Morphology and Optics of Human Embryos from Light Microscopy

Niels Holm Olsen

21st January 2003
Abstract

Evaluation of fertility of living embryos is of practical importance in the daily routines at fertility clinics and an ongoing biological research topic. Living embryos may be studied through a light microscope. By focusing the microscope at different optical sections, the three-dimensional structure of the embryo may be studied. The purpose of the Fertimorph project is to develop methods for fertility evaluation via measurements of the three-dimensional embryo morphology.

Reconstruction of the 3D structure of living embryos from optical sectional images, must be based on a model of the image formation. The image formation is a result of the optical characteristics of the embryo and the microscope optics used. Human embryos are "large" refractive objects, so the usual model of rectilinear light propagation in a 3D Euclidean geometry does not hold. Since this model of light propagation is central in the linear translational invariant models of light microscopy, these traditional models of image formation and corresponding techniques for 3D reconstruction cannot be used. Instead the morphology has to be inferred directly from the images.

A major challenge in such inferences, in addition to modelling the image formation, is modelling of the morphology and shape variability. Models of shape variability are needed as model structures for analyzing object shape and as tools for inferring and describing the task specific prior knowledge of the imaged objects, often essential for successful inference.

I present investigations of embryo optics, Hoffman Modulation Contrast image formation, and 3D reconstruction of human embryo morphology from HMC images - all in the context of the Fertimorph project. In addition to this, general Lie group models of point set variability inducing well-defined shape variability are discussed, as basic tools in models of shape; and the descriptive power of linear transformation groups for modelling shape variability is analyzed.
Preface

This thesis documents parts of my scientific studies at 3D-Lab, School of Dentistry, University of Copenhagen, enrolled as a PhD student, at the IT University of Copenhagen, and working for Image House A/S, Denmark in an industrial research project called Fertimorph on fertility and morphology analysis of human embryos. The studies were conducted during the period from December 1st 1999 to December 1st 2002, with Mads Nielsen from the IT University of Copenhagen, as my supervisor and Jon Sporring from 3D-Lab, School of Dentistry, University of Copenhagen as co-supervisor.

Working in an application oriented research project has been very rewarding. Fertimorph has been a very inspiring source of computer vision problems. Their origin in a real life application has been my security that their modelling, understanding and solution would be of practical importance. But also the interdisciplinary character of Fertimorph has been a big challenge with very interesting problems spread over many different research areas.

The tour:

from image analysis in computer science, through first the mathematics of probability theory, Bayesian inference and decision theory and next the mathematics and optics of light microscopy image formation; then through the physics of light, light propagation and light-matter interaction, only to move on to embryo specific optics and biological variation and mechanical models of internal embryo interactions; and finally from here back again through the nonlinear differential geometry of shapes to Lie group models of shape ensemble variability, as foundations for few-sample inference techniques of probabilistic models of ensembles showing large natural variations,

is a VERY LONG JOURNEY. It crosses many interdisciplinary boundaries and passes right through central parts of the different disciplines. Spanning so many disciplines is difficult. Consequently I have only been able to study parts of it in detail, and forced to skip many other parts and alternative routes.

So my initial fear that this project would not be a “research” project but only a project of “simple application” of already known techniques, was soon relieved but substituted with the opposite fear, that due to the complexity of the open research problems, it would not be possible to find practical solutions to the original Fertimorph application.

The choice of route has been a compromise, weighting research relevance, Fertimorph applicability and chance of success. Only by taking major shortcuts, have we been able to go part of the way. I truly hope that these compromises have served both research and Fertimorph application.
Acknowledgements

We rarely write I,
since we prefer we.

I may write I,
but should maybe,
rather write we.

Last, but not least,
You and I,
could also be we.

I would here like to express my thanks to the many people who have helped me during my PhD studies: First of all to my two supervisors and my collaborators in the Fertimorph project, for helping me and making it possible to study the many interesting problems of “3D reconstruction, morphology and optics of human embryos from light microscopy”; but also to my very friendly colleagues at 3D-Lab and IT-C, and my collaborators in the Natural Shape Project, for discussing research ideas and joint work on finding solutions. Many other researchers both from the local image analysis, optics and physics community and from around the world deserve thanks, for fruitful discussions at summer schools, workshops and conferences and for answering my e-mails with detailed comments and relevant references to literature and commercial suppliers. Last, but not least, I would like to thank my mother and farther for detailed comments on the manuscript, fruitful discussions on the physics of image formation and patient listening to my ideas and views of image analysis in general.
Thesis Overview

This thesis has been organized as an introductory part followed by 3 parts, corresponding to the 3 subprojects of my studies:

**Introduction** In this part I first present “The Fertimorph Project” (chapter 1). Next, an abstract “Bayesian Model of Fertimorph Inference and Decisions” (chapter 2) is introduced, discussing the chosen methodology by way of examples. This model provides a common frame for discussing and describing in detail the various models and computational methods for automated image analysis I have studied. The complexity and possible structurings of the Bayesian model of Fertimorph is further discussed and analyzed in the following chapter on “Fertimorph Model Complexity” (chapter 3). This allows me to present in more detail various structures of the abstract Bayesian inference model for image analysis in Fertimorph and my understanding of the difficulties motivating their substructuring.

**Towards 3D reconstruction** In this part I present my investigations of 3D reconstruction. First the “Microscope Image Formation” (chapter 4) is presented and potential approximate models shortly discussed. Next “Models and Methods for 3D-Reconstruction” (chapter 5) are analyzed with respect to computational complexity and applicability to different physical modalities, microscope contrast techniques and regimes of refractive objects. Then “A Study of Embryo Optics” (chapter 6) investigating which regime of refractive objects embryos belong to, is presented. Finally, linearity under the approximate model of anomalous diffraction is analyzed in “Reconstructing the Optical Thickness from Hoffman Modulation Contrast Images” (chapter 7), which is an article [37] written by me, Jon Sporring, Mads Nielsen, Christina Hnida and Søren Ziebe.

**Segmenting Embryos in HMC-Images** Unfortunately the model for reconstruction is not suited for modeling embryos with multiple blastomers with different positions along the direction of light propagation. Instead direct “3D Blastomer Detection in HMC-Images of Human Embryos” (chapter 8) using a simple spheres model of blastomer geometry together with a 2D theoretical model of infocus images of individual spheres is presented. This automatic detection algorithm has been implemented, in collaboration with Jon Sporring and Peter Berg Larsen, in a segmentation program with a user interface and an overlap consistency check for investigating and correcting the automatic sphere detections allowing for “Semiautomatic Sphere Model Selection” (chapter 9). Finally an “Evaluation of Semiautomatic Blastomer Detection” (chapter 10) in embryos is presented, based on 21 embryo image sequences recorded and annotated by Søren Ziebe and Christina Hnida.

**Shape Models** This part contains work on shape and shape ensemble modelling, using Lie groups to model well defined point set shape variability. The first chapter, which is an article written by me and Mads Nielsen [35], introduces the “Lie Group Model of Point Set Shape Variability” (chapter 11), for point sets in the plane, and analyzes a particular case of bending of 3 points. In the second chapter, which is an article [46] written by Andrew Swann and me, the expressive power of the Lie group model of shape variability is analyzed by studying “Linear Transformation Groups and Shape Space” (chapter 12) both for point sets in the plane and higher dimensions.
Contents

1 Introduction 1

1 The Fertimorph Project ........................................ 3
  1.1 Background ............................................. 3
  1.2 Fertimorph Goals ....................................... 4
  1.3 This PhD project in Fertimorph ......................... 5
    1.3.1 Input Images ....................................... 6
    1.3.2 Morphology Models ................................ 7

2 Bayesian Model of Fertimorph Inference and Decisions 9
  2.1 Modelling Experiments by Probability Distributions .... 9
    2.1.1 Models of Probability Distributions ............... 10
  2.2 Computational Model of Inference ....................... 10
    2.2.1 Forward and Well-Posed Problems ................. 11
  2.3 Probabilistic Model of Inference ....................... 11
    2.3.1 Maximum A Posteriori Probability ................. 12
    2.3.2 Probabilistic Model of Image Formation .......... 12
    2.3.3 Bayesian Model of Inference ...................... 12
    2.3.4 Bayesian Inference in Fertimorph ................. 13
  2.4 Bayesian Decisions .................................... 15
    2.4.1 Weighting Consequences - the Loss Function .... 15
    2.4.2 Minimal Risk Decisions ........................... 16
    2.4.3 Fertimorph Treatment Decisions ................. 16
    2.4.4 Induced Loss Function for Morphology Model Decisions 17
    2.4.5 Maximum A Posteriori Probability Decisions .... 18
  2.5 Summary .............................................. 18
    2.5.1 Prior Knowledge and Models ...................... 18
    2.5.2 Assumptions ...................................... 18
    2.5.3 Inferred Distributions and MAP Estimates ....... 19
    2.5.4 Optimal Treatment and Morphology Model Decisions 20
  2.6 Conclusions ........................................... 21
    2.6.1 Outline of Typical Automatic Procedure .......... 21
  2.7 Thesis Topics .......................................... 21

3 Fertimorph Model Complexity 23
  3.1 Identifying Modeling Challenges ....................... 23
  3.2 3D-reconstruction - Why and What? ..................... 24
    3.2.1 Irrelevant Imaging Conditions .................... 25
    3.2.2 Translational and Rotational Invariance ......... 25
    3.2.3 Well Defined Object Measurements ............... 26
    3.2.4 Object Models and Shape Models ................. 26
3.3 Isolating Complexity ........................................... 26
  3.3.1 Multilevel Bayesian Model ................................. 27
  3.3.2 Levels in the Image Formation ........................... 28
  3.3.3 Levels in the Object Model ............................... 29
3.4 Stepwise Inference ............................................ 30
  3.4.1 Stepwise MAP Decisions .................................. 30
  3.4.2 Smoothed A Posteriori Probability ....................... 30
3.5 How to Find Suitable Models ................................. 31
  3.5.1 Theoretical .................................................. 31
  3.5.2 Empirical .................................................... 31
  3.5.3 Image Formation ............................................ 32
3.6 Thesis Overview Revisited .................................... 32

II Towards 3D Reconstruction ................................ 33

4 Microscope Image Formation .................................. 35
  4.1 Light Microscopy ............................................. 35
    4.1.1 Zernike's Phase Contrast .............................. 37
    4.1.2 Hoffman Modulation Contrast ......................... 37
  4.2 Physics of Light Microscopy ................................. 38
    4.2.1 Approximate Models of Light and Light Matter Interaction .. 39

5 Models and Methods for 3D Reconstruction ................. 41
  5.1 Direct Reconstruction of a Volumetric Model .............. 41
    5.1.1 Translational Invariance ............................. 41
    5.1.2 Linearity .................................................. 42
    5.1.3 Reconstruction by Deconvolution ..................... 42
  5.2 3D Light Microscopy by Deconvolution .................... 43
    5.2.1 Fluorescent Emission ................................... 43
    5.2.2 Absorption ............................................... 43
    5.2.3 Refraction ............................................... 43
    5.2.4 Diffraction .............................................. 44
    5.2.5 Deflection ............................................... 45
    5.2.6 Anomalous Diffraction ................................. 45
    5.2.7 Limits of Applicability ................................. 46
  5.3 Reconstruction of Human Embryos ......................... 46
    5.3.1 Alternate Models for 3D Reconstruction .............. 47
  5.4 Conclusion .................................................. 48

6 A Study of Embryo Optics .................................... 49
  6.1 Introduction .................................................. 49
  6.2 Microscope Equipment and Image Acquisition ............. 50
    6.2.1 Hoffman Modulation Contrast .......................... 50
    6.2.2 Paraxial Bright Field - Spatially Coherent Illumination .. 51
    6.2.3 Ring shaped Illuminating Aperture .................... 51
    6.2.4 Adjusting for Köhler Illumination .................... 51
    6.2.5 Observed Image Focus Sequences ...................... 52
  6.3 Observations .................................................. 52
    6.3.1 Violations of Rotational Symmetry .................... 52
    6.3.2 The Lens Effect of a Cell ............................. 53
6.3.3 The Beckeline and Non-Central Images of the Illuminating Aperture 54
6.3.4 Towards Qualitative Reconstruction of Optical Thickness 54
6.4 Conclusions 54

7 Reconstructing the Optical Thickness from Hoffman Modulation Contrast Images 57
7.1 Abstract 57
7.2 Introduction 57
7.3 Hoffman Modulation Contrast 58
7.4 HMC Model Approximations 60
7.5 Automated, Fast and Regularized Reconstruction 60
7.6 Evaluation and Conclusion 61

III Segmenting Embryos in HMC-Images 65

8 3D Blastomer Detection in HMC-Images of Human Embryos 67
8.1 3D Detection of Spherical Deviations in Refractive Index 67
8.1.1 Blastomer Detection in Human Embryos 67
8.2 Assumed Model of Image Formation 68
8.2.1 Hoffman Modulation Contrast 68
8.2.2 HMC Image of Homogeneous Sphere 69
8.2.3 Noise and Background Model 71
8.2.4 The Final Image Formation Model - of a Single Blastomer 72
8.3 Single Blastomer Detection and Inference 72
8.3.1 Detection 72
8.3.2 Inference 73
8.3.3 Analyzing the Energy Term 73
8.4 Extending to 3D Image Focus Sequence Model 74
8.4.1 Automatic Cell Detection - What do we Detect? 75
8.5 Multiple Blastomer Detection - Model Match Measure 75
8.5.1 Approximate Linear Image Formation 75
8.5.2 Ignoring Blastomer Image Interactions 75
8.5.3 Conclusion 76

9 Semiautomatic Sphere Model Selection 77
9.1 Design Philosophy 77
9.2 Multiple Blastomers 77
9.2.1 Blastomer Interactions - No Shape Interaction 78
9.2.2 Prior of Blastomer Sizes and Positions 78
9.2.3 Interactive Inspection 79
9.3 Implementation Considerations 79
9.3.1 Automatic Detection of the HMC Analysis Direction 79
9.3.2 Automatic Registration of the Images in the Sequence 80
9.3.3 Automatic 3D Detection of Spherical Blastomers 80
9.3.4 Inferring the Size 81
9.3.5 Filtering Part 81
9.3.6 Local Extrema Detection 82
9.3.7 The No-Overlap Criteria 82
9.4 Conclusion 83
10 Evaluation of Semiautomatic Blastomer Detection

10.1 What and How To Evaluate

10.1.1 Evaluating Validity and Quality

10.1.2 Evaluating User Interaction

10.1.3 3 Evaluation Iterations

10.2 Evaluation I - Automated Cell Detection

10.2.1 Analysis

10.2.2 Input

10.2.3 Manual Evaluation

10.2.4 Output

10.2.5 Manual Evaluation of Detections

10.2.6 Detection Statistics

10.2.7 Types and Causes of False Detections

10.2.8 Statistics of Detection Error Types and Causes

10.2.9 Observations

10.3 Evaluation II - First Improvements

10.4 Evaluation III - Combinatoric Search

10.5 Conclusions

IV Shape Models

11 Lie Group Models of Point Set Shape Variability

11.1 Introduction

11.1.1 Shape

11.1.2 Kendall’s Manifold of Shapes of Point Sets in the Plane

11.1.3 Linear Point Set Models of Local Shape Variation

11.1.4 Nonlinear Models of Point Set Shape Variability

11.2 Variability Modeled by Group of Transformations

11.2.1 Continuous Variability Generated by a Vector Field

11.2.2 Well Defined Shape Variability Generated by Linear Vector Fields

11.2.3 1-parameter Groups Generated by Complex Linear Vector Fields

11.3 Preserved Length Ratio - Nonlinear “Bending”

11.3.1 Centered Bending of 3 Points in the Plane

11.4 Conclusions

12 Linear Transformation Groups and Shape Space

12.1 Introduction

12.1.1 Models of Point Set Shape Variabilities

12.1.2 Non-Linear Shape Ensembles

12.1.3 Linear Transformation Groups and Shape Space

12.1.4 Outline

12.2 Shape Space

12.3 Existence of Symmetries

12.4 Example Symmetries

12.5 Maximality of Symmetry Groups

12.6 Shape Stabilisers

12.7 One-Parameter Groups and Geodesics
Part I

Introduction
Chapter 1

The Fertimorph Project

This chapter presents the background and goals of the Fertimorph project, and how this PhD project is related to Fertimorph.

1.1 Background

Today, couples having difficulty obtaining pregnancy and thus getting children are treated at fertility clinics around the world. Depending on the origin of the difficulty, different treatments are used. Some of these are done “in vitro”, which means that the meeting of the egg (oocyte) and sperm cell(s) (spermatozoos) occurs artificially in a “test tube” (Petri dish) outside the woman. Often more than one egg (usually in the range 1-20) are extracted from the woman (aspirated) and tried fertilized. However to avoid multiple siblings, as few as possible (preferably not more than 2) of the fertilized eggs (embryos) are set back into the woman. The conflict of interests:

- to maximize the chance of getting a child, and
- to minimize the risk of (multiple) siblings

leads to the wish for selecting only the 1 or 2 fittest embryos for implantation. This selection is part of the daily routines at fertility clinics, and is based on manual inspection and scoring of the embryos using an inverted light microscope (see Figure 1.1).

Figure 1.1: Inverted light microscope with needles for ICSI treatment (Intra Cytoplasmic Sperm Injection).

The evaluation of the fertility and fitness for survival of living embryos is thus of practical importance in the daily routines at fertility clinics and still an ongoing biological research topic. As living embryos are transparent, and cannot be stained or fluorescently marked if they are to be used, they are often studied at fertility clinics using the Hoffman Modulation Contrast (HMC) light microscope. Figure 1.2 shows such an HMC image focused at an embryo. However
the process of manual evaluation of the embryo fertility, is prone to human errors and differences in individual subjective judgments. Furthermore, when focusing the microscope at different optical\textsuperscript{1} sections, the clinician can only mentally reconstruct the three-dimensional structure of the embryo, since only one optical section is in focus at a time.

