the hypotonic shock are particularly quick in CLCN2 RNAi cells, which is unexpected.

This is an interesting preliminary result, and further investigation may lead to the discovery of unknown functions of CLCN2.

8.6.2 Larger scale experiments

Similar experiments were conducted on much larger scale data, consisting of more than 4000 cells, each with one of 101 different genes suppressed by RNAi. Finding relationships within these and determining which genes control volume regulation is still an open area of biological research. We approach this as a retrieval problem,

finding genes with the most similar volume response to a query gene. The retrieval approach provides an alternative to the clustering approaches commonly used in these kinds of experiments, which can enforce arbitrary associations between genes if clustering algorithms reach local minima.

The full signals are used as feature vectors, and euclidean distance used to rank results. As the initial volume change and recovery after osmotic shock occurs within a relatively small proportion of the duration of the signal, distances between signals are Gaussian weighted to place more emphasis on the early change.

Figure 8.10 shows the retrieval results for several genes. The figure shows that the retrieval system consistently returns genes with similar responses, especially in the initial region prioritised by the Gaussian weighting. For each query, different genes are retrieved for hypertonic and hypotonic shock, which may suggest that genes respond differently to different kinds of shock.

8.7 Summary

We have presented a fully automated method for estimating the volume of spherical mitotic cells, within given bounds, based on a machine learned membrane classifier and a circular Hough transform. The method is shown to perform well on a test set of sequences tracking volume changes in mitotic cells over time, after osmotic shock.

