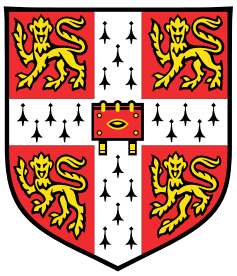
Using 3D image data to improve 2D Brightfield image segmentation

A method of 3D projection



Nicholas Piano

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Master of Philosophy

St Edmund's College November 2015

I would like to dedicate this thesis to my long-suffering supervisor, Y. Y. “Shery" Huang, for her help and advice. Without her guidance, this would not have been possible.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 15,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Nicholas Piano
November 2015

Acknowledgements

And I would like to acknowledge the support and patience of my parents, my colleague Cristina Bertulli, and Xiaohao Cai.

Abstract

Table of contents

List of figures

Chapter 1

Introduction

1.1 Cell segmentation in microbiology

An important part of live cell microbiology is the accurate measurement and tracking of cell morphology during an experiment. Using a microscope, there are many different ways of observing the cells ranging from brightfield microscopy to 3D fluorescence reconstructions. Both 2D and 3D shape data from the cells along with their speed and directionality can provide information on the effectiveness of drugs or other agents in the experiment. The processing of cell data often relies on the quality of Cell Segmentation, or the automatic or manual differentiation of Objects of Interest, such as cells, from the background. Many algorithms and software packages, such as Cellprofiler and ImageJ, are used to segment cells automatically, yielding variable quality.

A key limitation in widely used software that this paper seeks to address is the inability to account for consistent features that cannot be easily located in 3D data. 3D image data, such as from a confocal microscope, contains information about an environment on many focal planes [ref]. Objects can appear blurred or in focus depending the current focal plane. Consequently, features that are useful for segmentation; dark edges, uniform bright interiors, and other features such as fluorescent markers placed within cells are subject to any fluctuations in focus or the movement of objects vertically in the environment. This prevents consistent segmentation of the cell.

The environment used in this study is a microfludics chip built to simulate a human blood vessel. This type of chip, a microchannel framework printed on a PDMS substrate, is widely used in the medical industry to mimic body tissue [ref]. The 3D nature of this setup requires the use of confocal microscopy or similar methods to observe objects in all parts of the environment. In this case, a confocal microscope was used to record data in both a brightfield channel and a fluorescent GFP channel. A number of limitations on the

image quality, discussed in Chapter 2, prevent more powerful direct 3D methods to be used for segmentation. This necessitates a deeper investigation into how the currently available brightfield and GFP information can be used to segment the cells consistently.

1.2 Current work in context

The problem of consistently recognizing objects in a 3D environment using the brightfield is partially addressed in the 2009 study by Selinummi et al. [ref] They attempted to remove the problem of finding objects in 3D by studying the variance of vertical brightness profiles in the brightfield. This had the effect of simplifying 3D data into a 2D plane containing relevant object shapes, from which segmentation could be done more easily. Although this was effective in their case, it has several disadvantages that the current work aims to solve. Notably, it performs poorly in a multicellular environment, necessary for this and many other studies. This is discussed in more detail in the method in Chapter [ref, method].

The current work builds on the idea of studying vertical intensity profiles by applying this concept instead to the GFP. In the current data, the low quality GFP does not show the outlines of cells accurately. This is due to the internal cellular distribution of the GFP, described in Chapter 5. The central regions of the cell (not including the nucleus) are highlighted, so it can be used to locate the general bulk of the cell. The brightfield, on the other hand, shows the edges of the cell more clearly, but only if the cell is in focus (at the correct level). Edges of the cell start to fade along protrusions. In this study, any useful segmentation done must include accurate outlines of long cell protrusions to determine cell motility and behaviour.

The method described here uses the bulk of the cell visible in the GFP to locate the level in the data needed to ensure consistently clear dark edges in the brightfield, yielding accurate cell shape. The GFP and brightfield data represent the same physical space. This correspondence is exploited to allow information about the GFP channel to aid searching for features in the brightfield. The correct edges are located by building a vertical GFP intensity distribution for each pixel and selecting pixels from the brightfield data that match the level of the distribution peak. Several other properties of this GFP "profile" can be used to find different information about the cells. This is described in more detail in Chapter 6. This new method is one of image pre-processing; to prepare images for segmentation by software, in this case, Cellprofiler. It is not a method of segmentation, although as outlined in Chapter 7, it has the potential to become one.