There is thus a wish for quantitative 3D morphology analysis of human embryos. Both as part of clinical routines as well as a firm basis for quantitative statistics in fertility research.

1.2 Fertimorph Goals

The overall goal of the Fertimorph project is to develop a computer vision system integrating:

- Image acquisition
- Patient Database
- 3D Morphology measurements
- Fertility evaluation
- Quality control

for research and clinical use at fertility clinics. To achieve this, Fertimorph is a collaboration between:

\textsuperscript{1}I use the term “optical section” rather than “section” to emphasize that the “sections” of the embryos are not produced by cutting or slicing, but rather by optical focusing - the embryo of course must stay intact.
with collective expertise within the commercial, biological and technological aspects of the projects.

The biological objective is to improve implantation success rates, moving towards single-embryo-transfer, without compromising the fertility rate. This should be achieved by determination of quantitative morphological parameters, which can be used as a basis for objective quality assessment of human embryos in fertility treatment, and selection of the optimal embryo.

The technological objective is to develop computational methods for automated 3D reconstruction, segmentation and morphology analysis of human embryos, based on sequences of optically focused HMC light microscopy images.

This thesis will mainly deal with the technological objective of developing automated procedures resulting in quantitative morphological measurements.

1.3 This PhD project in Fertimorph

With the above description of the biological and technological objectives of the Fertimorph project, this PhD project can shortly be stated as:

3D

- Reconstruction,
- Segmentation, and
- Shape analysis

of Human Embryos from a stack of focused HMC light microscopy images. So, where the long term goal of Fertimorph is to improve the fertility treatment in terms of higher success rates of having only one child, I will concentrate on the analysis of the images with a 3D morphological description as the goal. The last step, consisting of analyzing the morphology and predicting the fertility, I leave to the collaborating biologists at the fertility clinic.

Similarly Fertimorph defines certain limits regarding what the morphology models should be based on. Since the end product of Fertimorph should be easy to integrate with existing equipment and procedures at fertility clinics, the morphology inference must be based on a few simple image measurements obtainable using a non-invasive contrast technique available on standard microscopes at fertility clinics. For example it would not be feasible to require adjustments of the embryo pose.

I guess this would be a good place to introduce in more detail the two interfaces between this PhD project and the Fertimorph project:

- The input images
- The output morphology models

Though they were not completely predefined, they were to a large extend dictated because of considerations to the Fertimorph project.
1.3.1 Input Images

The input images are produced, using the FertiGrab system\(^2\) equipped with the Multi Focus module, as a series of images focused at different optical sections of the same embryo. First the microscope is focused at what we call the "equatorial plane", and a first image of this is acquired. The rest of the images are taken with the microscope focused at planes e.g. 5 micrometers apart starting from e.g. 40 micrometers below (away from the light source) to 40 micrometers above the equatorial plane. Figure 1.3 illustrates such an example image sequence.

Figure 1.3: Stack of HMC images of Human Embryo - The input for 3D reconstruction, segmentation and shape analysis. The “equatorial plane image” is displayed in the topmost row. Next follows from left to right within each row the acquired sequence of images focused 5 micrometers apart. The top left image is focused 40 micrometers below the “equatorial plane”. The bottom right image is focused 35 micrometers above the “equatorial plane”.

\(^2\)From IH-Medical, Image House A/S, Denmark.
1.3.2 Morphology Models

Figure 1.4 illustrates the morphology models that the biologists use. Though the models are

drawn in 2D, the biologists know they represent 3D objects and are interested in 3D embryo
morphology measures like the volumes and surface areas of the different subparts. At day 0,
before fertilization, they are specifically interested in the presence of the first polar body, as well
as the size and shape of the egg-cell (oocyte). On day 1 after fertilization but before the first
cell-divisions, the presence of a second polar body and 2 pro nuclei are good signs. At day 2,
after the fertilization, they are interested in the number of and sizes and shapes of the divided
cells (blastomers), and their geometrical orientation relatively to each other. Also the degree and
characteristics of the fragmentation are of interest, specifically ways of distinguishing between
fragments and blastomers e.g. based on their sizes and shapes. In this context it is also of interest
to detect the presence of a nucleus in the blastomers.

All of these investigations share the overall goal of detecting signs of deviations from a
normal healthy development of the embryo, so the fittest for survival can be selected and used
for implantation.
Chapter 2

Bayesian Model of Fertimorph
Inference and Decisions

The purpose of this chapter is to provide a common theoretical frame for presenting and analysing the computational methods studied in the rest of this thesis. Describing this frame allows me to explicitly identify the model constructs and computational methods which I will study, and thereby restrict the number of approaches considered. It is here presented in the setting of Fertimorph.

It is however not a necessary foundation for reading and understanding the more concrete studies in the other parts of this thesis. Being rather abstract, reflecting my general attitude to image analysis and thus at times philosophical, I recommend readers in search for more concrete and practical studies to skip this chapter and go directly to the relevant part and chapters.

2.1 Modelling Experiments by Probability Distributions

I shall make extensive use of the notion of random variables and their probability distributions, as a convenient way of modelling measurements in connection with an experiment. The motivation for using a probabilistic model is that repeating an experiment rarely leads to the same measurements - often because we are not able to or maybe not even planning to repeat it under the same conditions.

For example, if I consider “observation of an image sequence of an embryo” as the experiment and the image pixel values as associated measurements, I have, amongst a lot of other conditions like internal electric fluctuations of the camera, not specified which embryo or at what time after fertilization it is observed. The next time I “observe an image sequence of an embryo” it may be another embryo - possibly originating from another woman and probably positioned differently in the microscope. This experiment can thus have a lot of different results and consequently lead to a lot of different measurements of image pixel values. The probability distributions describe which measurements are more likely than others. If e.g. the measured image sequence is denoted by $I$, I shall use the common shorthand notation of $P(I)$ for the probability distribution of $I$.

To make the distinction between an experiment and the associated measurements more clear, I can think of any number of measurements represented by different random variables. In fact I shall also use random variables to model “thought measurements” which are possibly never performed directly, like e.g. the embryo morphology $M$. In order to model possible dependencies between measurements (e.g. between $I$ and $M$) associated to the same experiment, I need also the simultaneous distribution $P(I,M)$ of the simultaneous “measurement” $(I, M)$, as well as conditional distributions like e.g. $P(I|M)$ representing the distribution of $I$ under the condition that $M$ is known.
2.1.1 Models of Probability Distributions

It is not trivial to make the notion of a probability distribution mathematically precise. In Kolmogorov’s model of probability theory we think of the image measurement $I$ as a stochastic vector or measurable mapping $I : \Omega \rightarrow I$ from a probability space $(\Omega, P)$ to a measurable space $I = \mathbb{R}^N$ or $I = \{0, \ldots, 255\}^N$, where $N$ is the total number of measured pixel values. The elements of $\Omega$ are all the possible results of the experiment. The probability measure $P$ models the probability $P(A)$ of all possible events represented by measurable subsets $A \subset \Omega$ of experiment results. Together $\Omega$ and $P$ models the experiment. A common way of specifying an experiment event in terms of a measurement event $I_0 \subset I$ is simply as $A = I^{-1}(I_0) = \{ \omega \in \Omega | I(\omega) \in I_0 \} \subset \Omega$. In this setting the above notation $P(I)$ can be seen as sloppy notation of something like $P(\{I = I_0\}) = P(I \in \{I_0\}) = p(I^{-1}(\{I_0\}))$ or more generally $P(I \in I_0) = P(I^{-1}(I_0))$, which is just the distribution of $I$ expressed as the image distribution of $P$ under the mapping $I$. Similarly the sloppy notation $P(I|M)$ for conditional distributions can be interpreted as the probability of the event $A = (I|M^{-1}(M_0))^{-1}(I_0)$ characterized by $I = I_0$ and $M = M_0$ in the model $P(I|M^{-1}(M_0))(A) = P(A)/P(M)$ of the conditional experiment where we already know that $M = M_0$. So in terms of the introduced notation $P(I|M)$ can be defined by $P(I|M) = P(I,M)/P(M)$.

When dealing with discrete distributions on finite sets this theory is straightforward. However, when we want to consider measurements with values in a continuous set we run into the difficulties of measure theory.\textsuperscript{1} It is then convenient to interpret $P(I)$ as the density of $I$, and write

$$P(I \in I_0) = \int_{I_0} P(I) dI. \quad (2.1)$$

Regardless of whether $P(I)$ denotes a probability or a density I shall refer to it as the likelihood of $I$.

Since this thesis is not on the sound mathematical models of probability theory, I shall not be too concerned about the exact mathematical representations but rather use it as a convenient notation for expressing the models used for image analysis.

2.2 Computational Model of Inference

Since the overall purpose of the analysis of the images is to arrive at a 3D model of the “embryo morphology”, we must be able to decide which one of the many possible 3D morphology models, is the right one for a concrete embryo.

If one denotes the set of possible 3D embryo morphology models by $M$, and the set of observable images by $I$, we need either a “deciding map”:

$$D : I \rightarrow M, \quad (2.2)$$

or more generally an energy or “badness of fit” function $E$ returning a measure of “how badly a particular morphology model fits the observed image”:

$$E : I \times M \rightarrow \mathbb{R}. \quad (2.3)$$

From an algorithmic point of view, having a deciding map $D$ or an energy function $E$ is not completely equivalent. Where $D$ represents an algorithm and thus the knowledge of “how” to find the morphology model explaining the observed image, an energy function $E$ only represents

\textsuperscript{1}Not all subsets of a continuous set are events that can be given a probability. All events consisting of only a finite number of measurement results have zero probability with respect to a smooth probability distribution given by a density with respect to some natural uniform measure on the continuous set, like the Lebesgue measure on $\mathbb{R}$.
a computational theory of “what” we want (the minimizing \( M \)). In addition to \( E \) we need a minimization algorithm that knows “how” to find the model with the minimum energy for a given observed image. However if we have a deciding map \( D \) it is easy\(^2\) to define an equivalent energy \( E \), and if \( M \) is finite we can, in theory, use global search to define \( D \) from a given \( E \). By studying energy functions we try to isolate the “what” part of the problem from the “how” part. This corresponds to the attitude:

We should know “what” we want, when trying to solve “how” to do it.

### 2.2.1 Forward and Well-Posed Problems

Studying energy functions can also be seen as a step towards studying the “forward problem”\(^3\):

Predicting the image \( I \) corresponding to a known model \( M \),

and a regularized version of the “inverse problem”:

Inferring the model \( M \) corresponding to an observed image \( I \), by minimizing a regularized energy function.

In many situations of image analysis\([6]\), the “forward” problem is easier to model, because of its nice properties like:

- existence of solutions \( I \) for each \( M \)
- uniqueness of solution (at least in theoretical ideal noise-less models)
- continuity of solution function \((I(M))\)

and, as introduced by Hadamard (see e.g. \([6][p.5]\)), is then said to be well-posed, while the “inverse” problem is ill-posed in the sense of lack of at least one of the above properties of well-posed problems. So, when the inverse problem is ill-posed, we will have problems with existence, continuity or uniqueness when trying to define a decision function \( D : I \rightarrow M \). When this happens, we instead consider the problem of minimizing an energy function, which contains a regularizing term \((E(M))\):

\[
\]  

(2.4)

Here \( E(I|M) \) denotes the energy function without regularization.

Rather than considering the image given, and using the energy function for evaluating “how badly the different models fit the observed image”, we are free to study the energy function in the setting where the model is given and we are evaluating “how badly the model fits different images”. However to really study the forward problem we would rather like to evaluate “how badly the different images fits the model”.

### 2.3 Probabilistic Model of Inference

The above discussion of “the direction” of an energy function as modelling a “forward problem” or the “inverse problem”, is more intuitive and clear, when the energy functions under consideration can be interpreted as Gibbs energies given by negative log-likelihoods of simultaneous or

\(^2\)Simply let \( E(I, M) = 0 \) if \( D(I) = M \) and 1 otherwise.

\(^3\)Which direction is “forward” and which is “backwards” or “the inverse problem” is of course a matter of viewpoint. I have deliberately used “forward” rather than “inverse” here, since this is how it is traditionally viewed. When only one of the “forward” and “inverse” problems is well-posed, it is customary to consider the well-posed problem the forward problem.
conditional distributions:

\[ P(I, M) = \exp(-E(I, M)) \in \mathbb{R}_+ \]  \hspace{1cm} (2.5)

\[ P(I|M) = \exp(-E(I|M)) \in \mathbb{R}_+ \]  \hspace{1cm} (2.6)

\[ P(M|I) = \exp(-E(M|I)) \in \mathbb{R}_+ \]  \hspace{1cm} (2.7)

I have here borrowed the notation of the distributions for the notation of the 3 different energy functions: \( E_{(I,M)}(I, M) = E(I, M) \), \( E_{I|M}(I, M) = E(I|M) \), and \( E_{M|I}(I, M) = E(M|I) \).

### 2.3.1 Maximum A Posteriori Probability

From the interpretation of \( P(M|I) \) as the probability distribution of \( M \) when the image is known to be \( I \), it is clear that it is really the energy function \( E(M|I) \) which we would like to minimize in order to infer the most probable model \( M \) given the observed image \( I \). Since \( P(M|I) \) is known as the a posteriori distribution, this morphology model:

\[ M_{MAP} = \arg \max_{M \in M} P(M|I), \]  \hspace{1cm} (2.8)

is also known as the “Maximum A Posteriori” (MAP) estimate.

### 2.3.2 Probabilistic Model of Image Formation

However, the a posteriori distribution \( P(M|I) \) is extremely task specific. The probabilities \( P(M|I) \) reflect what models we are taking images of in the experiment. Because of this, and because we can often expect the problem of finding \( M \) from \( I \) to be ill-posed, it is difficult to model \( P(M|I) \).

It is therefore often easier and of more general interest to model the forward problem of image formation:

\[ P(I|M) = \frac{P(I, M)}{P(M)}. \]  \hspace{1cm} (2.9)

This is a family of conditional distributions \( P(I|M) \) indexed by the model \( M \), each describing how likely it is to see different images \( I \) of the given model \( M \). In this sense \( P(I|M) \) is a probabilistic model of the image formation, which is independent of the task as characterized by the probability distribution \( P(M) \) of the task specific objects under study.

Still, having a model \( P(I|M) \) of the image formation is not enough. If the inverse problem is ill-posed we need to deal with this by some sort of regularization. Furthermore we still would prefer to infer \( M \) from \( I \) based on the task specific model of \( P(M|I) \) when we know the task. The Bayesian model of inference describes how to model and combine this task specific knowledge with the model of the image formation.

### 2.3.3 Bayesian Model of Inference

In the Bayesian model of inference, the image formation is extended with a prior probabilistic model of the embryo morphology models,

\[ P(M), \]  \hspace{1cm} (2.10)

which expresses the important task-specific prior knowledge of what objects we are studying in the experiment, and which is often essential for successful inference. Though exact prior models often are not known\(^4\) more general models may be available and hopefully sufficiently accurate.

\(^4\)The interest in inferring particular instances of models from observed images, is often motivated by a study of the prior model \( P(M) \) itself. In fact this is the case of the scientific part of the biological part of the Fertimorph project. Here the goal is to learn \( P(M) \) or at least conditional versions \( P(M|F) \) for different classes of embryos with varying fertility measures \( F \).
and specific to work as regularizers of the otherwise ill-posed problem of inferring $M$ correctly from $I$.

This prior knowledge is incorporated using Bayes rule for the a posteriori probability of a model $M$, given the observed image $I$:

$$P(M|I) = \frac{P(I|M)P(M)}{P(I)}.$$  \hfill (2.11)

Here $P(I) = \int_M P(I|M)P(M)dM$, which is often called the partition function, is the normalizing probability of the event of observing the image $I$ for (any) unknown embryo morphology $M \in \mathbf{M}$.

The Bayesian model of inference describes how to recombine the task independent model $P(I|M)$ of the (well-posed) forward problem of image formation and the task specific prior $P(M)$ of studied objects to a task specific regularized model $P(M|I)$ of the (otherwise ill-posed) inverse problem.

### 2.3.4 Bayesian Inference in Fertimorph

Keeping the long term goal of the Fertimorph project in mind - predicting the fertility of embryos, and selecting the optimal one(s) - we could extend the Bayesian model all the way to a model $F$ of an “embryo fitness” or “embryo implantation success” measure. Although I consider this extension outside the central focus of this PhD project, I will expand on this below. It allows me to study in detail an example of how two Bayesian inference models interface, and specifically how the results of the morphology analysis can be used for the following fertility analysis and treatment decisions.

It is not obvious just how to quantify the “embryo fitness” $F$. To keep things as simple as possible, I shall use a minimalistic binary representation of $F \in \mathbf{F} = \{\text{child}, \text{no child}\}$, and think of $F$ as essentially telling whether this embryo can become a child or not\(^5\). The over-all Bayesian model can then be written:

$$P(F|I) = P(I|F)\frac{P(F)}{P(I)}.$$  \hfill (2.12)

Now, splitting up on intermediate morphology models as possible “causes” or “explanations” using:

$$P(F, I) = \int_M P(F, I, M)dM,$$ \hfill (2.13)

we get:

$$P(F|I) = \frac{P(F, I)}{P(I)} = \frac{\int_M P(F, I, M)dM}{P(I)} = \int_M \frac{P(F, I, M)P(I, M)}{P(I)}dM = \int_M \frac{P(F, I)P(I, M)}{P(I)}dM = \int_M P(F|I, M)P(M|I)dM.$$  \hfill (2.18)

\(^5\)An obvious (and more realistic) extension of this would be to consider probability distributions on \{child, no child\} as the representation of $F$. As we shall see, this is in fact what is communicated in form of the complete a posteriori distribution of embryo fitness when using the full Bayesian model of inference.
The Fertimorph Inference Assumption

We now assume\(^6\) that knowing the embryo morphology model \(M\), is all there is to know from \(I\) about \(F\) in the sense\(^7\)
\[
P(F|M,I) = P(F|M). \tag{2.19}
\]

From this assumption we get:
\[
P(F|I) = \int_M P(F|M)P(M|I)dM. \tag{2.20}
\]

Full Bayesian Model of Inference

I interpret this formula as stating that once we (based on the observed image \(I\)) have inferred the a posteriori probability \(P(M|I)\) of the embryo morphology model \(M\) we should use it together with a model \(P(F|M)\) relating embryo morphology to embryo fertility in order to calculate the a posteriori probability of embryo fertility given the observed image \(I\).