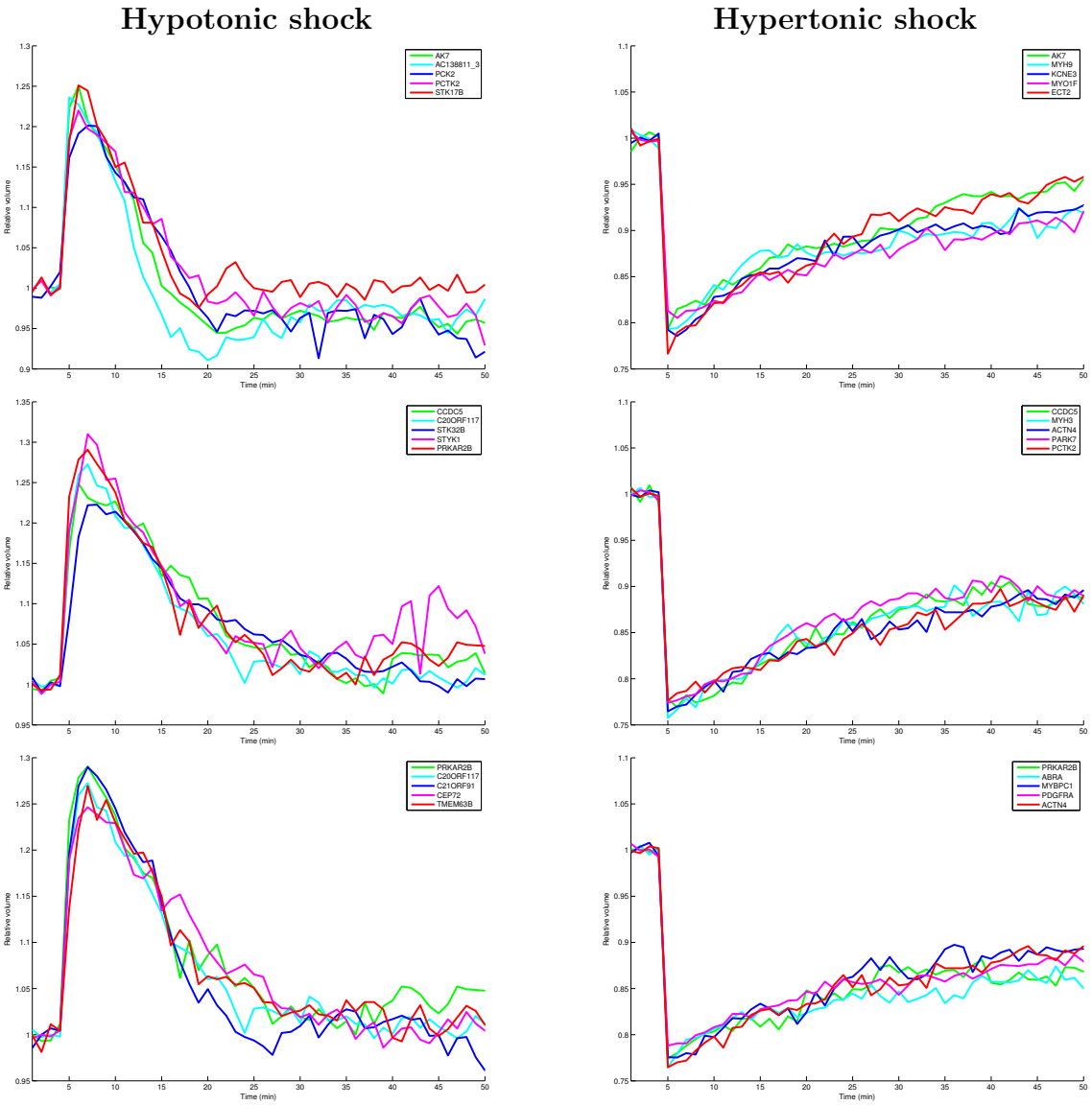


Figure 8.10: Volume response retrieval. Example retrieval results for three different genes' response to osmotic shock. Green: query gene. Other colours: four genes with most similar responses sorted by order of similarity in the legend. Left: hypotonic shock. Right: hypertonic shock.

The method is easily scalable, using multiple cores to process the images in parallel, making it suitable for high-throughput analysis for screens of thousands of cells.

The method is also applied in an experiment to determine the effects of suppression of the CLCN2 gene on cell volume response to osmotic shock, leading to interesting preliminary results. It has further been applied to much larger scale data, with the aim of determining the genes responsible for cell volume control (unpublished work). For this application we describe a simple retrieval system for finding genes with similar volume responses to osmotic shock.

This method is restricted to cases where cells are in prometaphase or metaphase and the spherical assumption is acceptable. Our approach will no longer work if cells progress to anaphase, and so is not to analysis of cell size changes throughout mitosis. Furthermore, while the Hough transform is reasonably robust to image noise and can cope with cells touching (groups of up to three adjacent round cells in our dataset), increased cell density could potentially cause some errors.

CHAPTER 9

CONCLUSIONS

In this chapter, we summarise the contributions of this thesis and discuss potential areas of improvement through further research.

9.1 Contributions

This thesis has addressed three areas of biological image analysis: (i) Mitotic phase labelling, (ii) Mitotic defect detection, and (iii) Cell volume estimation.

9.1.1 Mitotic phase labelling

In the first part of our work, in Chapters 4 and 5, we developed an automated system for mitotic phase labelling, taking the novel approach of using temporal features evaluated over the whole of the mitotic phases rather than over single frames, thereby capturing the distinctive behaviour over the phases. We described and compared three machine learned temporal models for mitotic phase labelling: (i) aligning temporal signals of simple features directly to a reference signal using Dynamic Time Warping; (ii) learning a Hidden Markov Model from feature signals; (iii) a discriminative Semi-Markov Model incorporating information over temporal segments.

We also introduced an alternative SMM loss function to the original of Shi et al. (2011) which is robust to inconsistencies in training data annotation at phase transition boundaries. This new loss function outperforms the original in this context. Using this loss function, the SMM approach resulted in better labellings than an independent frame classification approach using a significantly larger feature set.

Additionally, we showed that the Semi-Markov model's segment based approach is more robust to differences between the training and testing datasets than other models, due to it not explicitly modelling phase duration. Considering variations of features over whole mitotic phases allowing the labelling to cope with changes in experimental and imaging conditions.

This robustness means that the method is suitable for use in biological research experiments with minimal interaction from users, as the model can simply be trained once on a large dataset and will still give high quality labellings on new data. In addition to being directly applied to experiments investigating mitotic progression, the labellings can also be used as a basis for further analysis of cell behaviour within phases, e.g. as we do in Chapters 6 and 7.

In addition to the phase labelling methods, we presented a novel segmentation approach for fluorescence images of histone GFP tagged cells. To complete the overall system we also include a simple tracking method and a classifier for determining whether or not a given cell track contains a mitosis event.