1.3 Thesis outline

Firstly, a brief description of the microfluidics environment, its purpose, and its limitations will be given in Chapter 2. The system used and its properties are the original reason this study into image processing was undertaken. The 3D nature of the system prevents conventional 2D microscopy, and the live cell environment meant that more detailed cell fixing techniques could not be used. The lack of image quality and coarseness of the data led to the need to investigate if any useful information could be extracted reliably from it.

The current method depends heavily on the distribution of GFP within a cell. This, along with an understanding of cell shape and its optical properties and responses to the light used in the microscope is important for determining how useful this method can be in changing conditions. If the type of microscope or environment is changed, this method should be able to operate under a new set of parameters. A description of the places GFP can be found within a cell can inform adaptations of this method and help find conditions where the method could fail or have reduced effectiveness. An example of this is cytosolic versus cytoskeletal GFP, found in the central cytosol and the external cytoskeleton respectively. These two different distributions of GFP can effect this method which relies on peak positions of the profiles. Peaks in different distributions can give misleading answers regarding the positions of parts of the cell. A description of the brightfield reveals similar situations. A key example is the superposition of two objects in the environment. The brightfield light is mixed in the final data, so data representing superimposed objects cannot be divided into specific shapes of both objects. This limitation cannot be resolved with the GFP data. This and other problems will be described in Chapter 3.

To explain the context of this type image pre-processing, a background in image processing and some techniques used by Cellprofiler will be given in Chapter 4. Basic image features can be combined into larger features that can delimit objects and parts of objects. Software like Cellprofiler can use this information to measure shapes and other properties of cells in an image. It can also be used to track cells from one frame to another by looking for similar cell shapes and proximity. The method described here depends on the way a recognizer like Cellprofiler works, although with further work, the information used for pre-processing in this method could allow a more powerful segmentation algorithm to be created. This will be investigated finally in Chapter 7.

Chapter 6

Finally, testing will be used to compare this method to previous experimental work and commonly used methods of image pre-processing. Although this is not yet a fully capable segmentation method, the pre-processing can be compared by segmentation. Data from several experiments was collected and cancer cells within them segmented. Several properties

of the cancer cell shapes, such as the projected 2D area, can be used to compare images on their ability to represent the cells.

Chapter 2

The microfludics environment

2.1 Description of the environment

The microfludics environment used for this study consisted of a microchannel framework printed using soft lithography on a PDMS chip. PDMS, or Polydimethylsiloxane, is a a silicone gel that can be molded and used to create microscopic structures, such as channels [ref]. This microchannel technique is widely used in the biomedical industry to mimic body tissue for complex in vitro [italic] studies. In this case, the channel was set up to model a human blood vessel. The diagram in Figure [ref] below shows part of the channel framework with two PDMS pillars on either side of a gap. On one side of the gap, a liquid medium meant to simulate blood plasma is pumped in where it can remain static, or be used to simulate the flow inside a blood vessel. On the other side of the gap, collagen gel, used to simulate the extra-cellular matrix surrounding a blood vessel, provides an anchor point for the endothelial cells, or cells found in the blood vessel wall, to attach to. These endothelial cells are added to the environment prior to the experiment. The experiment is monitored when cancer cells that are marked with a fluorescent GFP, or Green Fluorescent Protein, are added to the environment in the medium channel, from which they can cross the endothelial cell barrier into the collagen gel, as they might in vivo [italic] during metastasis [ref].

The entire channel is approximately 100 microns high [ref], or about 5 times the height of a typical cancer cell used in the experiments [ref]. The width of the gap containing the endothelial cell barrier is between 100-200 microns [ref]. A typical blood vessel would be tubular, but this setup, while possible [ref], would be difficult to study and to image with a microscope. Opting instead for the simplified setup, the vertical wall of cells can be observed much more easily and still be used to provide information on how the cancer cells cross the barrier.

2.2 Purpose and the need for segmentation

The purpose of this environment is to determine how cancer cells cross the endothelium (barrier) during metastasis.

2.3 Limitations and 3D structure

Chapter 3

Background

3.1 Introduction

3.2 Cell microscopy, optical structure, and GFP distributions

3.3 Image processing and segmentation

3.4 Cellprofiler and segmentation software

3.5 Cell tracking

3.6 The Selinummi brightfield profile method

Chapter 4

Method

4.1 The GFP profile

4.2 Optimum features for cell recognition

4.3 zMod and zBF

4.4 Artificial edges for segmentation: zDiff and zEdge

4.5 Summary

Chapter 5

Results and discussion

5.1 zMod parameters: R, Sigma, and Delta-z

5.2 zDiff and zEdge

5.3 Comparison with common methods

5.4 Errors and limitations

Chapter 6

Conclusion

6.1 Summary

6.2 Observations

6.3 Further work

References