In such a fully Bayesian model of inference the posterior probability,
\[
P(M|I) = \frac{P(I|M)P(M)}{P(I)}, \tag{2.21}
\]
is the end result of the Bayesian morphology analysis of the images.

To complete the inference of embryo fertility we must perform multiple Bayesian inferences of \(F\) from \(M\):
\[
P(F|M) = \frac{P(M|F)P(F)}{P(M)}, \tag{2.22}
\]
and combine the results to the final posterior distribution of embryo fertility
\[
P(F|I) = \int_M P(F|M)P(M|I)dM. \tag{2.23}
\]

Here the task specific priors \(P(M)\) and \(P(I)\) are induced from the task specific prior \(P(F)\) using the two models \(P(M|F)\) and \(P(I|M)\):
\[
P(M) = \int_F P(M|F)P(F)dF \tag{2.24}
\]
\[
P(I) = \int_M P(I|M)P(M)dM. \tag{2.25}
\]

Maximum A Posteriori Inference

In the deterministic and invertible situation the morphology model \(M\) is uniquely identified by the image. The inferred a posteriori distribution \(P(M|I)\) will then be degenerate with all the probability mass centered on that model, and the integral over \(M\) reduces to \(P(F|M)\). So in this situation we can consider \(M\) as a known measurement result, and we arrive at the

\(^6\)This is an implicit assumption of Fertimorph, that we can infer the embryo fertility \(F\) from the images \(I\) by first extracting the embryo morphology \(M\) from \(I\) and then infer \(F\) from \(M\). The validity and accuracy of this assumption of course depends on what we understand by the embryo morphology. If you don’t like it as an assumption, think of it as an approximation.

\(^7\)This is equivalent to assuming that \(M\) is a “sufficient statistic” of \(I\) for inferring \(F\) in the sense \(P(I|M,F) = P(I|M)\), which can be read as: once we know \(M\), additional knowledge of \(F\) does not alter the distribution of \(I\).
standard Bayesian formula for inferring the posterior distribution of embryo fertility $F$ after having calculated $M$ from the image.

In practice we try to mimic the deterministic case by first finding the maximum a posteriori (MAP) estimate of the embryo morphology,

$$M_{\text{MAP}} = \arg \max_{M \in M} P(M|I), \quad (2.26)$$

and then from this inferring the posterior distribution of embryo fertility,

$$P(F|M = M_{\text{MAP}}) = \frac{P(M = M_{\text{MAP}}|F)P(F)}{P(M = M_{\text{MAP})}}, \quad (2.27)$$

and possibly again reducing to the MAP estimate of $F|M = M_{\text{MAP}}$

$$F_{|M = M_{\text{MAP}}}_{\text{MAP}} = \arg \max_{F \in F} P(F|M = M_{\text{MAP}}). \quad (2.28)$$

### 2.4 Bayesian Decisions

Above I have described how a Bayesian inference of the full posterior distribution of embryo morphology models in theory can be used as a basis for further inference of the posterior distribution of embryo fertility. However the inferred posterior probability distribution $P(M|I)$ or $P(F|I)$ "only" gives the probabilities of the different models being the "cause" of the observed image. We are not told "which model is the right or best one to use". To answer this question, we need more information about our goal. In fact, different goals may have different "optimal choices". Which is "the right or best decision" depends on the consequences of taking it, and how the different consequences are weighted against each other.

#### 2.4.1 Weighting Consequences - the Loss Function

In classical decision theory (see e.g. [31][p.8-18]) this weighting of the consequences is described by a loss function $L(M^d, M^t)$ or $L(F^d, F^t)$, giving the loss value of (the consequences of) deciding on the model $M^d$ or fertility $F^d$ when the true model or fertility is in fact $M^t$ or $F^t$. Here the decision has been described as a decision of inferred embryo morphology or fertility, but also more general decisions from an abstract decision space $D$ may be analyzed by a general loss function $L: D \times I \rightarrow \mathbb{R}$, where $I$ may be substituted by $M$ or $F$ depending on what grounds the decision is made and the (expected) loss evaluated.

**Fertimorph Treatment Consequences**

Again keeping the over all Fertimorph goal in mind, it may be more relevant to consider the treatment decision of how many and which embryos to use. Assuming we have $N$ embryos from a couple and an image sequence $I_1, \ldots, I_N$ of each, the over-all Bayesian model could be written:

$$P(F_1, \ldots, F_N|I_1, \ldots, I_N) = P(I_1, \ldots, I_N|F_1, \ldots, F_N) \frac{P(F_1, \ldots, F_N)}{P(I_1, \ldots, I_N)}. \quad (2.29)$$

In that case (see e.g. [39][p.40]) the treatment loss function $L_t$ would be defined on the product space of a treatment decision space $D^N_1$ (like e.g. $D_t = \{\text{use, don’t use}\}$) and the model space of (true) embryo fertility $F^N$:

$$L_t: D^N_1 \times F^N \rightarrow \mathbb{R}. \quad (2.30)$$
Now the loss function could be designed to express the weighting of the (expected) consequence of having 0, 1, 2, or more children.

Please note that, in the previous section on Bayesian inference in Fertimorph, we only considered one embryo, in order to keep the intuition clear and formulas simple. When discussing treatment decisions, we must consider them simultaneously, since the decisions of using the different embryos are highly dependent on each other. To keep notation simple and general, I will use short notation like \( \mathbf{F} = (F_1, \ldots, F_N) \) and \( \mathbf{F} = F^N \) from now on.

### 2.4.2 Minimal Risk Decisions

The Bayesian paradigm of having a prior distribution of models, allows one to define optimality of a decision function \( D : \mathbf{T} \rightarrow \mathbf{D} \), as one that decides in a way that minimizes the risk \( r(D) \) as measured by the expected loss ([39][p.50]) under the Bayesian prior model:

\[
r(D) = \int_{\mathbf{F}} P(\mathbf{F}) \int_{\mathbf{T}} P(\mathbf{T} | \mathbf{F}) L(D(\mathbf{T}), \mathbf{F}) d\mathbf{T} d\mathbf{F}. \tag{2.31}
\]

Here \( D(\mathbf{T}) \) denotes the decision taken based on the observed images. In order to minimize the risk one can “simply” ([39][Theorem 2.5 p.50]) decide on a decision \( D(\mathbf{T}) = d \) which minimizes the posterior expected loss \( \rho(d | \mathbf{T}) \):

\[
\rho(d | \mathbf{T}) = \int_{\mathbf{F}} P(\mathbf{F} | \mathbf{T}) L(d, \mathbf{F}) d\mathbf{F}. \tag{2.32}
\]

So the Bayes optimal decision in minimal risk sense is given by:

\[
D(\mathbf{T}) = \arg \min_d \rho(d | \mathbf{T}). \tag{2.33}
\]

### 2.4.3 Fertimorph Treatment Decisions

From the inferred posterior distribution \( P(\mathbf{F} | \mathbf{T}) = \int_{\mathbf{M}} P(\mathbf{F} | \mathbf{M}) P(\mathbf{M} | \mathbf{T}) d\mathbf{M} \) of embryo fertility calculated from the posterior embryo morphology distribution inferred from the images, the expected treatment loss \( \rho_t(d_t | \mathbf{T}) \) of the treatment decision \( d_t \) can be calculated:

\[
\rho_t(d_t | \mathbf{T}) = \int_{\mathbf{F}} \int_{\mathbf{M}} P(\mathbf{F} | \mathbf{T}) P(\mathbf{M} | \mathbf{T}) d\mathbf{M} L_t(d_t, \mathbf{F}) d\mathbf{F}. \tag{2.34}
\]

When the inferred embryo morphology posterior distribution is degenerate with all the probability mass located at \( \mathbf{M} = \arg \max_{\mathbf{M} \in \mathbf{M}} P(\mathbf{M} | \mathbf{T}) \), we get the simpler formula:

\[
\rho_t(d_t | \mathbf{M}) = \int_{\mathbf{F}} P(\mathbf{F} | \mathbf{M}) L_t(d_t, \mathbf{F}) d\mathbf{F}. \tag{2.35}
\]

We see that the posterior expected loss \( \rho_t(d_t | \mathbf{T}) \) of a treatment decision given the observed images is a weighted average of these posterior expected losses \( \rho_t(d_t, \mathbf{M}) \) of the treatment decision given the morphology models:

\[
\rho_t(d_t | \mathbf{T}) = \int_{\mathbf{M}} P(\mathbf{M} | \mathbf{T}) \int_{\mathbf{F}} P(\mathbf{F} | \mathbf{T}) L_t(d_t, \mathbf{F}) d\mathbf{F} d\mathbf{M} \tag{2.36}
\]

\[
= \int_{\mathbf{M}} P(\mathbf{M} | \mathbf{T}) \rho_t(d_t | \mathbf{M}) d\mathbf{M}. \tag{2.37}
\]

We thus see that optimal treatment decisions, in the minimum risk sense, should be made by choosing the treatment decision \( d_t \) which minimizes a weighted average of posterior expected losses given the embryo morphology models. The weights are given by the posterior distribution of morphology models inferred from the observed images. The expected loss for a given set of morphology models is again given as a weighted average of the loss function of the treatment decision given the fertility of the embryos, where the weighting is given by the posterior distribution of embryo fertility inferred from the morphology models.
2.4.4 Induced Loss Function for Morphology Model Decisions

Above I have reviewed how the posterior distribution of embryo fertility together with a loss function weighting the consequences of different treatment decisions, defines the posterior expected loss; and finally how optimal treatment strategies in the minimal risk sense decides by minimizing this posterior expected loss. However, computing and communicating a full posterior distribution of embryo morphology models is likely to be intractable. Instead, we would like to mimic the deterministic and invertible situation where there is only one inferred morphology model. Hopefully, the inferred posterior distributions are close enough to being degenerate, so it makes sense to approximate it by a degenerate distribution representing a single morphology model \( \mathcal{M} \). But if this is not the case, we may need to take into account the goal of Fertilimorph as expressed by the treatment decision loss function, when deciding on a single morphology model to base the treatment decision on rather than simply deciding on the MAP estimate.

Assumption of Optimal Treatment Decisions

In this case, the treatment decision will only be based on the inferred morphology model \( \mathcal{M} \) - not the full posterior distribution. It is thus natural to assume that the treatment decision \( d_t \) will be chosen as to minimize the posterior expected loss\(^8\) given the (decided) morphology models \( \mathcal{M}^d \):

\[
D_t(\mathcal{M}^d) = \arg \min_{d_t} \rho_t(d_t | \mathcal{M}^d). \tag{2.38}
\]

Induced Loss Function

This treatment decision strategy will induce a posterior expected loss function \( \rho_M \) on the set of morphology model decisions \( D_M(\mathcal{T}) = \mathcal{M}^d \in \mathcal{M} \) given the observed images \( \mathcal{T} = (I_1, \ldots, I_N) \), which is given by:

\[
\rho_M(\mathcal{M}^d | \mathcal{T}) = \int_{\mathcal{M}} P(\mathcal{M}' | \mathcal{T}) \rho_t(D_t(\mathcal{M}^d | \mathcal{M}')) d\mathcal{M}'. \tag{2.39}
\]

We see that this corresponds to an induced loss function for morphology model decisions,

\[
L_M(\mathcal{M}^d | \mathcal{M}') = \rho_t(D_t(\mathcal{M}^d | \mathcal{M}')), \tag{2.40}
\]

which is based on knowledge of the following treatment decision strategy \( D_t : \mathcal{M} \rightarrow \mathcal{D}_t \) and the posterior expected loss function \( \rho(d_t, \mathcal{M}) \) of a treatment decision given the exact morphology models.

Assumption of Surjective Treatment Decision Strategy

Assuming that \( D_t \) is surjective, we can force any treatment decision by a suitable morphology model decision, and are thus able to achieve the same minimal risk when performing optimal morphology model decisions as when making optimal treatment decisions based on the full posterior distribution of morphology models.

---

\(^8\)At this moment, I cannot rule out that the treatment decision can be made (slightly) more cleverly, by taking into account that it will be based on a forced degenerate model approximation of the a posteriori distribution of morphology models. The idea should be that the treatment decision strategy and the morphology model approximation decision should be coordinated. For the point I am trying to make here, this is not essential. The point being that once the two decision strategies have been coordinated the actual treatment decision will be made at the point of deciding on the morphology model approximation, since the following treatment decision will be made solely on this morphology model.
2.4.5 Maximum A Posteriori Probability Decisions

When the task specific information on how the decision consequences are weighted against each other is not available we will have to do with a more general principle for deciding on the morphology model. In this case it is common to simply decide on the morphology model $M_{\text{MAP}}$ with “Maximum A Posteriori” (MAP) probability:

$$M_{\text{MAP}} = \arg \max_{M \in \mathcal{M}} P(M|I).$$

(2.41)

According to [31][p.16-17], this corresponds to the so-called (0-1) loss function, which attributes the loss 1 to all model decisions except the true model, which is given a loss of 0. When $P(M|I)$ is a density, the loss function is modelled by a negative Dirac delta distribution located at the true model, which integrates to -1 on any set containing the true model and to 0 on all other sets.

2.5 Summary

In order to realize the above extreme degree of dedication of embryo image analysis and morphology inference to embryo fertility assessment and fertility treatment, we need to know both the task specific priors and the goal specific (expected) consequences as well as the imaging model.

2.5.1 Prior Knowledge and Models

Embryo Fertility Specific

<table>
<thead>
<tr>
<th>$\mathcal{M}$</th>
<th>Set of embryo morphology models.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{F}$</td>
<td>Set of embryo fitness measures.</td>
</tr>
<tr>
<td>$P(M</td>
<td>F)$</td>
</tr>
<tr>
<td>$P(F)$</td>
<td>Prior distribution of embryo fertility.</td>
</tr>
</tbody>
</table>

Fertility Treatment Specific

<table>
<thead>
<tr>
<th>$\mathcal{D}_t$</th>
<th>Set of treatment decisions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_t(d_t</td>
<td>F)$</td>
</tr>
</tbody>
</table>

Image Formation Specific

<table>
<thead>
<tr>
<th>$\mathcal{I}$</th>
<th>Set of images from a couple.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{M}$</td>
<td>Set of embryo morphology models.</td>
</tr>
<tr>
<td>$P(I</td>
<td>M)$</td>
</tr>
</tbody>
</table>

2.5.2 Assumptions

(A1) $P(F|M, I) = P(F|M)$ The Fertimorph inference assumption: $M$ is a sufficient statistic of $I$ for inferring $F$.

(A2) $D_t(M^d) = \arg\min_{d_t} \rho_t(d_t|M^d)$ The basis for induced morphology decision (expected) loss function: The morphology based treatment decision follows a minimal risk strategy.

(A3) $D_t(M) = \mathcal{D}_t$ The basis for existence of optimal morphology decisions: We can reach any treatment decision by a suitable morphology model decision $M^d$. 

18
2.5.3 Inferred Distributions and MAP Estimates

From the fertility specific prior knowledge we may infer induced morphology specific prior knowledge, needed together with the model of the image formation, in order to infer the embryo morphology and fertility:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(M)$</td>
<td>Induced prior distribution of embryo morphology.</td>
</tr>
<tr>
<td>$P(F</td>
<td>M)$</td>
</tr>
<tr>
<td>$P(I)$</td>
<td>Induced prior distribution of embryo images.</td>
</tr>
<tr>
<td>$P(M</td>
<td>I)$</td>
</tr>
<tr>
<td>$P(F</td>
<td>I)$</td>
</tr>
<tr>
<td>$F_{\text{MAP}}$</td>
<td>Maximum a posteriori estimate of embryo fertility from observed image under assumption A1.</td>
</tr>
<tr>
<td>$M_{\text{MAP}}$</td>
<td>Maximum a posteriori estimate of embryo morphology from observed image.</td>
</tr>
<tr>
<td>$(F</td>
<td><em>{M=\text{MAP}})</em>{\text{MAP}}$</td>
</tr>
</tbody>
</table>

This is illustrated in Figure 2.1.
2.5.4 Optimal Treatment and Morphology Model Decisions

From the fertility treatment specific goals we may further infer induced morphology specific expected treatment losses, optimal morphology specific treatment decisions and prior knowledge of the expected losses of morphology model decisions, needed together with the model of the image formation in order to decide optimally on the treatment and the morphology model in a minimal treatment risk sense:

| $\rho(d_t|\mathcal{M})$ | The morphology specific posterior expected losses of treatment decision $d_t \in \mathcal{D}$. |
|-------------------------|----------------------------------------------------------------------------------|
| $D_t(\mathcal{M})$      | The optimal treatment decision strategy based on known morphology achieving minimal posterior expected loss $\rho(d_t|\mathcal{M})$. |
| $L_M(\mathcal{M}^d|\mathcal{M}^t)$ | The induced (expected) loss function for morphology model decisions $\mathcal{M}^d \in \mathcal{M}$, assuming the following treatment decisions are taken (optimally) according to $D_t(\mathcal{M})$. |
| $D_t(\mathcal{I})$      | The optimal treatment decision strategy based on observed images achieving minimal posterior expected loss $\rho(d_t|\mathcal{M})$. |
| $D_M(\mathcal{I})$      | The optimal morphology decision strategy based on observed image achieving minimal posterior expected loss $\rho(d_t|\mathcal{M})$. |

This is illustrated in Figure 2.2.