9.1.2 Mitotic defect detection

In Chapter 6, we presented a novel method for detecting two specific types of very rare and subtle mitotic errors in large volumes of microscopy images. The system relies on detecting and tracking only the subset of cells in the data which undergo mitosis to reduce the number of cells to analyse. Candidate defects are proposed, and subsequently filtered, based on size and appearance, as well as by mitotic phase as labelled by out system from Chapter 4. Despite the low precision of predicted defects, our method is useful as a filter for further biological analysis, as it reduces the amount of data to be analysed manually by several orders of magnitude.

We investigated an unsupervised approach to this problem in Chapter 7. By fitting a Gaussian mixture model to image features evaluated around detected mitotic cells, a model for normal cell surroundings is learned. Images which contain regions with low probability under this model can be identified as likely defects. Within a retrieval framework, this approach highly ranked mitotic defects, along with other events which occur infrequently in the data. With further work this method may be used to find more general deviations from normal mitosis than the specific ones focused on here.

9.1.3 Cell volume estimation

In Chapter 8, we presented a fully automated method for estimating the volume of spherical mitotic cells, within given bounds, based on a machine learned membrane classifier and a circular Hough transform. The method is shown to perform well, with zero-mean error with standard deviation less than one pixel, on a test set of sequences tracking volume changes in mitotic cells over time, after osmotic shock.

The method is easily scalable, using multiple cores to process the images in parallel, making it suitable for high-throughput analysis for screens of thousands of cells, and has already been applied to two experiments.

The first is an experiment, with around 100 cells measured by our system, to determine the effects of suppression of the CLCN2 gene on cell volume response to

osmotic shock, leading to the discovery of an unexpected phenotype. It has further been applied to much larger scale data, around 4000 cells tested and 101 different genes suppressed with RNAi, with the long term aim of fully determining the genes responsible for cell volume control.

9.2 Further Work

Biological image analysis is a constantly changing field and is strongly affected by changes in imaging techniques, image quality and cell culture conditions. However, the methods presented in this work are sufficiently general to be applicable to other imaging conditions with minimal adaptation. In this section we suggest areas of potential further development to our work to improve performance.

9.2.1 Mitotic phase labelling

The most simple extension of these methods would be to explore the effects of additional features on performance. Experiments with DTW and HMM using simple size and shape signals showed improved performance, but due to the nature of the encoding used for SMM the memory requirements quickly become prohibitive as the number of features increases.

Looking into alternative feature encodings for the SMM is another potential area of improvement. There are a range of potential duration invariant ways of representing the raw feature signal e.g. fitting a polynomial to the signal and using the coefficients as the feature encoding. It may also be worthwhile investigating the performance of the SMM with explicit duration modelling, e.g. by concatenating a value of $e^{\frac{-(d-\mu)^2}{\sigma^2}}$ to the feature vector, where d is the phase duration and μ and σ are the mean and standard deviation of the phase duratation as measured from training data. While this may reduce the robustness of the method, performance for similar train and test sets may be improved.

Finally, the area with the most potential for development of the ideas in this work is to adopt an unsupervised approach to learning mitotic phases. This has recently been done at the individual frame level by Zhong et al. (2012), but extending this to temporal segments in a similar way to our approach to supervised labelling methods could potentially yield similar performance improvements.

9.2.2 Mitotic defect detection

There are a number of possible avenues for developing the mitotic defect detection portion of this work. One potential approach is to develop a supervised machine learning method for detecting the misaligned chromosomes. However, this would

likely need a much larger number of examples of defects for training than are available in the current dataset.

Another possibility would be to expand on the approach or ranking examples based on their likelihood in the data. This could be done using SVM ranking (Joachims, 2002) for example, and could also potentially be done in a semi-supervised manner to reduce the amount of necessary manual annotation.

9.2.3 Cell volume estimation

One potential area of refinement for our volume estimation system, assuming the availability of image sequences with a histone GFP, would be to detect cell nuclei as in Chapters 4 to 6 and restrict the Hough transform computation to the regions surrounding cells, significantly reducing the memory requirements. This would make it feasible to extend the approach to 3D - either by applying the circular Hough transform to each focal plane separately, or using a spherical Hough transform.