Figure 2.2: Fertilimorph minimal risk decision overview
2.6 Conclusions

Since I do not have most of the prior knowledge, and even if I had it would be very complicated to follow the analysis through to optimal model and treatment decisions in a minimal risk sense, I will have to make simplifying approximations.

Instead I shall consider maximum a posteriori estimation.

2.6.1 Outline of Typical Automatic Procedure

With the above Bayesian framework, a typical procedure for inference of an embryo morphology model $M$, will thus be based on:

1. A prior model of expected embryo morphology models
   \[ P(M). \] (2.42)

2. A model of the image formation:
   \[ P(I|M). \] (2.43)

3. An optimization algorithm for finding the MAP estimate
   \[ M_{\text{MAP}} = \arg \max_{M \in M} P(M|I). \] (2.44)

This is all very general up to now. I have not specified what sets to use for $M$ and $I$, nor how to represent and approximate $P(I|M)$ and $P(M)$.

2.7 Thesis Topics

In this thesis I will describe and analyze:

- possible structures and
- specific choices

of approximations to the models

- $M$,
- $P(M)$,
- and $P(I|M)$,

and

- the challenges they put on the search for $M_{\text{MAP}} = \arg \max_{M \in M} P(M|I)$. 

21
Chapter 3

Fertimorph Model Complexity

In this chapter I analyze in general terms the origins of complexity of the abstract Bayesian models of inference presented in chapter 2, when applied to image analysis and morphology inference in the Fertimorph project. In particular the origin of the need for and difficulty of 3D reconstruction is analyzed. This analysis suggests a number of general ways of isolating and controlling model complexity by structuring the models and methods needed for inferring embryo morphology from transmitted light microscopy images.

3.1 Identifying Modeling Challenges

If either of the three ingredients, the prior distribution of embryo morphology $P(M)$, the model of image formation $P(I|M)$ and the optimization algorithm $\arg \max_{M \in \mathbb{M}} P(M|I)$, in the Bayesian model of Fertimorph MAP inference (see Figure 2.1 page 19 in chapter 2) are inadequate, the procedure for inferring the embryo morphology $M$ from images $I$ as a whole will fail. Specifically it is essential that the inferred a posteriori probabilities $P(M|I)$ of the model are in fact maximal for the “true embryo models” of an image. Also the optimization algorithm must be able to find them. These three ingredients often pose the following qualitative challenges:

1. Modelling the prior $P(M)$ is complicated by the large amount of natural variations of embryo morphology, yet all adhering to the physical constraints between different parts. To my knowledge we do not have such a quantitative model. We will have to do with an approximation and consequently possible discrepancies from reality and optimality.

2. Modelling the image formation $P(I|M)$ is complicated by at least the following effects:

- The image formation from partial morphology models is ill-posed: When using an approximate partial model $M_{\text{approx}}$ of embryo morphology, the unmodelled morphology variability induce far from degenerate and complicated conditional distributions $P(I|M_{\text{approx}})$ of images. This is a consequence of the complexity of $M$ and $P(M)$ and the large information content of $M$ brought to the the images $I$ by the image formation.

The set of images is large, and due to the natural variations of embryo morphology, the set of embryo images is a large and complicated subset thereof - embryos don’t all look the same and the restrictions on their morphological variations induce complicated dependencies between different parts of the images.

The standard similarity measure or metrics, like the inner product on the set of images, $I = L_2(\mathbb{R}^2, \mathbb{R})$, modelled by square integrable real functions on $\mathbb{R}^2$:

$$< I_1, I_2 > = \int I_1(x, y)I_2(x, y)dx dy, \quad (3.1)$$

23
and the induced norm:

\[ ||I||^2 = \langle I, I \rangle, \]  

are not well suited to express these large allowable variabilities.

- Embryo morphology is mixed with other imaging conditions by the image formation: The images \( I \) do not only depend on the embryo morphology \( M \). The observed images also depend strongly on a number of imaging conditions, which can vary independently of the embryo morphology. The set of imaging conditions, which I shall denote by \( C \), includes the position and pose of the embryo.

- The mixing of embryo morphology with position and pose is complicated: Even though the images can be considered functions defined on a 3D space, which in a natural way can be related to the 3D space of the imaged embryo, the intensity values at a specific point in an image, and the physical properties at the corresponding real 3D point of the embryo, are not directly related. The intensity at an image point is not simply a function of the physical property at the corresponding real spatial point.

- The inverse problem of the image formation is ill-posed: There are aspects of the morphology \( M \) which are not imaged by the image formation and consequently not visible in the images \( I \). Which aspects of the morphology are invisible, depends on the imaging conditions.

Again we will have to do with an approximation.

3. When there are non-linearities in the image formation, direct calculation of the solution of the inverse problem will often not be possible. In those cases a search based optimization is complicated by the large number of embryo morphology models making global search for the MAP model in the full space of morphology models impossible; and local search methods are likely to get stuck in one of the multiple local extrema of typical approximate posterior probability distributions.

This is illustrated in Figure 3.1.

### 3.2 3D-reconstruction - Why and What?

Based on the above understanding of the modelling challenges, we are able to answer the questions:

- Why do we need a 3D-reconstruction?

- and what exactly is it?

- Why can’t the images be used directly?

The short answer to these questions could be, that the image measurements are not directly related to the structure we want to study. They are only images of the structure. We need to calculate the structure that gives rise to the measurements we make. This calculation is called reconstruction. Since the structure we are interested in is a 3D spatial structure we talk of 3D-reconstruction.

24
3.2.1 Irrelevant Imaging Conditions

To better understand the need for 3D-reconstruction, we can for a moment analyze which properties an expedient 3D-reconstruction has. Of course the data that the following morphology and fertility analysis is based on should reflect the object and not the arbitrary imaging conditions, like e.g. the position and pose of the object. Since the object studied is a volume object in 3D-space the “reconstructed object model”, should have a similar 3D volumetric structure.

3.2.2 Translational and Rotational Invariance

This can be made more concrete and precise in terms of translations and rotations of the object. Suppose the object were to be translated slightly. This obviously wouldn’t change the object and so we wouldn’t like the 3D-reconstruction to change except maybe corresponding to a similar translation. This requirement is in fact to a fairly good approximation met by the originally observed images. However, if we were to rotate the object a bit, the observed images would not simply be rotated versions! The observed image intensities do not constitute a volumetric dataset, where each pixel intensity corresponds to the measurement of some physical property of the corresponding volume of the object. 3D-reconstruction is the process of inferring such a volumetric dataset from the observed images. We need this reconstruction to make sure that the data we analyze has to do with the object, and not the arbitrary choice of pose and image

---

1This is mainly, but not solely, due to the geometry and physics of the focusing image formation by the microscope optics from the light emerging from the studied embryo. I believe that the global morphology of the studied embryo is also needed to model the dependencies on the physical properties of the rest of the embryo. This is a very subtle statement. If the intensity at an image point depends on the physical properties of all points of the embryo, of course you need to know the full embryo. What I am trying to say is, that you also need to know the full embryo in order to figure out what these dependencies are. That is the dependencies depend on the embryo. In terms of a linear translational invariant model this means that the weights depend on the embryo. So I guess I am saying a linear translational invariant model doesn’t hold.
3.2.3 Well Defined Object Measurements

One may take the above considerations of desired invariance properties a step further, and apply them to the overall task of inferring embryo morphology or performing other measurements on a single embryo from a microscopy focus image sequence of the embryo at a single position and pose. The question is simply, if it is at all possible:

Does there exist a non-trivial\(^2\) function taking an image sequence as input and giving a true object measurement as output, in the sense, that the result of the function is the same for any possible pose and position of the same embryo?

Of course this is not the case in a strict sense, since obviously the embryo needs to be in the field of view of the microscope. We should properly also allow for some noise on the measurement. But even if we loose up on the requirement of invariance with respect to position and pose, in order to tackle these problems, it is not obvious whether the image formation brings enough information about the object, that we can actually achieve well defined object measurements by analyzing the images for a particular position and pose. Since we want to use these measurements for choosing between individual embryos, it is important that we are in fact able to measure individual\(^3\) object properties. It would not be satisfying if the decision depended on the arbitrary position and pose of the imaged embryos.

3.2.4 Object Models and Shape Models

Often when modelling an object, not only the object is represented but also the position and pose relative to an arbitrary coordinate system is represented either implicitly or explicitly. The above discussion of well defined object measurements is really about isolating an objects intrinsic properties like the spatial variation of a physical property from the arbitrary imaging conditions in general and in particular its arbitrary position and pose.

What we really want is an “object model” and not a combined “object and imaging conditions model”. The abstract mathematical way of constructing an “object model”, from a combined “object-, position- and pose- model”, is to consider equivalence classes with respect to translations and rotations.

When the interest is restricted to geometrical object properties, and one further wants to ignore the size by grouping objects that are identical up to a scaling, one talks of the shape of the object. Since many of the interesting morphological parameters of the embryos are actually shape parameters, I have worked in parallel on general models of shape and shape variability. This mainly in the setting where the geometry is represented by identifiable point sets, but also on extensions to planar curves with the hope that they can be applied to modelling of embryo morphology, after the 3D reconstruction and segmentation.

3.3 Isolating Complexity

Up to now, I have not specified what sets to use for \(M\) and \(I\), nor how to represent and approximate \(P(I|M)\) and \(P(M)\).

Let me illustrate the general possibility of trading increased embryo morphology model complexity (increasing size of \(M\), and \(P(M)\)) for decreased complexity of the image formation model

\(^2\)Constant functions are not really interesting as object measurements!
\(^3\)If we were only interested in statistics of embryos, we could use the theory of stereology to estimate average morphological parameters.
as described by the probability distributions \(P(I|M)\) of images conditional on the different morphology models. Suppose I want to use a very simple model of embryo morphology, only considering the number of blastomeres, and disregarding everything else as not being considered part of the embryo morphology. In this case, I could use the set of natural numbers \(N\) as my set of embryo morphology models \(M = N\). The prior model \(P(M)\) would then be a simple discrete probability distribution on \(N\) with essentially all the probability concentrated on the numbers between 0 and 8, corresponding to the prior knowledge that we only study embryos up to two days after fertilization and that they rarely have more than 8 blastomeres.

However, as a consequence of the simplicity of the morphology model \(M\), the model of the image formation is more complicated. All the visible variations of embryos with e.g. two blastomeres, like e.g. position, size and shape, span a huge set of possible images of embryos with two blastomeres. Since these geometrical variations are extremely nonlinear in the measured image pixel intensities, the conditional distribution \(P(I|M = 2)\), is difficult to describe, without explicitly modelling these geometrical variations. The point is, that even if we don’t consider certain variabilities as part of the morphology model, we need to consider them anyway as hidden parameters or nuisance parameters in the image formation model.

As another extreme, one could take the point of view, that whatever is visible in the images, is due to the morphology of the embryo, and thus choose to model the morphology of an embryo by its image, and so take \(M = I\). In this case the image formation model \((P(I|M))\) would be trivial, but so would the possible inferences with this model. The burden of extracting interesting parameters, like the number of blastomeres, would still have to be solved, and so we are left where we started with the problem of extracting morphology parameters from images.

If we choose \(M\) sufficiently rich to make the forward problem approximately deterministic \(I \approx I_M(M)\) and well-posed, the remaining variability can often be approximately modelled by a noise process expressed by a simple distance measure on the set of images \(I\) regarded as an Euclidean vector space of functions

\[
P(I|M) = P(\Delta I = I - I_M(M)) = \exp(-||I - I_M(M)||^2)). \tag{3.3}
\]

### 3.3.1 Multilevel Bayesian Model

The idea is now, based on the two previous observations of extreme morphology models and image formation models, to get the simplicity of both of them while avoiding their complex parts, by introducing multiple levels in the model. I shall denote the different levels by: \((I = )L_1, \ldots, L_{N_L}(= F)\). In this model \(L_n\) is considered “the image” of \(L_{n+1}\) as described by the model \(P(L_n|L_{n+1})\).

The image formation from \(L_{n+1}\) to \(L_m\) (assuming \(m < n + 1\)) is then described by (multiple application of):

\[
P(L_m|L_{n+1}) = \frac{P(L_m, L_{n+1})}{P(L_{n+1})} = \int_{L_n} dL_n P(L_m|L_n, L_{n+1})P(L_n|L_{n+1}) \tag{3.4}
\]

\[
P(L_m|L_n, L_{n+1}) \approx P(L_m|L_n). \tag{3.5}
\]

Similarly the posterior distribution of \(L_{n+1}\) may be inferred from the posterior distribution of \(L_n\) given an observed image \(I\) using (multiple applications of):

\[
P(L_{n+1}|I) = \frac{P(L_{n+1}, I)}{P(I)} = \int_{L_n} dL_n P(L_n|I)P(L_{n+1}|L_n, I) \tag{3.6}
\]

\[
P(L_{n+1}|L_n, I) \approx P(L_{n+1}|L_n) \tag{3.7}
\]

\[
P(L_{n+1}|L_n) = \frac{P(L_n|L_{n+1})P(L_{n+1})}{P(L_n)}. \tag{3.8}
\]
The priors \( P(L_n) \) of each level \( L_n \) can in principle\(^4\) be induced all the way from the prior \( P(L_{N_L}) \) of \( L_{N_L} \) using:

\[
P(L_n) = \int_{L_{n+1}} dL_{n+1} P(L_n|L_{n+1})P(L_{n+1}). \tag{3.9}
\]

With this model it is no longer obvious how to distinguish between the image formation model, and the prior morphology model. As we just saw, taking \( L_1 = I \) makes modelling of (the last step in) the image formation trivial. Similarly, taking \( L_{N_L} = N \) to be the number of blastomers, makes modelling of the prior morphology simple.

However, we still have to find suitable intermediate levels, for which the step from one level to the next is manageable to model.

### 3.3.2 Levels in the Image Formation

To arrive at a multilevel model of the image formation in Fertimorph lets first consider in general terms how the images are formed by the microscope and recorded by the camera (see Figure 3.2):

---

\(^4\)Please note that this would be the preferred priors in a clinical application setting, where these priors have been found from thorough statistical analysis of a lot of embryos and observation of their morphology and fertility. However, if the inferred morphology models are to be used in such a statistical investigation of these priors themselves, it would be important that the inference of the embryo morphology to a larger extend is based on the observed images rather than the particular choice of prior.

\(^5\)We are using transmitted light microscopy.

---
3. The “emerging light from the object” is focused, depending on the focus position of the microscope objective lens, to “an image on the ccd-chip” of the camera.

4. The camera measures the “image intensity” for each pixel.

Based on this, one could use the following intermediate levels of the model:

\[ L_1 = I \] The measured pixel intensity image.

\[ L_2 = L_e \] The light at the camera ccd-chip.

\[ L_3 = L_e \] The light emerging from the object.

\[ L_4 = V_{phys} \] The 3D volumetric model of the object physics.

### 3.3.3 Levels in the Object Model

What remains is an analysis of how to divide the object morphology model into useful levels. This has already implicitly been described by the planned computational steps of 3D reconstruction, segmentation and shape analysis:

1. The first step of segmentation consists of going from the reconstructed volumetric model \( L_4 = V_{phys} \) of the physical properties of the object to a volumetric model \( L_5 = V_{seg} \) of the local object segment or part type\(^6\).

2. The next step could consist of reducing to approximate surface models \( L_6 = O_{surf} \) of the segments.

3. In the case where the full object surface geometry \( O_{surf} \) cannot be inferred due to e.g. problems with the 3D-reconstruction of \( V_{phys} \), it may be necessary to introduce an intermediate partial geometry model \( L_7 = O_{contour} \) of for example object contours, which can be inferred directly from the images\(^7\).

4. From here one may reduce the description further to derived object and shape measures \( L_8 = O_{shape} \) like surface area, segment volume or number of parts.

To summarize, we have the following levels in the object model:

\[ L_4 = V_{phys} \] The 3D volumetric model of the object physics.

\[ L_5 = V_{seg} \] The 3D volumetric model of the object segments or parts.

\[ L_6 = O_{surf} \] The object surface models.

\[ L_7 = O_{contour} \] The object contour models.

\[ L_8 = O_{shape} \] The object surface or contour shape models.

---

\(^6\) Relevant segment types could be: Surrounding medium, zona pellucida, cytoplasm, ooplasm, 1. polar body, 2. polar body, fragment, blastomer, pronucleus, nucleus, nucleoli.

\(^7\) To fully appreciate the inclusion of this level, you should be aware of the negative conclusions of both 3D-volumetric reconstruction and 3D model-based object (surface) reconstruction under the approximate thin object model of light matter interaction known as anomalous diffraction. This lack of models for 3D-reconstruction will be discussed in more detail in the part on models of 3D reconstruction.
3.4 Stepwise Inference

We shall not use a fully Bayesian approach communicating the full prior and posterior distributions between the levels in the multilevel model, for the following reasons:

1. We do not have the necessary prior models $P(F)$, $P(\text{shape}|F)$, and $P(\text{phys}|\text{seg})$. Finding these are actually part of the biological objective of the Fertimorph project.

2. Since part of the biological objective of the Fertimorph project is to use the inferred morphology models $M$ for inference of these prior models $P(M|F)$ of embryo morphology, the morphology models cannot be inferred based on these. We can at most use, otherwise proposed priors, and would prefer these to be as "weak"\footnote{What precisely is meant by “weak” or how to define it precisely is not obvious. As a first shot at defining it one could try with the entropy of the distribution. However, this is only defined relative to a uniform measure or natural parameterization, which would then need to be axiomatically characterized for example in terms of invariance with respect to one or more symmetry groups.} as possible, in order for the inference to be as close as possible in character to an actual measurement rather than being a qualified guess mainly based on expectations.

3. Performing full inference in such a model is infeasible due to the many hidden parameters.

3.4.1 Stepwise MAP Decisions

Instead it is hoped that the fully Bayesian inference can be well approximated by stepwise model decisions of intermediate level models, based on weak or non-specialized priors. This should be the case, if the a posteriori probability distribution is well concentrated around a single model at the level where the Bayesian inference is concluded by a decision. I shall only consider MAP decisions because:

- We do not have a concrete loss function $L: M \times M \to \mathbb{R}$ at our disposal.
- Hopefully the posterior distribution is sufficiently concentrated (with respect to the metric $d$) that a near deterministic model of the inference computations makes sense.

Under all circumstances we need both the prior models and the models of the image formation, at each level.

3.4.2 Smoothed A Posteriori Probability

When the model space $M$ is large and the inferred posterior distribution $P(M|I)$ tends to be noisy, it may be preferable to use a regularized MAP decision or estimate of the model, by smoothing the posterior distribution according to some distance measure $d: M \times M \to \mathbb{R}$, prior to the determination of the MAP model.

Assuming for simplicity that $M$ is a finite dimensional vector space and $d$ is the Euclidean metric induced by the norm $||\cdot||$, this smoothing can for example be achieved by considering the linear Gaussian scale space, which I will try to express using the notation for an extra/intermediate level in the Bayesian model,

$$P(M_\sigma|I) = \int_M P(M|I)P(M_\sigma|M)dM,$$

where $\Delta M = M - M_\sigma$ is assumed to be Gaussian distributed with spread $\sigma$:

$$P(M|M_\sigma) = \frac{1}{Z_\sigma} \exp\left(-\frac{||M - M_\sigma||^2}{2\sigma^2}\right)$$

$$Z_\sigma = \int_M \exp\left(-\frac{||\Delta M||^2}{2\sigma^2}\right)d\Delta M.$$
The most important practical motivations for regularizing the posterior distribution in this way, prior to deciding on the MAP estimate of $M_\sigma$, are as follows:

1. It is believed that the later optimal decisions and inferences in the deterministic approximation depend continuously (with respect to the metric $d$) on the inferred model $M$. That is, models close to each other as measured by $d$, lead to the same optimal decisions and inferences.

2. I believe that the later decisions and inferences which optimally should be based on the posterior distribution $P(M|I)$ can be done approximately optimal, based on the approximate posterior distribution $P(M|M_\sigma,\text{MAP})$.

3. Regularizing the posterior distribution by smoothing will hopefully extend the catchment basin of attraction to the global maximum, making optimization by local searching more likely to find the global maximum.

3.5 How to Find Suitable Models

So far I have only discussed the model structure. Now I will shortly discuss alternative ways of finding suitable models for this structure. I see two different ways of attacking this problem.

3.5.1 Theoretical

Without looking at the concrete images, I could try to find a suitable model based on theoretical knowledge. I consider this a kind of qualified guessing. However, since it is rooted in well established theories, it will have the advantage of being related to these, and associated with a number of assumptions under which the model is known to hold. The success of this approach will depend on the accuracy and agreement between model and observed images - that is finding a theoretical model for which all the assumptions are fulfilled. Another quality of this approach, is that the task specific parts will be separated from the more general parts, and so we will get more generally applicable methods. Of course this theoretical approach can only take us as far as the theories go. As soon as we start to deal with effects that have not been described theoretically or is difficult to model, we may have to use the more pragmatic empirical approaches at modelling.

3.5.2 Empirical

Rather than (qualified) guessing at a model and just using it for inference, I could try to estimate a good model, from empirical observations. That is, if we have suitably many examples of images annotated with the true morphology models, we can try to use statistical methods for estimating models of the relationships between embryo morphology and observed images. Of course such an estimation would have to be done within some model frame, which would somehow have to be guessed or otherwise proposed.

The advantage of this strategy for finding models, is that it can potentially model effects which we didn’t think of theoretically, and thus is more likely to match the real observed images accurately. An obvious disadvantage is that we need a large representative set of examples which have been annotated with the true morphology models. Another disadvantage is that we can only hope to model the morphology which we are able to annotate by hand - that is, we must by some other means be able to arrive at the true model. This means we can at best hope to automate existing manual procedures, and will not be able to arrive at 3D morphological measurements that cannot be performed by hand.
3.5.3 Image Formation

Regarding the modelling of the image formation, the theoretical approach would imply that I should study the literature on microscopy, and the physical theories of light and optics, and hope to find an approximate model which is both computational feasible and accurate enough to explain the images we see.

3.6 Thesis Overview Revisited

I conclude this last chapter by giving a short overview of the remaining 3 parts of this thesis, relating the different contributions to the different structures of the abstract Bayesian inference model of image analysis in Fertimorph.

There are three remaining parts of this thesis:

II Towards 3D reconstruction

III Segmenting embryos in HMC-images

IV Shape analysis

In part II, theoretical models of the image formation and methods for reconstruction are analyzed. In chapter 4, light microscopy in general and two contrast techniques for studying refractive objects are presented together with a short discussion of potential approximate models of light microscopy image formation ($P(L_e|V_{phys}, L_i)$, $P(L_e|L_i)$ and $P(I|L_e)$). In chapter 5 assumptions, applicability and complexity of existing models of reconstruction ($P(V_{phys}|I)$, $P(V_{phys}|L_e, L_i)$, and $P(L_e|L_i)$) are analyzed with respect to the different approximate models of light microscopy. In chapter 6 a study of light microscopy images of human oocytes is presented and analyzed with respect to dominating optical phenomena in the image formation. In chapter 7, an approximate linear model of the HMC image formation $P(I|V_{phys})$ is presented, which can be analyzed analytically and allows for a fast reconstruction of an optical thickness profile, which is the part of $V_{phys}$ which is not lost in the imaging process according to the model.

In part III, a model (chapter 8) and user interactive method (chapter 9) for direct segmentation of blastomers in embryo, based on an approximate direct model of the image formation $P(I|O_{sphere, surf})$ from an approximate morphology model $O_{sphere, surf}$ of spherical blastomers together with a simple prior model of non-overlapping blastomer morphology $P(O_{sphere, surf})$, is presented and evaluated (chapter 10).

In part IV, a Lie group model of point set variability, capable of modelling globally well defined shape variabilities is presented and analyzed for point sets in 2D (chapter 11) and in general dimensions (chapter 12). The intended purpose of these models of variability is to serve as model constructs in generative models $P(O_{points}|O_{point,shape})$ of prior point set distributions $P(O_{points})$ and the induced shape distributions $P(O_{point,shape})$. My main motivation for studying models in the ill-posed generative direction $P(O_{points}|O_{point,shape})$ rather than the other well-posed direction $P(O_{point,shape}|O_{points})$, is the intended use in few sample inferences of empirical models of $P(O_{point,shape})$. 
Part II

Towards 3D Reconstruction
Chapter 4

Microscope Image Formation

In this chapter I will discuss models of the image formation. The image formation is a consequence of light emission from the light source of the microscope, light propagation, light-matter interaction (with both the object and the lenses and filters of the microscope) and finally absorption when measuring the light intensities of the image. I thus need to present theories from physics, optics, and microscopy. Since there is plenty of literature on these well-established theories, I will give references to these and only shortly review central models and results and instead discuss which aspects of the theories may be of relevance to the modelling of the image formation in light microscopy of human embryos.

In chapter 7 on “Reconstructing the optical thickness from HMC-Images” I will discuss in more detail the Hoffman Modulation Contrast technique and various image formation model approximations.

4.1 Light Microscopy

General references to light microscopy include:

- “An Introduction to the Optical Microscope”[12]
- “Contrast Techniques in Light Microscopy”[13]
- “A symmetrical representation of the geometrical optics of the light microscope”[22]
- “Video Microscopy”[26]

In addition to the principles of construction and operation of the microscope (see Figure 4.1), central themes are the generation and control of

- Magnification
- Resolution
- Contrast

in the formed images. Magnification is discussed in terms of a geometrical optics ray model of the large scale geometry of image formation, as illustrated in Figure 4.1 for a microscope using infinity corrected optics. For quantitative accounts of the geometrical optics model of image formation and aberrations by lenses I refer the reader to [11, 50].

Since the purpose of light microscopy is the study of details of small objects with dimensions down to the wavelength of light, resolution in light microscopy images has since Lord Rayleigh
Figure 4.1: Left: Microscope principle of image formation. Notice how parallel rays impinging on the object originate from the same point in the front focal plane of the condenser, and parallel rays emerging from the object are focused to a “conjugate” point in the back focal plane of the objective. Middle: Principle of controlling object illumination. Here shown for the most often used Köhler illumination, where an auxiliary lens collects the light from the source and focuses it in the front focal plane of the condenser. This achieves an incoherent superposition (from the different parts of the light source) of planarly coherent light waves impinging on the object, although it is more often emphasized that it achieves a uniform object illumination. The directions of illumination are controlled using the illuminating aperture. The field of illumination is controlled using the field diaphragm. Right: Geometrical ray model of image formation in light microscope using infinity corrected optics. The objective lens forms the image of the object “at infinity”, where the lens has been optimized to produce the best image. This “image” is then imaged by the auxiliary lens on to the camera ccd chip. In this figure the distance from the objective lens to the auxiliary lens achieves a magnification of 2.

and Abbe (cf. [55, p.687]) been discussed in terms of a wave model of diffraction effects in the focusing of fine details. The high-quality optics used in microscopes are produced with an accuracy so high that they are diffraction limited. This means that they have been refined to an extent that they can focus the light so precisely that it is the wave-nature of light and the finite aperture of the microscope which because of diffraction effects set the limits of resolution. It is this source of diffraction effects in light microscopy which achieves the most attention in the microscopy literature. Quantitative models of lens image formation taking diffraction into account are most easily expressed using the theory of Fourier optics (see e.g. [21, 20, 50]). Shortly stated: In this theory a scalar monochromatic model of the electromagnetic field in a plane is Fourier analyzed with respect to the two spatial coordinates in the plane. The Fourier transform of the scalar field in the plane is called the angular spectrum, since it can be interpreted as weights in a superposition of plane waves, propagating in different directions.

Discussion of contrast in terms of more quantitative modelling of intensity variations is much more task specific. It depends on both the specific object characteristics and the contrast technique used. As light microscopy is often used for viewing biological specimen much literature on microscopy is in this context. Since biological specimen are often transparent, much effort is put into staining or fluorescent marking of specific parts of biological interest. There are however

---

1 Lens aberrations are however according to [50] most often treated using geometrical optics in the microscopy literature although the Fourier optics treatment is more general and accurate.

2 Spatial frequencies higher than $1/\lambda$, $\lambda$ being the wavelength, does not correspond to plane waves but are known as evanescent waves, as they are strongly attenuated when propagating[21, p. 51]
also a number of non invasive contrast techniques useful for studying the refractive properties of objects. See [13] for an overview. Below I shall only shortly review the principles of Zernike’s Phase Contrast and Hoffman Modulation Contrast.

4.1.1 Zernike’s Phase Contrast

The most often used construction of Zernike’s Phase Contrast technique[55, 56], is illustrated in Figure 4.2. The principle of contrast generation is to change the phase of the undiffracted light (the directly transmitted illumination) by about 1/4 of a wave, and thereby allowing it to interfere constructively or destructively with the waves diffracted by the object. For a quantitative treatment of the contrast generation in Zernike’s Phase Contrast technique the reader is referred to [4] and [11, p.472-476]. It will also be discussed slightly more quantitatively in terms of the Born approximation in section 5.2.4 of chapter 5.

4.1.2 Hoffman Modulation Contrast

Hoffman Modulation Contrast (described first in [24, 25, 23]) is a light microscopy contrast technique, well suited for in vivo studies of biological specimen, because of its non-invasive contrast generation for transparent but refracting objects[23]. This kind of objects are often referred to as “phase objects” in the microscopy literature, since they interact with the light by only influencing the phase of the light. This is opposed to semi transparent absorbing objects, which mainly influences the amplitude.

The HMC technique generates contrast by converting “phase gradients” or “optical gradients” into intensity variations in the microscopy images[24] such that positive gradients show up bright and negative gradients show up dark. Figure 4.3(LEFT) shows an example HMC image. This results in images that give a human observer the familiar but in this situation false impression[24] of looking at the highlights and shadows resulting from a surface with height variations, shined upon from one side.
4.2 Physics of Light Microscopy

Central questions in the search for a hopefully linear model of the image formation $I(V_{phys})$ have been, what (approximate) models of the

- Light
- Light emission
- Light propagation
- Object
- Light-object interaction
- Microscope image formation
- Camera intensity measurement

to use. In general to answer these questions we should know which of the optical phenomena:

- Polarization
- Color
- Diffraction
- Temporal Coherence
- Spatial Coherence

are showing their effects in the image formation, and which of the optical properties of the object, like:
are dominating the light-object interaction.

Realizing that the dominating physical light-object interaction is refraction due to spatially varying light speed in different parts of the object and the surrounding medium, I have spent quite some time trying to find an approximate model of the image formation, which can be used to reconstruct the 3D spatial variation of the light speed or equivalently the 3D spatial variation of the refractive index.

4.2.1 Approximate Models of Light and Light Matter Interaction

In the case of a refractive light-object interaction there are still quite a few different approximate models available depending on which regime of refractive objects the studied object belongs to. The regime is characterized in terms of both the spatial dimensions of the object as well as the magnitude of variation of the refractive index. The problem is not so much to find a model that is accurate, but rather to find one that is also computational feasible while still being accurate enough.

From physics we have the following (approximate) models of light propagation and interaction with transparent objects:

- Quantum Electro Dynamics (The most general model)
- Maxwell’s equations (general $n$)
- Scalar Diffraction theory ($n$ smooth or with small variations) In the scalar wave theory of light one ignores the vector nature of the electromagnetic field, and simply considers a scalar field representing the “total amplitude” of the electromagnetic field. Of course all aspects of polarization and related phenomena are ignored in this model.
- Anomalous Diffraction (thin lens approximation) ($|n - 1| << 1$, $\beta >> 1$ while $\beta (n - 1)$ approximately constant)
- Rayleigh-Gans theory (Born Approximation) ($|n - 1| << 1$ and $\beta |n - 1| << 1$)
- Rayleigh scattering ($\beta << 1$ and $n \beta << 1$)
- Geometrical Optics ($\beta >> 1$). Geometrical Optics is the limiting theory of Maxwell’s equations in the limit of small wavelengths.
- Euclidean Geometry ($n$ constant).

Here the regimes of applicability (with good approximation accuracy) have been noted in terms of the following light and object characteristics:

- speed of light: $c$
- wavelength: $\lambda$

---

3Variations in light speed.
4Changes in polarization.
• Characteristic object size (Diameter): \( D \)

• Size parameter: \( \beta = \frac{2\pi D}{\lambda} \)

• Relative refractive index: \( n = \frac{n_{\text{object}}}{n_{\text{surrounding}}} = \frac{c_{\text{surrounding}}}{c_{\text{object}}} \)

A full treatment of these is outside the scope of this thesis. Except for the theory of QED, which is described qualitatively in [19], they are all described quantitatively in [11], to which the reader is referred. Their applicability to 3D reconstruction of human embryos is discussed in chapter 5.
Chapter 5

Models and Methods for 3D Reconstruction

In this chapter assumptions, complexity and applicability of existing methods and models of reconstruction are analyzed. Focus is on the most attractive linear translational invariant model, which would allow for direct reconstruction of a volumetric model by use of a deconvolution method. This corresponds to direct inference of \( P(V_{\text{phys}}|I) \), based on a simple model of \( P(I|V_{\text{phys}}) \).

5.1 Direct Reconstruction of a Volumetric Model

The possibility of having a large number of volumetric morphology model parameters \( M = V_{\text{phys}} \), that are also “closely related” to the observed image intensities \( I \) is very appealing. Both as a way of delaying the incorporation of task specific modelling, as well as from a computational point of view.

5.1.1 Translational Invariance

If an image is represented as a real function of 3 real \(^1\) (spatial) coordinates (say row, column, and focus position):

\[
I = \{ I : \mathbb{R}^3 \rightarrow \mathbb{R} \}.
\]

(5.1)

What I mean by a “closely related” morphology model is, that it has the same structure:

\[
M = \{ M : \mathbb{R}^3 \rightarrow \mathbb{R} \},
\]

(5.2)

and obeys a translational invariant image formation model:

\[
P(I|M) = P(\tau_{\Delta x}^{-1}(I)|\tau_{\Delta x}^{-1}(M)), \forall \Delta x \in \mathbb{R}^3,
\]

(5.3)

where \( \tau_{\Delta x} \) is the translation operator:

\[
(\tau_{\Delta x}^{-1}(f))(\vec{x}) = f(\vec{x} - \Delta x), f : \mathbb{R}^3 \rightarrow \mathbb{R}, \vec{x}, \Delta x \in \mathbb{R}^3.
\]

(5.4)

The idea is that \( M \) should represent the spatial variation of some physical property (like e.g. mass density, fluorescence concentration, linear attenuation coefficient or in the case of transparent

\(^1\)In reality we only have a discrete number of regularly spaced spatial samples within a compact domain, but this continuous model is mathematically convenient to work with - at least when discussing translational invariance.
phase-objects like human embryos, the refractive index) of the embryo, which gives rise to the observed image, and that the image only depends on the position of the embryo by a similar positioning of the image.