It may also be possible to improve results by moving beyond the spherical shape assumption used throughout Chapter 8. A simple way to do this could be to explore the use of an ellipsoidal shape representation which would add some more flexibility.

If more refined measures of shape are needed, the Hough transform approach can be used to determine an initial circular interest region, and then an optimal path can can be found through it e.g. using dynamic programming. This kind of

approach would allow for much finer measurement of cell area/volume, but it is worth considering if there is any biological research requirement for this degree of precision.

9.3 Potential extensions and applications

A simple further application of the phase labelling methods would be to test their applicability to images with different proteins tagged, e.g. as Held et al. (2010) do with tubulin images. This of course is dependant on having reliable segmentation and tracking methods for the new image types.

In sequences with two proteins fluorescently tagged temporal features could either be learned jointly, to improve accuracy, or independently, to determine if delays are caused by a specific protein. The independent approach would be especially interesting in an unsupervised learning context, where normal and delayed phases can be discovered as different modes of cell behaviour.

Combining the phase labelling and defect detection into a unified framework has the potential to give a very powerful model of mitotic behaviour. For example, using an approach like the dynamic topic model of Hospedales et al. (2012), to jointly learn a generative model of both mitotic phases and modes of their appearances could give

better results than the individual methods used here. This would also make it easier to discover deviations from normal behaviour.

Another interesting application would be to combine the phase labelling and volume estimation methods. If the volume estimation method can be generalised to cope with the non-circular appearance of interphase cells, then it would be possible to analyse the behaviour of cell shape over phases.

BIBLIOGRAPHY

A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(23):90 – 126, 2006.

Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Transactions on Biomedical Engineering*, 57(4):841–852, 2010.

B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Essential Cell Biology*. Garland Science, 2009.

R. Ali, M. Gooding, T. Szilágyi, B. Vojnovic, M. Christlieb, and M. Brady. Automatic segmentation of adherent biological cell boundaries and nuclei from bright-field microscopy images. *Machine Vision and Applications*, 23(4):607–621, July 2012. ISSN 0932-8092.

Y. Altun, I. Tsochantaridis, and T. Hofmann. Hidden markov support vector machines. In *Proceedings of the International Conference on Machine Learning*, volume 3, pages 3–10, 2003.

B. Amberg and T. Vetter. GraphTrack: Faster than realtime tracking in videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman. Learning to detect cells using non-overlapping extremal regions. In N. Ayache, editor, *International Conference on Medical Image Computing and Computer Assisted Intervention*, Lecture Notes in Computer Science, pages 348–356. MICCAI, Springer, 2012.

M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188, February 2002.

M. Benzeghiba, R. De Mori, O. Deroo, S. Dupont, T. Erbes, D. Jouvet, L. Fissore, P. Laface, A. Mertins, C. Ris, R. Rose, V. Tyagi, and C. Wellekens. Automatic speech recognition and speech variability: A review. *Speech Communication*, 49 (1011):763 – 786, 2007.

J.-P. Bergeest and K. Rohr. Fast globally optimal segmentation of cells in fluorescence microscopy images. In *Proceedings of the International Conference on*

Medical Image Computing and Computer Assisted Intervention, 2011.

J.-P. Bergeest and K. Rohr. Efficient globally optimal segmentation of cells in fluorescence microscopy images using level sets and convex energy functionals. *Medical Image Analysis*, 16(7):1436 – 1444, 2012. ISSN 1361-8415.

D.S.C. Biggs. 3D deconvolution microscopy. In *Current Protocols in Cytometry*. John Wiley & Sons, Inc., 2001.

A. W. Bird and A. A. Hyman. Building a spindle of the correct length in human cells requires the interaction between tpx2 and aurora a. *The Journal of Cell Biology*, 182(2):289–300, 2008.

C. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 2006.

S. S Blackman. Multiple hypothesis tracking for multiple target tracking. *Aerospace and Electronic Systems Magazine, IEEE*, 19(1):5–18, January 2004.

H. A. K. Blom. An efficient filter for abruptly changing systems. In *Proc. 23rd IEEE Conference Decision Contr.*, pages 656–658, 1984.

Y. Boykov and M. P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada*, volume 2, pages 105–112, 2001.

M. Brand and V. Kettnaker. Discovery and segmentation of activities in video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):844–851, 2000.

L. Breiman. Random forests. *ML Journal*, 45(1):5–32, 2001.

A. Carpenter, T. Jones, M. Lamprecht, C. Clarke, I. Kang, O. Friman, D. Guertin, J. Chang, R. Lindquist, J. Moffat, P. Golland, and D. Sabatini. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology*, 7(10):R100, 2006.

T. F. Chan and L. A. Vese. Active contours without edges. *Image Processing, IEEE Transactions on*, 2001.

C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.

T. Chen, Y. Zhang, C. Wang, Z. Qu, M. Cai, F. Wang, and T. Syeda-Mahmood. Local complex phase based level set and its application to dic red blood cell segmentation. In *IEEE International Symposium on Biomedical Imaging*, pages 187–190, 2011.

X. Chen, X. Zhou, and S.T.-C. Wong. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *Biomedical Engineering, IEEE Transactions on*, 53(4):762–766, 2006.

J. Cheng, M. Veronika, and J. C. Rajapakse. Identifying cells in histopathological images. In *Proceedings of the International Conference on Pattern Recognition*, ICPR'10, pages 244–252, Berlin, Heidelberg, 2010. Springer-Verlag.

L.P. Coelho, A. Shariff, and R.F. Murphy. Nuclear segmentation in microscope cell images: a hand-segmented dataset and comparison of algorithms. In *Proceedings of the Sixth IEEE international conference on Symposium on Biomedical Imaging: From Nano to Macro*, pages 518–521, 2009.

D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5): 603–619, 2002.

D. A. Compton. Mechanisms of aneuploidy. *Current Opinion in Cell Biology*, 2011.

C. Conrad and D. W. Gerlich. Automated microscopy for high-content rnai screening. *The Journal of Cell Biology*, 188(4):453–61, February 2010.

O. Debeir, P. Van Ham, R. Kiss, and C. Decaestecker. Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes. *IEEE Trans Med Imaging*, 24(6):697–711, 2005.

R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications ACM*, 1972.

A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillen-Aghion, J-C Olivo-Marin, and C. Zimmer. Segmenting and tracking fluorescent cells in dynamic 3-d microscopy with coupled active surfaces. *Image Processing, IEEE Transactions on*, 14(9): 1396–1410, 2005.

O. Dzyubachyk, W.A. Van Cappellen, J. Essers, W.J. Niessen, and E. Meijering. Advanced level-set-based cell tracking in time-lapse fluorescence microscopy. *Medical Imaging, IEEE Transactions on*, 29(3):852–867, 2010.

A. El-Labban, A. Zisserman, Y. Toyoda, A. W. Bird, and A. Hyman. Dynamic time warping for automated cell cycle labelling. In *Microscopic Image Analysis with Applications in Biology*, 2011.

A. El-Labban, A. Zisserman, Y. Toyoda, A. W. Bird, and A. Hyman. Discriminative semi-markov models for automated mitotic phase labelling. In *IEEE International Symposium on Biomedical Imaging*, 2012.

A. El-Labban, C. Arteta, A. Zisserman, A. W. Bird, and A. Hyman. Mitotic phase based detection of chromosome segregation errors in embryonic stem cells. In *IEEE International Symposium on Biomedical Imaging*, 2013.

W. Gander, G.H. Golub, and R. Strebel. Least-squares fitting of circles and ellipses. *BIT Numerical Mathematics*, 34(4):558–578, 1994.

A. Genovesio, T. Liedl, V. Emiliani, W.J. Parak, M. Coppey-Moisan, and J.-C. Olivo-Marin. Multiple particle tracking in 3-D+t microscopy: method and application to the tracking of endocytosed quantum dots. *IEEE Transactions on Image Processing*, 15(5):1062–1070, 2006.

G. Gonzalez, F. Fleuret, and P. Fua. Learning rotational features for filament detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

H. Gray. *Anatomy of the Human Body*. Lea & Febiger, 1918. URL http://www.bartleby.com/107/. Bartleby.com, 2000.

S. Grunder, A. Thiemann, M. Pusch, and T. J. Jentsch. Regions involved in the opening of CIC-2 chloride channel by voltage and cell volume. *Nature*, 360, 1992.