The hope is, that having such a general morphology model with a structure closely related to the images, allows for a simple image formation model where the image intensities at a given position in the image, depends mainly on only a few model values around a corresponding position in the model of the spatially varying physical property. Furthermore, by having such a general morphology model, it is hoped that the variabilities of the images not explained by it will be of a kind that can be explained by stochastically independent noise on the image intensities. The conditional probabilities \( P(I|M) \) of the image given the model can then easily be modelled by the standard \( L_2 \) metric on the set of images represented as functions:

\[
P(I|M) = \frac{1}{Z} \exp\left( -\frac{\|I - I_m(M)\|^2}{\sigma^2} \right),
\]

where \( Z \) is a normalizing constant, and \( I_m(M) \) denotes the model predicted ideal noiseless image of an embryo with “morphology” \( M \).

In the setting of light microscopy imaging, I believe this property of translational invariance is a reasonable approximation of the image formation.

5.1.2 Linearity

The math gets really simple, if we can further assume linearity of the image formation, in the sense

\[
\forall a, b \in \mathbb{R}, M_1, M_2 \in \mathbf{M} : I_m(aM_1 + bM_2) = aI_m(M_1) + bI_m(M_2).
\]

Linearity and translational invariance makes the image formation particularly simple, since it then has the form:

\[
I_m(M) = \iiint_{\mathbb{R}^3} M(\vec{x}) \tau_{\vec{x}}(I_m(\delta_0)) d\vec{x},
\]

where \( I_m(\delta_0) \) is the image of a unit point object \( \delta_0 \) only located at origo. Intuitively this can be understood as follows: Any object model \( M \) is considered a linear combination of point objects, and due to the linearity the image is then just this linear combination of images of the point objects. The translational invariance ensures that all these images are just translated versions of one and the same image.

5.1.3 Reconstruction by Deconvolution

Under the assumption of linearity and translational invariance, we essentially only need to know what the image of a point looks like in order to model the image formation. Such linear translational invariant models have been used for many years in the field of optics and communication.

\(^2\)The notion of “a corresponding position” is not really well defined. However the assumption of translational invariance ensures that there is a notion of spatial structure of the dependencies independent of the position. This will become even more concrete in terms of the point spread function under a further assumption of linearity of the image formation.

\(^3\)I will not go into the problem of defining how to normalize, in the general setting, where \( I \) is a set of functions. This would lead to a need for defining more exactly what I mean by probability distributions, and the mathematics of measure theory or distribution theory, or maybe even further?... Instead I will note that when we at the end need to compute things, we will deal with finite sets, since they should be representable in a computer, and consequently I can think of the values as probabilities of a finite set of images, and thus simply require that they add to one.

\(^4\)This belief is backed up by the fact that 3D reconstruction by deconvolution is partly based on this assumption, and that this technique has proved useful to an extent that there are commercial software solutions available for 3D reconstruction by deconvolution. However talking to Gabor Szekely, ETH, Zürich, when he visited 3D-Lab, I got the impression, that it was more often violations of translational invariance than lack of linearity that was the biggest problem for these methods.
theory of electrical engineering (cf. [21]). Their basis on Fourier analysis allows for analytical analysis of the inverse problem, and regularized reconstruction by deconvolution, as described in e.g. the first part of “Inverse Problems in Imaging”[6] on “Image Deconvolution”.

5.2 3D Light Microscopy by Deconvolution

Application of image deconvolution to 3D-reconstruction from a stack of optically focused light microscopy images has been investigated since the 1980’s[1, 2], and developed to an extend that there are now a large number of (commercial) software products available\(^5\). However, they are all based on a translational invariant linear model of the image formation.

In order to discuss the validity of the linear translational invariant model we need to settle on the physical object modality to reconstruct and check if it is in agreement with the physical models of the different steps in the image formation.

5.2.1 Fluorescent Emission

The linear translational invariant model works well for fluorescence microscopy where the spatial variation of fluorescent concentration is the physical property of the object which is reconstructed. This is the physical object modality that these methods were developed for.

Since a fertilized human embryo, which hopefully at some point should become a child, for obvious reasons cannot be fluorescently marked, the direct application of 3D reconstruction of the fluorescent concentration wouldn’t make sense. The use of autofluorescence is also not a possibility. As far as I understand, this would entail the usage of a triggering light source in the (near) ultraviolet spectrum, which is not healthy for the embryos.

5.2.2 Absorption

Instead it was believed that the main reason for the observed contrast in the images was due to absorption of light. And so the original plan was to use these linear models of image formation, on a suitably logarithmically transformed image, since the end intensity of a ray would qualitatively depend on the product of the traversed voxel (Volume × Element) transmittances. After a logarithmic transformation these products would become sums or integrals and consequently we would have the basis for a linear model, like in X-ray computed tomography (see e.g. [6, p.195]) but with a different geometry and mixing of the ray integrals. It is however not the absorptive properties of human embryos, that are dominating the image formation. Instead it is the refractive properties of human embryos together with diffraction effects from the wave nature of light which is the dominating light matter interaction and cause of image intensity variations.

5.2.3 Refraction

Human embryos are, like many microscopic biological objects, to a first approximation transparent. That is, only a small amount of the incoming light is either reflected or absorbed - most of it is transmitted. Still the light is not completely unaffected by a human embryo. The light propagates at different speeds inside different parts of the embryo. In itself this change of speed cannot be measured directly, since the change of speed only lasts while the light is inside the embryo. It is only secondary interference effects due to the diffraction propagation of light that can create intensity variations in images of the object. This can e.g. be achieved by imaging

it slightly out of focus, or using a more sophisticated contrast technique like Zernike’s Phase Contrast or Hoffman Modulation Contrast.

To understand how these delays inside the embryo have affected the transmitted light and how the used microscopy contrast technique works,

it is central to model the wave nature of the light

and, to do this (using a simple approximate model)

it is important to know which regime of refractive objects embryos belong to.

Approximate models for different regimes are listed in section 4.2.1 p. 39.

5.2.4 Diffraction

In the limit of very weakly refractive objects, only the phase of the light is influenced by the object, and the intensity variations are caused by diffraction. For this regime of objects the light matter interaction and light propagation is modelled well by the Born Approximation (see e.g. [11, p.699]) also known as Rayleigh Gans Theory in the Optics Field[49]. In this model the light matter interaction is described as if every inhomogeneity in the refractive index emits a small “diffracted” spherical wave which is 1/4 wave (π/2 radians) out of phase with the incoming light[13, p.61]. The amplitude of these diffracted waves depend linearly on the deviations of the refractive index under the Born approximation. So there is hope for a linear model, but we have only come to the emerging light \( L_e \) (see figure 4.1 p. 36). We are looking for a linear model all the way to the image intensities \( I \). When these diffracted waves are brought to a focus and interfere with the undiffracted light, they unfortunately only change the phase of it. Since the measured image intensity is proportional to the time average of the squared modulus of the electro magnetic field, it is independent of the phases of the light. We thus see, that it is only at the very last step in the image formation, that the camera’s non-linear mapping of the electro magnetic field \( L_e \) on the ccd-chip to the measured image intensity \( I \), spoils the possibility of a non-trivial linear model of the mapping from a volumetric model of the refractive index \( V_{phys} = n(x, y, z) \) to the image intensity deviations \( I - I_{bg} \).

In fact, a small weakly absorbing object can also be treated in the Born approximation model as “emitting” a small diffracted wave which is 1/2 wave (π radians) out of phase [13, p.60]. When these waves are brought to focus they interfere destructively with the undiffracted light and cause intensity deviations as expected.

Zernike’s Phase Contrast

The regime of objects, where the generated contrast depends linearly on the total phase change when using Zernike’s phase contrast technique, is characterized by the maximum phase difference between the delayed and undelayed light being no bigger than approximately 15 deg or 1/24th of a wave[4, p.13]. This linear model of Zernike’s phase contrast is also treated in [11, p.476].

For this regime of objects Zernike’s phase contrast technique generates images equivalent to bright field images of the corresponding virtual amplitude object which Zernike’s phase contrast technique “transforms” the phase object into[55, p.697]:

The phase grating will therefore appear as an amplitude grating of exactly corresponding structure ...

\(^6\)As measured by the relative deviation from background intensity

44
In [56] the validity of the theory is extended from periodic phase gratings to arbitrary phase objects (but still modelled by the Born approximation). We should thus in theory be able to follow our original plan of direct reconstruction of a volumetric model of the refractive index by deconvolution\(^7\) of the observed images, provided we use the phase contrast technique, and the studied embryos fall in the regime of the Born approximation.

5.2.5 Deflection

For large phase objects the total phase changes experienced by the light is likely to be so large, that the linear approximations of diffraction break down. Instead there is hope that the geometrical optics model of ray propagation and deflection at interfaces between volumes with different refractive indices, as described by the law of refraction (Snell's law) and the Fresnel formulae for the transmitted intensities (see e.g. [11, p.38-43]), can approximately describe the image formation.

Modelling the refraction effects by ray deflections using the geometrical optics model, is problematic to interface with a typical regular grid discretization of a volumetric model. The geometrical optics model of ray deflections interface more directly with a geometrical model of the refractive object discontinuities. The ray deflections are critically dependent on the local geometry of the discontinuities of the refractive index. Taking the geometrical model to its extreme, the rays are not deflected when meeting a discontinuity perpendicularly, which would be the approximate case for rays running near parallel to one of the voxel grid axes. Using a discrete volumetric sampling of the refractive index, one would need to smoothly interpolate before modelling the propagation of light rays. However interpolating smoothly would also smooth away all the discontinuities, and so the discrete deflections of rays according to Snell’s law would have to be substituted by a differential equation of light rays (see [11, p.130]) of the form:

\[
\frac{d}{ds} \left( n \frac{d\vec{r}}{ds} \right) = \nabla n. \tag{5.8}
\]

5.2.6 Anomalous Diffraction

We notice that the right hand side of equation 5.8, being the gradient of the refractive index, depends linearly on \(n\). So there is here a small hope of approximate linearity if we formally consider the mapping of \(n\) to the integrated ray deflections (for all rays):

\[
\int n(\vec{r}(s')) \frac{d\vec{r}}{ds'} ds' = \int (\nabla n)|_{\vec{r}(s')} ds'. \tag{5.9}
\]

Noticing that the integrals run along ray paths which obviously depend on the magnitudes of the previous deflections, it is very difficult to keep the hopes up. However for objects in the regime of anomalous diffraction, there might just be a possibility, although it comes at the price of reducing the object model to an essentially 2-dimensional optical thickness profile. This model is discussed further in chapter 7.

Even, if we this way are able to make the geometrical optics model interface with the volumetric model of the refractive index, and possibly in an approximately linear way when considering the mapping to integrated ray deflections, we still need to remember the last step of the image formation - that of intensity measurement. For small deflection angles, which we definitely would need to assume if we want a chance of linearity, the intensity in the image by a ray is approximately independent of the direction of it. We are in a situation parallel to the one where

\(^7\)It seems that a logarithmic transformation is not needed under the first Born approximation, since it only considers diffraction from the assumed undisturbed incoming light, and not second order diffraction known as multiple scattering. The second order Born approximation and multiple scattering is discussed in [11, p.708-710].
the linear Born approximation model was spoiled by the non-linear intensity measurement. The cure for that was Zernike’s Phase Contrast technique. The “cure” now is Hoffman Modulation Contrast.

Hoffman Modulation Contrast
Studying the literature of Hoffman Modulation Contrast [23] a central ingredient in understanding the contrast technique is the finite change in direction experienced by a ray when deflected by the object. The linearity of the image intensities created using this contrast technique is discussed further in chapter 7.

5.2.7 Limits of Applicability
It is worth noting here, that violation of the central assumption of rectilinear ray propagation in the linear models of image formation when dealing with fluorescent or absorptive properties of an object, is the central ingredient in the model of HMC image formation. When the refractive index deviations are not in the regime of the Born approximation, the wave propagation of light no longer follows an approximate Euclidean geometry. It is this violation that is the origin of the major difficulties of reconstruction of refractive objects.

These difficulties have been noted in [47] p. 305. The linear translational invariant model has its limits of application, when dealing with:

- reconstruction of a “Thick Specimen”
- reconstruction of “phase object”.

The problem is lack of linearity in the image formation due to deviations from rectilinear light propagation.

5.3 Reconstruction of Human Embryos
It is the wave nature of light together with spatial differences in light speed, which results in the refraction of the light as it passes between different parts of the embryo.

According to [41] the refractive index of mammalian cells\(^8\) is around 1.37. Compared to the refractive index of water, which is approximately 1.33, this is only about 3 percent higher. This means that visible light with a wavelength of \(1/2\mu m\) in vacuum will have a wavelength of \(\frac{1/2\mu m}{1.37} \approx 0.38\mu m\) in water, it will have a slightly shorter wavelength of \(\frac{1/2\mu m}{1.33} \approx 0.36\mu m\) inside the embryo. Not only are the variations of refractive index between inside an embryo and a surrounding aqueous medium small, the spatial extend of the different structures is also small compared to daily life measures (the diameter of a human embryo is approximately 120 micrometers). However, compared to other human cells, the embryo is very big - about an order of magnitude. It is the spatially integrated difference in light speed (or refractive index) which results in the spatial differences in the phase of a coherent wave. Using the above numbers we get a maximum phase difference of approximately:

\[
120\mu m\left(\frac{1.37}{1/2\mu m} - \frac{1.33}{1/2\mu m}\right) = \frac{120\mu m}{1/2\mu m} \times (1.37 - 1.33) = 10(waves).
\]

So the variations in light speed within a human embryo are both rather big and rather small, depending on what you compare to. Compared to the regime of linearity of Zernike’s phase

---

\(^8\)Finding specific literature on the spatial variations of refractive index of human embryos is not easy. This general number for mammalian cells is however in approximate accordance with the number of 1.38 concluded from a study of embryo optics reported in chapter 6 of this thesis.
contrast, the delays of light in an embryo are about two orders of magnitude too large. But
compared to the dimensions of the embryo itself, the corresponding difference in optical thickness
is about an order of magnitude smaller:

\[
\frac{10(\text{waves}) \times 1/2\mu m}{120\mu m} = 1/24.
\] (5.11)

As human embryos are both large and refractive, and it is the refractive property of the embryo
which is the main origin of contrast in Hoffman Modulation Contrast images, the first planned
route of making a linear translational invariant model of the image formation seems only possible
using the anomalous diffraction model, in combination with the Hoffman Modulation Contrast
technique.

### 5.3.1 Alternate Models for 3D Reconstruction

If one wishes a true 3D object model reconstruction, a more accurate model of the image forma-
tion is needed and possibly using another contrast technique. Such a model would then have to
be analyzed with respect to a method of reconstruction. However, leaving the linear translational
invariant models of image formation, it is likely that the problem of reconstruction will be much
more complicated to analyze and compute. Based on the above observation, that it was the
non-linear intensity measurement of the electromagnetic field which introduced the problems, I
believe it will be beneficial to study the different steps \( P(L_c | f) \), \( P(L_e | L_c) \), and \( P(V_{\text{phys}} | L_e, L_i) \) in
reconstruction separately.

Modelling \( P(I | L_c) \) and \( P(L_c | L_e) \) are non-trivial problems, especially if diffraction phenomena
in temporally partial coherent light outside the regime of quasi-monochromatic light are to be
treated. They have been studied and modelled for many years, using both the wave model
of light and the geometrical optics based ray model [21, 20, 50, 11]. When dealing with either
completely incoherent or completely coherent illumination in the wave model, linear models of
diffraction limited image formation are the basis for linear methods of reconstruction (see e.g.
[6, p.60-64]) by deconvolution.

Because of the non-Euclidean geometry of light propagation within refractive objects, it is
my impression that the main difficulty of true 3D reconstruction of embryos lies in modelling
\( P(L_e | V_{\text{phys}}, L_i) \) in a way that is both sufficiently accurate and computational simple to recon-
struct \( V_{\text{phys}} \) from \( L_e \). The problem of reconstructing the refractive index from the emerging light
field is the topic of both:

- Inverse diffraction tomography, and
- Inverse scattering theory.

### Diffraction Tomography

Inverse diffraction tomography is however traditionally (see e.g. [11]) based on the Born approx-
imation, which we have already argued as insufficient for modelling the optics of embryos. Both
[11, p.715] and [6, p.216] discuss the possible advantage of basing the model on the first Rytov
approximation (see e.g. [11, p.726-729] for a short account), which to my understanding is a
linear approximation of the phase of the wave field rather than the field itself as in the Born
approximation, when dealing with large objects. But they both also emphasize that there is
a relationship with the first Born approximation, and that it leads to the same mathematical
problem as the Born approximation. My understanding of the lack of applicability of the Born
approximation to embryo optics, as due to multiple scattering causing a delay of the light with

---

\( ^9 \)White light from an electric bulb has a longitudinal coherence length of approximately 1 micrometer[33]
an order of magnitude of 10 waves, also applies to the Rytov approximation. Furthermore, from the abstract [27] I got the impression that Rytov approximation based diffraction Tomography leads to a non-linear relationship between the scattering amplitudes and the object model. Based on these two hints of problems and difficulty of 3D reconstruction of human embryos using the theory of diffraction tomography based on the Rytov approximation, this has not been investigated further. Anyway we would also still have to solve the problem \( P(L_c|I) \), and in particular \( P(L_c|I) \).

**Inverse Scattering Theory**

The mathematical theory of 3D reconstruction of both absorptive and refractive objects from far field scattering measurements based both on the full electromagnetic theory of light matter interaction as described by Maxwell’s equations and the acoustic theory of scalar wave propagation have been analyzed in [15]. Discussing the problem of 3D reconstruction of human embryos with among others, Reiner Kress and Thorsten Hohage, at the “TMR Workshop on Inverse Problems (26-27 June, 200)”, I got the impression, that it would not be a simple task to apply this theory in Fertimorph. In particular, the computational methods studied by Thorsten Hohage used a regularly sampled discretization of the object and the light field model, for solving the forward problem. The grid spacing being in the order of magnitude of 1/10’th of a wavelength or smaller. Since embryos are rather big objects when measured in terms of optical wavelengths, with a characteristic dimension of 240 wavelengths, such a discrete model would need a grid with at least about \( 2400^3 = 13.824 \times 10^9 \) nodes, which is still somewhat above the capabilities of computers today. Instead they suggested to investigate geometrical optics models or alternative representations of the object and the light field using specialized basis functions.