R. M. Haralick. Statistical and structural approaches to texture. *Proceedings of IEEE*, 67(5):786–804, May 1979.

N. Harder, F. Mora-Bermúdez, W. Godinez, J. Ellenberg, R. Eils, and K. Rohr. Automated analysis of the mitotic phases of human cells in 3d fluorescence microscopy image sequences. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2006*, volume 4190, pages 840–848, 2006a.

N. Harder, B. Neumann, M. Held, U. Liebel, H. Erfle, J. Ellenberg, R. Eils, and K. Rohr. Automated recognition of mitotic patterns in fluorescence microscopy

images of human cells. In *IEEE International Symposium on Biomedical Imaging*, pages 1016–1019, 2006b.

N. Harder, F. Mora-Bermúdez, W. J. Godinez, A. Wúnsche, R. Eils, J. Ellenberg, and K. Rohr. Automatic analysis of dividing cells in live cell movies to detect mitotic delays and correlate phenotypes in time. *Genome Research*, 19(11):2113–24, 2009.

M. Held, M. H. A. Schmitz, B. Fischer, T. Walter, B. Neumann, M. H. Olma, M. Peter, J. Ellenberg, and D. W. Gerlich. Cellcognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods*, August 2010.

T. Hofmann. Probabilistic latent semantic indexing. In *Special Interest Group on Information Retrieval*, 1999.

T. Hospedales, S. Gong, and T. Xiang. Video behaviour mining using a dynamic topic model. *International Journal of Computer Vision*, 98(3):303–323, 2012.

Z. M. Huang, C. Prasad, F. C. Britton, L. L. Ye, W. J. Hatton, and D. Duan. Functional role of CLC-2 chloride inward rectifier channels in cardiac sinoatrial nodal pacemaker cells. *J. Mol. Cell. Cardiol.*, 47:121–132, Jul 2009.

S. M. Huber, C. Duranton, G. Henke, C. Van De Sand, V. Heussler, E. Shumilina, C. D. Sandu, V. Tanneur, V. Brand, R. S. Kasinathan, K. S. Lang, P. G. Kremsner, C. A. Hubner, M. B. Rust, K. Dedek, T. J. Jentsch, and F. Lang. Plasmod-

ium induces swelling-activated ClC-2 anion channels in the host erythrocyte. *J. Biol. Chem.*, 279:41444–41452, Oct 2004.

N. İkizler and D. A Forsyth. Searching for complex human activities with no visual examples. *International Journal of Computer Vision*, 80(3):337–357, 2008.

M. Isard and A. Blake. Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29:5–28, 1998.

T. Joachims. Optimizing search engines using clickthrough data. In *Proc. KDD*, pages 133–142, 2002.

T.R. Jones, A.E. Carpenter, M.R. Lamprecht, J. Moffat, S.J. Silver, Jennifer K. Grenier, A.B. Castoreno, U.S. Eggert, D.E. Root, P. Golland, and D.M. Sabatini. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proceedings of the National Academy of Sciences*, 106 (6):1826–1831, 2009.

N.N. Kachouie, P. Fieguth, J. Ramunas, and E. Jervis. Probabilistic model-based cell tracking. *International Journal of Biomedical Imaging*, 2006.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions A.S.M.E., Journal of Basic Engineering*, pages 35–45, March 1960.

V. Karimäki. Effective circle fitting for particle trajectories. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 305(1):187–191, 1991.

M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 1987.

U.M. Landau. Estimation of a circular arc center and its radius. *Computer Vision, Graphics, and Image Processing*, 38(3):317–326, 1987.

C. Li, C.-Y. Kao, J.C. Gore, and Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *IEEE Transactions On Image Processing*, 2008a.

K. Li, E. Miller, L. Weiss, P. Campbell, and K. Kanade. Online tracking of migrating and proliferating cells imaged with phase-contrast microscopy. In *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW '06)*, pages 65 – 72, June 2006.

K. Li, M. Chen, T. Kanade, E. Miller, L. Lee Weiss, and P. Campbell. Cell population tracking and lineage construction with spatiotemporal context. *Medical Image Analysis*, 12(5):546 – 566, October 2008b.

X. Lou and F. A. Hamprecht. Structured learning for cell tracking. In J. Shawe-Taylor, R.S. Zemel, P. Bartlett, F.C.N. Pereira, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 1296–1304. 2011.

X. Lou, U. Koethe, J. Wittbrodt, and F.A. Hamprecht. Learning to segment dense cell nuclei with shape prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1012–1018, 2012.

K.E.G. Magnusson and J. Jalden. A batch algorithm using iterative application of the Viterbi algorithm to track cells and construct cell lineages. In *IEEE International Symposium on Biomedical Imaging*, pages 382–385, 2012.

S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11: 674–693, 1989.

K.Z. Mao, P. Zhao, and P.-H. Tan. Supervised learning-based cell image segmentation for P53 immunohistochemistry. *Biomedical Engineering, IEEE Transactions on*, 53(6):1153–1163, 2006.

M. Marcuzzo, P. Quelhas, A. Campilho, A. Maria Mendonça, and A. Campilho. Automated arabidopsis plant root cell segmentation based on svm classification and region merging. *Computers in Biology and Medicine*, 39(9):785–793, September 2009.

J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proceedings of the British Machine Vision Conference*, pages 384–393, 2002.

S.K. Nath, K. Palaniappan, and F. Bunyak. Cell segmentation using coupled level sets and graph-vertex coloring. In *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, MICCAI'06, pages 101–108, Berlin, Heidelberg, 2006. Springer-Verlag.

B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, and J. Ellenberg. High-throughput rnai screening by time-lapse imaging of live human cells. *Nature methods*, 3(5):385, 2006.

M. Ostendorf, V.V. Digalakis, and O.A. Kimball. From HMM's to segment models: a unified view of stochastic modeling for speech recognition. *Speech and Audio Processing, IEEE Transactions on*, 4(5):360 –378, sep 1996.

N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

D. Padfield, J. Rittscher, N. Thomas, and B. Roysam. Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. *Medical Image Analysis*, 13(1):143 – 155, 2009. ISSN 1361-8415.

D. Padfield, J. Rittscher, and B. Roysam. Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis. *Medical Image Analysis*, 15(4):650 – 668, 2011.

J. Pan, T. Kanade, and M. Chen. Learning to detect different types of cells under phase contrast microscopy. In *Microscopic Image Analysis with Applications in Biology (MIAAB) 2009*, September 2009.

H. Peng, X. Zhou, F. Li, X. Xia, and S.T.C. Wong. Integrating multi-scale blob/curvilinear detector techniques and multilevel sets for automated segmentation of stem cell images. In *IEEE International Symposium on Biomedical Imaging*, ISBI'09, pages 1362–1365, Piscataway, NJ, USA, 2009. IEEE Press.

H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally-Optimal Greedy Algorithms for Tracking a Variable Number of Objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

R. Poppe. A survey on vision-based human action recognition. *Image and vision computing*, 28(6):976–990, 2010.

L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*, 77:257–286, 1989.

G. Rätsch and S. Sonnenburg. Large scale hidden semi-markov svms. In *Advances in Neural Information Processing Systems*, pages 1161–1168, 2006.

T. Ridler and S. Calvard. Picture thresholding using an iterative selection method. *Systems, Man and Cybernetics, IEEE Transactions on*, 8(8):630 –632, aug. 1978. ISSN 0018-9472. doi: 10.1109/TSMC.1978.4310039.

J. Rittscher. Characterization of biological processes through automated image analysis. *Annual Review of Biomedical Engineering*, 12(1):315–344, 2010.

J. Rittscher, R. Machiraju, and S.T.C. Wong. *Microscopic image analysis for life science applications*. Artech House, 2008.

M. Rousson and R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. In *Proceedings of the Workshop on Motion and Video Computing*, MOTION '02, Washington, DC, USA, 2002. IEEE Computer Society.

S. Saito, K. Morita, A. Kohara, T. Masui, M. Sasao, H. Ohgushi, and T. Hirano. Use of BAC array CGH for evaluation of chromosomal stability of clinically used human mesenchymal stem cells and of cancer cell lines. *Human Cell*, 2011.

H. Sakoe. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26:43–49, 1978.

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. Tomanacak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676–682, 2012.