**5.4 Conclusion**

Different parts of the embryo may have different optical characteristics and thus interact with light in different ways. Differences in characteristic size also result in the qualitatively different wave or particle phenomena of diffraction and ray propagation. Depending on the optical phenomena under study and the preferred method of visualization, different optical microscopy techniques may be used. In general a detailed analysis of the mutual constraints between

- the biological structures under study
- the choice of optical modality to reconstruct
- the microscope optics chosen for image formation
- the complexity of modelling image formation
- the complexity of reconstruction

is needed, in order to uncover the possibilities of 3D-reconstruction.

In chapter 6 we shall investigate optical properties and phenomena in light microscopy of human embryos. In chapter 7 the Hoffman Modulation Contrast technique is analyzed with respect to linearity when dealing with anomalous diffracting phase objects, and a method for reconstruction using this model is presented.
Chapter 6

A Study of Embryo Optics

In this chapter I present experimental investigations of embryo optics and transmitted light microscopy image formation using bright field and Hoffman modulation contrast. The purpose of analysing these experiments is to gain insight into which regime of phase objects embryos belong, and which optical effects are present in the image formation.

6.1 Introduction

In the spring 2000, Christina Hnida and I (Niels Holm Olsen) recorded our first focused image sequences of a few human oocytes, which were left over from fertility treatments. The purpose of these first manual acquisitions of image sequences were both related to the technical procedures involved with operating the microscope as well as the planned biological investigations of fertility-related morphological structures. Equatorial images of the studied oocytes are shown in Figure 6.1. In this chapter I analyze the observed focus sequences (see appendix I), in order to investigate the following questions:

- How absorptive are embryos?
- What is the order of magnitude of the refractive index of a human embryo?
- How large are the consequent variations in optical delays of the light?
- How large are the resulting deflection angles of different light rays?
- Is the approximate model of geometrical optics with a radiometric model of light energy propagation a good description?
- Are interference effects and diffraction phenomena needed to explain the images?
- Is the spectral properties of the light needed to understand the images?
- Are there color dependent effects in the images?
- Are polarization effects needed to explain the images?

Based on analysis of bright field images, examples of which are shown in Figure 6.2, focused down to about 1000\(\mu m\) below the oocytes (away from the light source) and interpretations of the observed images as a superposition of images of the microscope illuminating aperture formed by local focusing of different parts of the oocytes, we conclude:
Figure 6.1: Transmitted light microscopy images focused at the equatorial plane of the studied objects. The first row contains Hoffman Modulation Contrast images. The second row contains bright field images using a small almost point like illuminating aperture approximately shaped like a polygon with 8 sides of equal length. The last row contains bright field images with oblique illumination from a ring shaped aperture.

- When observed in the surrounding aqueous medium, human oocytes, which have an approximate diameter of 60µm, belong to the regime of large but softly\(^\text{1}\) refractive objects with a relative refractive index of \(1 + 0.033\).

- Oocytes have a refractive index of about 1.38.

### 6.2 Microscope Equipment and Image Acquisition

We used a Nicon inverted light microscope, equipped with a long working distance 0.5 aperture condenser. It was equipped with both objective and condenser aperture turrets with matching pairs of objectives and condenser filters, allowing for fast and easy changing between different magnifications and contrast techniques. We used the FertiGrab system\(^\text{2}\) for image acquisition.

#### 6.2.1 Hoffman Modulation Contrast

We used a Nicon MC3 40X Hoffman Modulation Contrast objective together with the corresponding HMC condenser module filter in position for HMC image acquisition.

\(^1\)The model of anomalous diffraction is most accurate for the light passing through the central part of the cell. In the geometrical optics model, the light rays passing near tangentially to the surface of the cell, do experience arbitrary high gradients in the optical thickness profile. However, since this is at the contour of the cell, the absolute optical thicknesses are very small. Furthermore, diffraction effects due to the finite wavelength of the light, ensures that the “rays” are not infinitesimally thin and so they experience a diffraction-smoothed optical thickness profile.

\(^2\)From IH-Medical, Image House A/S, Denmark
Figure 6.2: Transmitted bright field light microscopy images focused at planes far below the equatorial plane (away from the light source) of the studied objects. The first row contains axial bright field images using a small almost point like illuminating aperture approximately shaped like a polygon with 8 sides of equal length. The second row contains oblique bright field images using a ring shaped illuminating aperture, normally used with an objective with a corresponding phase ring in its back focal plane for phase contrast.

6.2.2 Paraxial Bright Field - Spatially Coherent Illumination

In lack of a true bright field objective we also used the 40X HMC objective lens as an approximate bright field objective, but without the corresponding condenser filter. Instead we used a manually adjustable approximately circular 8-edged diaphragm in the front focal plane of the condenser to get maximally spatially coherent illumination by closing it to its minimal opening\(^3\). When the microscope is adjusted for Köhler illumination, this achieves an object illumination consisting of plane coherent waves, all propagating in approximately the same direction.

6.2.3 Ring shaped Illuminating Aperture

In order to do an experiment with other shapes of the illuminating aperture, we used the ring shaped illuminating aperture Ph1 intended for phase contrast.

6.2.4 Adjusting for Köhler Illumination

In order to adjust the microscope for Köhler illumination, we first focused on the “equator plane” of the egg, by adjusting the 40X Hoffman objective position, using the microscope focusing knob, until we found the observed image to be as “sharp” as possible.

Adjustment of “Köhler illumination” was only done with respect to focusing and centering of the “field diaphragm”\(^4\) in the field of view. We thus assumed that after such focusing and centering the image of the light source would then be centered and focused in the front focal plane.

\(^3\)When studying the first oocyte, we were somewhat puzzled to see that the manually adjustable circular condenser diaphragm we found above the condenser was not placed in the same plane as the HMC and phase contrast condenser optical filters in the condenser turret. It is because of this almost certain that this circular aperture was not placed in the front focal plane of the condenser. When studying the other oocytes we used another similar microscope, but with the adjustable circular diaphragm mounted in the condenser turret.

\(^4\)The diaphragm controlling which part of the object field is illuminated.
plane of the condenser. This is also the plane where the condenser diaphragm controlling the illuminating aperture should be placed. In order to have a fully open condenser aperture, we turned the condenser turret to the position with no mounted filter.

After this, we closed down the field diaphragm to its minimum, and adjusted the distance between the condenser and the in-focus object plane, by moving the condenser vertically (along the axis of the microscope) with all its filter mountings, until we saw a sharp image of the field diaphragm. Next we adjusted the condenser lens horizontally until the image of the field diaphragm was centered in the field of view. Having done this we finally opened the field diaphragm so the full field of view was illuminated. We did not check whether these adjustments in fact also resulted in the correct focusing of the image of the light source in the front focal plane of the condenser, as it should in Köhler illumination. With no other means of adjusting this than moving the condenser vertically, we assumed that the microscope was designed to have correct simultaneous focusing of the light source and the field diaphragm.

6.2.5 Observed Image Focus Sequences

By adjusting the vertical position of the objective lens, we focused at vertical positions from about $z = 100$ micrometers (above the “equator plane” of the egg) to about $z = -1000$ micrometers (below the egg), and acquired images at regular spaced positions. Since the egg has a diameter of about 120 micrometers, most of the images were focused at planes well below the egg. It was not possible to focus further above the egg, since the objective lens then hit the bottom of the petri dish.

The observed image sequences of the 4 oocytes using the 3 different contrast and illumination conditions can be found in appendix I.

6.3 Observations

In this section I discuss a number of observations, and analyze which conclusions can be drawn.

6.3.1 Violations of Rotational Symmetry

We have seen that the observed images with a small almost point like illuminating aperture, show an extreme degree of violations of 3D-rotational symmetry. When focusing on a plane as far as 1 mm below an embryo with a diameter of only about 0.1 mm, we are sure there is nothing but air, but we still see an image with spatial variations in intensity which are caused by the embryo. These effects of the embryo are only found below or on the backside of the embryo with respect to the illumination direction. They are not seen to the side of the embryo in its “equator plane”. Our understanding of this lack of rotational symmetry in the image formation - rotating the object results in a similar rotation of the image - goes as follows. Even though we are using an objective with a numerical aperture of about 0.5, we don’t see the classical defocusing or blurring of the in-focus image, as we would expect from a model of the image formation where each point on the object emits light in all directions with a position dependent intensity. In such a model, the defocusing could be approximated by a convolution with a circular disk of confusion with a radius proportional to the distance of the defocused image plane to the in-focus image plane. Using a 0.5 numerical aperture the constant of the proportionality would be 0.5. We don’t see such a blurring, since the image is not formed as a focused image of an object emitting

---

5We later learned that there is a special correcting lens between the field diaphragm and the condenser which is used together with the long working distance condenser lens. We still assume that the purpose of such a correcting lens is in fact to achieve simultaneous focusing of the field diaphragm in the object field and the light source in the front focal plane of the condenser.
light in all directions. The only thing emitting light is the light source of the microscope, and this is imaged/placed virtually at infinity and thus shines equally on all pixels in the image - if there were no object in between. But why do we then talk of a focused image? Because we can interpret the images we observe as focused images of the electromagnetic field in the plane we focus on. The violation of rotational symmetry in the image formation is actually a consequence of the lack of rotational symmetry of the illumination from the light source which is only a point at infinity and not the full hemisphere.

As a consequence of this extreme lack of rotational symmetry of the illumination, we get a consequent asymmetry in the images. The image of a structure or part of the object depends on the orientation relative to the illumination. The interface between inside and outside the cell is very clear at points on the interface where the light is close to tangential to the cell surface, and is not visible at all at points where the light is normal to the surface of the cell. This makes one speculate if there is any information at all about the parts of the cell surface with normals directed close to parallel to the light propagation direction.

6.3.2 The Lens Effect of a Cell

By the observation of the focused image of the illuminating aperture about 850 micrometers below the cell, we can say that yes, the central parts (with respect to the light propagation direction) of the cell surface do have an influence on the light - at least the second order structure of its geometry as measured by the local radius of curvature.

Based on this experiment, we can conclude that the central part of a spherical cell of an embryo may act like a lens that focuses the incoming light. Having adjusted the microscope for Köhler illumination we believe the light source of the microscope and the front focal plane of the condenser is imaged virtually by the condenser at infinity. Like any lens, the cells of an embryo focuses these apparently infinitely distant light sources to an image in the back focal plane of the cell. We have seen such a focused image of the light source shielded by differently shaped illuminating apertures, about $z = -850$ micrometers below what we believe to be an “equator plane” dividing the cell in halves. We believe this image is created by the central part of the cell acting as a lens, and conclude that the focal distance of this central part of the cell is about $f = 850$ micrometers. The “lens-makers formula” (see e.g. [18][p.23]) relates this focal length to the refractive index $n$ of the cell relative to its surroundings:

$$\frac{1}{f} = (n - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \quad (6.1)$$

where $R_1$ and $R_2$ are the signed radii of curvature of the entering and exiting surfaces. In the case of a sphere of radius $R$ we have $R_1 = R$ and $R_2 = -R$, and thus get:

$$\frac{1}{f} = (n - 1)(\frac{1}{R} + \frac{1}{R}) = \frac{2(n - 1)}{R} \quad (6.2)$$

$$f = \frac{R}{2(n - 1)} \quad (6.3)$$

$$n = 1 + \frac{R}{2f} \quad (6.4)$$

Using $R \approx 56 \mu m^6$ for the cell where we measured $f \approx 850 \mu m$ we get the refractive index of the cell relative to its surroundings to be:

$$n \approx 1 + \frac{56 \mu m}{2 \times 850 \mu m} \approx 1.033 \quad (6.5)$$

---

6Found by measuring the mean diameter of the oocyte to be about 405 pixels, and using a pixels pacing of 0.278 micrometers found from a calibration image of a micrometer scale, resulting in $R \approx 405/2 \times 0.278 \approx 56 \mu m$. 

53
Assuming that the surrounding aqueous medium has an absolute (that is, relative to vacuum) refractive index, which is similar to that of water, which is approximately \( \frac{4}{3} \) (see e.g. [49, p. 17]), the absolute refractive index of the oocyte can be calculated to be \( \frac{4}{3} \times 1.033 = 1.377 \). This is in good agreement with the general number of 1.37 for mammalian cells reported in [41].

6.3.3 The Beckeline and Non-Central Images of the Illuminating Aperture

We have also observed what we qualitatively interpret as a superposition of images of the illuminating aperture, focused by the non-central parts of the cell. We saw that these images were located in rings around the central image. The radius of these rings increased while the sizes of the images decreased as we observed them closer and closer to the equator plane of the cell, where they seemed to transform into the Beckeline with the same ring radius as the cell. We thus believe the Beckeline and these rings to be superpositions of images of the illuminating aperture, formed by the non-central parts of the cell, having the approximately same effective focal length. As we get further and further away from the central part the effective focal length of that part of the cell get shorter and shorter and consequently the images of illuminating aperture are focused closer and closer to the cell equator plane, and thus get smaller and smaller explaining the decrease in the width of the ring of images.

6.3.4 Towards Qualitative Reconstruction of Optical Thickness

In principle we should be able to interpret these local focal lengths of non-central parts of the cell as the result of a local second order (spatial) variation of the optical thickness experienced by the light.

However, this way of measuring the 3D-geometry of a cell is rather indirect and of a qualitative nature. Recognizing the 3D positions and sizes of images of the illuminating aperture does not seem like an easy task. And so we have been looking for alternative and more quantitative models of the image formation to base a reconstruction on. Especially, since the used HMC microscopy technique directly measures the first order variation of the optical thickness experienced by the light, it seemed more promising to base an inference of the second order variation of the optical thickness on HMC images rather than bright field images with a small illuminating aperture. After measuring the first order variation of the optical thickness, it seems like a detour to first infer the second order variations. In chapter 7 we shall investigate reconstruction of the optical thickness profile by regularized integration of a directional derivative of it as measured by the HMC technique.

6.4 Conclusions

- The order of magnitude of the refractive index of a human embryo is 1.033 relative to the surrounding aqueous nutrition medium.

- The consequent variations in optical delays of the light through an embryo with a diameter of \( 120\mu m \) relative to paths through the surrounding medium are from 0 to \( (1.033 - 1) \times 120\mu m \approx 4\mu m \).

- The maximal deflection angles occur for the light rays tangentially incident on an interface with a relative refractive index difference of 0.033. According to Snell’s law this corresponds to a small maximal deflection angle.

- We have not shown violations of a radiometric model of light energy propagation.
• Only when focusing close to points of tangentially incident light on the surface of the cell do we see images reminiscent of diffraction phenomena.

• We attribute this lack of observed interference and diffraction effects to the wide spectral composition of the light. We thus believe that it is essential to model the high degree of temporal or longitudinal incoherence.

• Using a single channel camera we have not been able to investigate the existence of color dependent effects in the images. We have however not observed any when inspecting directly through the eyepieces of the microscope.

• Because of the rather small maximal deflection angles, we don’t expect polarization effects to be needed to explain the images. We have however not investigated this quantitatively.

It seems that an anomalous diffraction model of the studied oocytes is a reasonable approximate model for most parts of the oocyte-light interaction, in combination with a geometric optics model of the image formation of large scale intensity variations.
Chapter 7

Reconstructing the Optical Thickness from Hoffman Modulation Contrast Images

This chapter describes a method for reconstructing the optical thickness profile from a HMC microscopy image of a transparent but anomalous refracting object. This is as close as we got at a "volumetric" 3D reconstruction. The method uses a simple model of the HMC image formation isolating the nonlinearities, and is applied to isolated human blastomers, in a study of their volumes. Previous and related works are path-length estimation in differential interference contrast images [38, 53] and general regularization techniques [48, 34].

7.1 Abstract

Hoffman microscopy imaging systems are part of numerous fertility clinics worldwide. We discuss the physics of a Hoffman imaging system from optical thickness to image intensity, implement a simple yet fast reconstruction algorithm using Fast Fourier Transformation, and discuss the usability of the method on a number of cells from a human embryo. Novelty is identifying the non-linearity of a typical Hoffman imaging system, and the application of Fourier Transformation to reconstruct the optical thickness.

7.2 Introduction

Hoffman Modulation Contrast (HMC) is a non-invasive light microscopy contrast technique, for the study of so-called phase objects, which only affects the phase of the transmitted light. Because of its non-invasive ability to generate contrast in images of otherwise transparent phase objects, it is widely used for in vivo studies of biological specimen. Specifically it is a standard technique used at many fertility clinics, for evaluation of embryo quality. The wish for quality assessment in fertility treatments, has led to the study of computational methods for 3D object reconstruction and quantitative morphological measurements, based on HMC images. Such methods may thus be of great clinical importance at the many fertility clinics around the world.

This chapter is organized as follows: In Sect. 7.3 the Hoffman Modulation Contrast technique is reviewed and the refined image formation model is presented. In Sect. 7.4 the implicit
approximations of light matter interaction underlying the Hoffman image formation model are discussed. In sect. 7.5 is presented a fast and regularized object reconstruction, and in sect. 7.6, the method is evaluated by reconstructing the optical thickness profiles of real isolated human blastomers.

### 7.3 Hoffman Modulation Contrast

The HMC technique generates contrast by converting “phase gradients” or “optical gradients” into intensity variations in the microscopy images[24] such that positive gradients show up bright and negative gradients show up dark. Figure 7.1(LEFT) shows an example HMC image. This results in images that give a human observer the familiar but in this situation false impression[24] of looking at the highlights and shadows resulting from a surface with height variations, shined upon from one side.