C. Schmid. Constructing models for content-based image retrieval. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 39–45, 2001.

J. A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, Cambridge, 1998.

P. J. Shaw. Comparison of widefield/deconvolution and confocal microscopy for three-dimensional imaging. In James B. Pawley, editor, *Handbook Of Biological Confocal Microscopy*, pages 453–467. Springer US, 2006.

Q. Shi, Y. Altun, A.J. Smola, and S.V.N. Vishwanathan. Semi-markov models for sequence segmentation. In *EMNLP-CoNLL*, pages 640–648, 2007.

Q. Shi, L. Cheng, L. Wang, and A. J. Smola. Human action segmentation and recognition using discriminative semi-markov models. *International Journal of Computer Vision*, 2011.

Y. Shi and W. C. Karl. Real-time tracking using level sets. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 34–41, 2005.

R. Skloot. *The Immortal Life of Henrietta Lacks*. Crown Books, 2010.

G. Srinivasa, M.C. Fickus, Y. Guo, A.D. Linstedt, and J. Kovačevic. Active mask segmentation of fluorescence microscope images. *IEEE Transactions on Image Processing*, 18(8):1817–1829, August 2009.

M. P. Stewart, J. Helenius, Y. Toyoda, S. P. Ramanathan, D. J. Muller, and A. A. Hyman. Hydrostatic pressure and the actomyosin cortex drive mitotic cell rounding. *Nature*, 469(7329):226–230, 2011.

K. Sung and T. Poggio. Example-based learning for view-based human face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1): 39–51, January 1998.

I. Tsochantaridis, T. Joachims, T. Hofmann, and Y. Altun. Large margin methods for structured and interdependent output variables. *JMLR*, 2005.

M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1–2):61–81, 2005.

A. Veeraraghavan, A. K. Roy-Chowdhury, and R. Chellappa. Matching shape sequences in video with applications in human movement analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(12):1896–1909, 2005.

C. Wahlby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson. Combining intensity, edge and shape information for 2d and 3d segmentation of cell nuclei in tissue sections. *Journal of Microscopy*, 215(1):67–76, July 2004.

R. F. Walker and P. T. Jackway. Statistical geometric features-extensions for cytological texture analysis. *Proceedings of the International Conference on Pattern Recognition*, 2:790–794, 1996.

X. Wang, W. He, D. Metaxas, R. Mathew, and E. White. Cell segmentation and tracking using texture-adaptive snakes. In *IEEE International Symposium on Biomedical Imaging*, pages 101–104, 2007.

Wikipedia. Fluorescence microscope — Wikipedia, the free encyclopedia, 2013. URL http://en.wikipedia.org/w/index.php?title=Fluorescence_microscope&oldid=564586844. [Online; accessed 25-July-2013].

X. Wu and S.K. Shah. A bottom-up and top-down model for cell segmentation using multispectral data. In *IEEE International Symposium on Biomedical Imaging*, pages 592–595, 2010.

L. Xu, E. Oja, and P. Kultanen. A new curve detection method: randomized hough transform (rht). *Pattern recognition letters*, 11(5):331–338, 1990.

F. Yang, M.A. Mackey, F. Ianzini, G. Gallardo, and M. Sonka. Cell segmentation, tracking, and mitosis detection using temporal context. In *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention*, MICCAI'05, pages 302–309, 2005.

Z. Yin, R. Bise, M. Chen, and T. Kanade. Cell segmentation in microscopy imagery using a bag of local bayesian classifiers. In *IEEE International Symposium on Biomedical Imaging*, pages 125–128, 2010.

Y. J. Zheng, T. Furukawa, T. Ogura, K. Tajimi, and N. Inagaki. M phase-specific expression and phosphorylation-dependent ubiquitination of the ClC-2 channel. *The Journal of Biological Chemistry*, 277:32268–32273, 2002.

Q. Zhong, A. G. Busetto, J. P. Fededa, J. M. Buhmann, and D. W. Gerlich. Unsupervised modeling of cell morphology dynamics for time-lapse microscopy. *Nature Methods*, advance online publication, 2012.

C. Zimmer and J.-C. Olivo-Marin. Coupled parametric active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1838–1842, 2005.