However it is also this coincidental familiarity that makes the HMC images preferred by humans for visual inspection and easy to (mis)interpret. As we shall see, for a certain regime of phase objects, a more accurate interpretation of a HMC image is as the directional derivative of the “optical thickness profile” of the object under study.

As described in [24] and [25], the HMC technique uses a rectangular slit illuminating aperture placed in the front focal plane of the condenser of the microscope, together with an aligned amplitude modulating filter with a dark, a grey and a bright region placed in the conjugate back-focal plane of the objective lens. In [23] this was refined using linear polarization filters giving control of the effective transmittance of a second slit placed next to the other slit. See Figure 7.1(RIGHT).

To reconstruct the optical thickness measured with Hoffman microscopes, the original model from [25, Eq. 10–11] has been adopted to the refined technique described in [23]. In this case, as illustrated by the graph in Figure 7.2, there are 5 intervals of local linearity of the function.
mapping deflection angles to image intensities. The function has 4 first order discontinuities at deflection angles $-2w/f, -w/f, 0, \text{ and } w/f$ corresponding to the 4 positions of the image of the slits in the HMC modulating filter shown in Figure 7.3. At the discontinuities the intensity may be found to be:

\[
\begin{align*}
I_{-2w/f} &= I_i \alpha T_D + I_i T_D, \\
I_{-w/f} &= I_i \alpha T_G + I_i T_D, \\
I_0 &= I_i \alpha T_B + I_i T_G, \\
I_{+w/f} &= I_i \alpha T_B + I_i T_B,
\end{align*}
\]  

where $f$ is the focal length of the microscope, $w$ is the width of the slits, $I_i$ is the incident intensity for each of the two slits, $\alpha$ is the transmittance of the semi-transparent slit, and $T_B, T_G$ and $T_D$ are the transmittances of respectively the bright, grey and dark region of the HMC modulating filter. Linear interpolation between these gives:

\[
I(S_\theta) - I_0 = \begin{cases} 
I_i(...); & S_\theta \leq -2\frac{w}{f} \\
I_i((T_G - T_D)(\alpha(\frac{w}{T_B}S_\theta + 1) - 1) + \alpha(-T_B + T_G)); & -2\frac{w}{f} \leq S_\theta \leq -\frac{w}{f} \\
I_i(\alpha(T_B - T_G) + (T_G - T_D))\frac{w}{T_B}S_\theta; & -\frac{w}{f} \leq S_\theta \leq 0 \\
I_i(T_B - T_G)\frac{w}{T_B}S_\theta; & 0 \leq S_\theta \leq \frac{w}{f} \\
I_i(...); & \frac{w}{f} \leq S_\theta
\end{cases}
\]

Note that $S_\theta = |\nabla S| \cos \beta = \cos \theta \frac{\partial}{\partial x}S + \sin \theta \frac{\partial}{\partial y}S$ is the directional derivative of the optical thickness in the slit width direction. Here the function $|\nabla S(x,y)|$ denotes the magnitude of the gradient of the optical thickness $S(x,y)$, and $\beta$ is the angle between the gradient of the optical thickness and the slit width direction, $\theta$. This model has apparently been refined for “negative” gradients.
7.4 HMC Model Approximations

In the following, the model approximations used in the imaging model will be made explicit:

**Transparent Refracting Object Model:** The object under study is assumed to be completely transparent and with spatially varying light speed \(c(x, y, z)\). The light-matter interaction is thus of purely refractive character and may be modeled by the non-complex refractive index \(n(x, y, z) = c_e/c\), where \(c_e\) is the light speed of the surrounding medium. For the majority of living cells and tissues, this is a fairly good approximation[12, p. 3].

**Weak Phase Object Approximation:** The refraction experienced by a ray of light is assumed to be well approximated by anomalous diffraction [49, p. 115], which is a good approximation **even** for objects having large finite extend, \(z_0\) in the sense \(2\pi z_0 \gg 1\), as long as they are soft, \(|n - 1| \ll 1\), and have \(2\pi(n - 1)\) approximately constant throughout the object. In this model the object is represented by the finite phase delay function, \(\phi(x, y) = \int_0^{z_0} \frac{2\pi}{\lambda}(n(x, y, z) - 1) \, dz\), where \(\lambda\) is the wavelength of light in the surrounding medium. The optical path length difference or “optical thickness”, \(S(x, y) = \frac{\lambda}{2\pi}\phi(x, y) = \int_0^{z_0} (n(x, y, z) - 1) \, dz\), has the advantage over the phase delay that, it is independent of the wavelength or frequency of the light, and thus only depends on the object.

**Gaussian Optics:** The approximate model of paraxial rays[11, p. 207] known as Gaussian Optics[11, p. 566] is adopted. Taking the z-axis along the optical axis of main ray propagation, a ray direction vector \(\vec{d} = (dx, dy, dz)\) is then expressed in polar coordinates as \(\vec{d} = (\cos(\eta) \sin(\xi), \sin(\eta) \sin(\xi), \cos(\xi))\). For paraxial rays the angle \(\xi\) is then assumed small enough such that \(\sin(\xi) \approx \xi\), and thus \(\vec{d} \approx (\cos(\eta)\xi, \sin(\eta)\xi, 1)\).

**Small Deflection Angles:** Suppose a paraxial ray with direction vector \(\vec{d} = (dx, dy, dz)\) is incident on an anomalous diffracting object with constant “optical gradient” \(\nabla S(x, y) = (k_1, k_2)\). The optical thickness profile is then given as \(S(x, y) = k_1 x + k_2 y + S_0\). The corresponding phase delay profile for a monochromatic paraxial ray with wavelength \(\lambda\) in the surrounding medium is then \(\phi(x, y) = 2\pi/\lambda S(x, y) = 2\pi/\lambda(k_1 x + k_2 y + S_0)\). If the phase function of the incoming wave is written: \(\phi_i(x, y, z) = 2\pi/\lambda(dx x + dy y + dz z)\) we see that the refracted wave in the \((x, y)\)-plane \((z = 0)\) will have the phase function \(\phi_r(x, y) = \phi_i(x, y, 0) + \phi(x, y) = (dx + k_1)x + (dy + k_2)y + S_0\), which we recognize as the phase function of a plane wave with direction cosines \((dx + k_1, dy + k_2)\).

**Geometrical Optics Image Formation:** Diffraction phenomena arising from the finite aperture of the image forming objective lens is not modeled. This is in contrast to the extensive theory of the resolution limits of image formation in diffraction limited systems. However, the objects and the resolution at which they are studied are assumed to be sufficiently larger than the wavelength of light. Human embryos have a diameter of approximately \(120\mu m = 240\lambda\).

7.5 Automated, Fast and Regularized Reconstruction

The HMC derivative direction may be estimated as the dominating direction of structure tensor [52] at global scale. In the present algorithm a variant of the structure tensor is used [54], where the direction of light is estimated as the largest eigenvector of the \(2 \times 2\) matrix \(T_{ij}(x, y, \sigma) = \int_\Omega P(x, y, \sigma)(S_{\theta, x, \sigma} S_{\theta, y, \sigma}) \, dx \, dy\) with \(P(x, y, \sigma) = k(S_{\theta, x, \sigma}^2 + S_{\theta, y, \sigma}^2)\), where \(*\) is the convolution operator, \(k\) is a normalizing constant such that \(P(x, y, \sigma)\) integrates to unity, \(S_{\theta, x, \sigma} = \frac{\partial}{\partial x} G_{\sigma} * S_{\theta}\), \(S_{\theta, y, \sigma} = \frac{\partial}{\partial y} G_{\sigma} * S_{\theta}\), and \(G_{\sigma}(x, y) = (2\pi\sigma^2)^{-1} \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right)\).
Given the HMC derivative direction, \( \theta \), the optical thickness in a noiseless image may be reconstructed using,

\[
S(x, y) = \int_{-\infty}^{t_x, y, \theta(x, y)} S_\theta(t_x, y, \theta(t)) \, dt,
\]

where \( l_x, y, \theta(t) \) is a line parameterized by \( t \) passing through the point \( (x, y) \) at an angle \( \theta \) w.r.t. the image coordinate system such that \( l_x, y, \theta(0) = (x, y) \), and \( S_\theta \) is the measured image. Unfortunately, noise in \( S_\theta \) will have profound influence on the numerical stability of the reconstruction. To remedy this, the integration is regularized by seeking \( L \) such that reconstructions where neighboring lines fluctuate violently are penalized. This is achieved by minimizing \( \int_\Omega F(S_\theta(x, y), L_x(x, y), L_y(x, y)) \, dx \, dy \), where \( L_x \) and \( L_y \) are the derivatives of \( L \) w.r.t the coordinate axis, \( \Omega \) the image domain,

\[
F(S_\theta, L_x, L_y) = \tau(L_x^2 + L_y^2) + (S_\theta - \cos(\theta)L_x - \sin(\theta)L_y)^2,
\]  

(7.6)

and \( \tau \) is the regularization parameter. Arguments for \( S_\theta \), \( L_x \), and \( L_y \) have been omitted for brevity. Using Euler-Lagrange it is found that the minimum must satisfy, \( 0 = \partial_x F L_x + \partial_y F L_y \). Solving this equation in the Fourier Domain leads to the following solution,

\[
\hat{L} = \frac{-2 \cos(\theta) iu - 2 \sin(\theta) iv}{2 \tau(u^2 + v^2) + 2 \cos^2(\theta) u^2 + 2 \sin^2(\theta) v^2 + 4 \cos(\theta) \sin(\theta) uv} \hat{S}_\theta,
\]

where \( \hat{L} \) and \( \hat{S}_\theta \) are the Fourier Transform of the \( L \) and \( S_\theta \), and \((u, v)\) are the frequency directions corresponding to \( (x, y) \).

### 7.6 Evaluation and Conclusion

We have reconstructed the optical thickness profile of 8 isolated human blastomeres using the method described above, and compared the results with the apparent geometrical volume, estimated from the 2D contour and assumption of spherical geometry.

To reconstruct the optical thickness, the measured intensity values by the used camera were assumed to be proportional to the light intensities \( I \) as long as they were within the dynamic range of the camera.

Based on simple measurements and calibration experiments we used the constants \( T_d = 0 \), \( T_g = 0.5 \), \( T_b = 1 \), \( w/f = 0.19 \), and \( T_d = 0.64 \) in the conversion of image intensities to deflection angles, by inversion of equation 7.5. Given an image of an object, where the uniform non-refracting medium is visible, the background intensity, \( I_0 \), is readily acquired. Further, experimenting with different values of the regularization parameter, \( \tau \), lead us to set it as low as possible without resulting in images visibly disturbed by noise. We found that increasing \( \tau \) after a certain threshold started to have a systematic decrease in optical thickness. We concluded by setting \( \tau = 0.05 \). An example of a reconstruction is given in Figure 7.4. The Fourier implementation assumes periodic boundary condition (torus structure), the result of which is visible in the reconstruction in the figure. To amend, a spatial finite difference scheme was implemented with fixed, zero boundaries, but otherwise identical optimization functional, in order to refine the initial estimate of the Fourier Method.

Using the reconstructed thickness profiles, we may estimate the diameter \( (D_z) \) of the blastomere along the z-axis from the central optical thickness \( S_c \) and the average relative refractive index, from: \( S_c = D_z(n/n_s - 1) \). In order to measure the “optical vertical diameter” we need to know \( n/n_s - 1 \). Unfortunately, it has not been possible to find these data in the literature for human blastomeres, except from general comments on the refractive index of biological and human cells being close to that of water and dependent on the protein concentration. Based on our
Figure 7.4: Reconstruction of the optical thickness of two human blastomers. LEFT: Image intensities converted to deflection angles or “optical gradients”. RIGHT: Reconstructed optical thickness profile, in approximate units of micrometers.

<table>
<thead>
<tr>
<th>Blastomer</th>
<th>Dc (µm)</th>
<th>Smax (µm)</th>
<th>Dz (µm)</th>
<th>1/2π(Dc/2)²/n³ (µm)³</th>
<th>VO/(n/ns - 1) 10³(µm)³</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>3.1</td>
<td>82</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>2.5</td>
<td>66</td>
<td>72</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>2.8</td>
<td>75</td>
<td>64</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>49</td>
<td>2.8</td>
<td>75</td>
<td>63</td>
<td>38</td>
</tr>
<tr>
<td>5</td>
<td>57</td>
<td>3.1</td>
<td>82</td>
<td>100</td>
<td>58</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>2.5</td>
<td>65</td>
<td>64</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>51</td>
<td>2.3</td>
<td>61</td>
<td>68</td>
<td>29</td>
</tr>
<tr>
<td>8</td>
<td>51</td>
<td>1.9</td>
<td>52</td>
<td>69</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 7.1: Blastomer data.

Analysis of other measurements, we have come to believe that human blastomers have an average relative refractive index to the surrounding water of approximately $n/ns = 1.38/1.33 = 1.04$. Using this value, we have calculated the diameters $D_z$ along the z-axis of 8 human blastomers, and compared them with a mean estimate $D_c$ from the 2D contour (see table 7.1). It is also possible to estimate the “optical volume” of the object by integrating the optical thickness over the support of the object image:

$$VO = \int_{\text{Object image support}} S(x, y)dA = \int_V (n/ns - 1)dV$$

(7.7)

where $V$ is the cylinder volume described by the object contour along the z-axis. These values have been computed for the 8 blastomers, and are presented in table 7.1 scaled with $n/ns - 1$ for ease of comparison with the geometrical volume estimate. The geometrically and optically estimated volumes differ up to a factor 2. We attribute this deviation in the number of relatively poorly calibrated parameters used in the reconstruction, however the spherical model applied in the geometrical estimate has not been validated, and certainly by just the 2D intersections it is apparent that they are not completely spherical. Other factors that should be considered when comparing these methods are differences in protein concentration, and the possibility that gravity might be controlling the orientation of non-spherical blastomers.

The presented algorithm is capable of retrieving 3D information otherwise lost in studies of 2D contours on Hoffman microscopy images. Future work will be spend on improving our models.
and methods to increase the accuracy of 3D information from the 2D Hoffman images especially in studying blastomers in embryos. Being non-invasive and part of numerous clinical fertility treatments, improved understanding of HMC images is expected to have an increasing effect on the success-rate of implantations.
Part III

Segmenting Embryos in HMC-Images
Chapter 8

3D Blastomer Detection in HMC-Images of Human Embryos

Since we have not been able to find a model of the image formation, which allows for reconstruction of a 3D volumetric model, we have been looking for a less ambitious embryo specific object model, which we can infer directly from the images. Such “reconstruction” is based on recognition of characteristic patterns in the images, which from experience can be related to the object geometry. However, transforming experience into computational models and methods is not an easy task. In order to implement such an image based method we still need a computational model of the image formation. In terms of the multilevel model of the image formation (chapter 3, Figure 3.2 page 28), the model of the image formation can be considered starting with $L_7 = O_{contours}$, and going all the way to the images $I$. Inspired by the biologists manual analysis of the embryo focus image sequences by annotation of the equator contours of the observed blastomers, we have developed a program called “select” for automatic detection of spherical blastomers and user interactive inspection and correction of the automatic detections. This program is described in the next chapter.

The purpose of this chapter is to present our theoretical understanding of the models and techniques used in “select” for the automatic blastomer detection in HMC-image focus sequences of human embryos.

8.1 3D Detection of Spherical Deviations in Refractive Index

We present a simple algorithm for detecting 3D spherical deviations in the refractive index of an object based on 3D image analysis of an optically focused sequence of Hoffman Modulation Contrast (HMC) light microscopy images. The algorithm is based on fast global search of maximum correlation with a theoretical model image. Since the search is global, it may be used for initialization of local optimization algorithms using a deformable model.

8.1.1 Blastomer Detection in Human Embryos

One application of this detection algorithm is detection of spherical blastomers in human embryos. In Figure 8.1 is shown an example of a HMC-image of an embryo. This image may be compared to the 2D embryo-model sketched in Figure 8.1, which identifies important morphological parts of a typical embryo on the second day after successful fertilization.
Figure 8.1: Left: Example of HMC-image of an embryo. Right: Morphology model of embryo at day 2 after fertilization.

### 8.2 Assumed Model of Image Formation

To derive our model of the image formation, we start from the same approximate image formation model assumptions:

- Gaussian Optics
- Transparent Refracting Object Model
- Weak Phase Object Approximation
- Small Deflection Angles
- Geometrical Optics Image Formation

discussed in [37] and here presented in chapter 7 p. 60.

In addition to these, we further assume that the studied object can be modelled as a collection of spheres each with homogeneous internal refractive index. We first consider how such a single sphere is imaged under Hoffman Modulation Contrast in order to arrive at a model image.

#### 8.2.1 Hoffman Modulation Contrast

The HMC-technique, as illustrated in Figure 8.2 is particularly useful for studying thin phase objects in the regime of anomalous diffraction. In this model of the light matter interaction, a phase object is modelled by its optical thickness profile. In the approximate case of only extreme paraxial incoming rays, all rays will “feel” the same optical profile $S$ along the $z$-axis (taken to be the optical axis). In terms of the spatially varying refractive index relative to the surrounding medium:

$$n(x, y, z) = c_{\text{surrounding}} / c(x, y, z),$$  \hspace{1cm} (8.1)

the optical thickness $S(x, y)$ relative to the surroundings and the corresponding phase shift $\phi(x, y)$ experienced by a light ray parallel to the $z$-axis and passing through the point $(x, y)$ in the object plane ($z = 0$) are given by

$$S(x, y) = \int (n(x, y, z) - 1) dz,$$  \hspace{1cm} (8.2)

and

$$\phi(x, y) = \frac{2\pi}{\lambda} S(x, y).$$  \hspace{1cm} (8.3)
According to the theory of Hoffman Modulation Contrast [23], [37] the image intensity \( I \) is a function of the directional derivative of the optical thickness profile of the object. For simplicity we here assume that the HMC contrast level has been chosen to give intensity deviations \((I - I_0)\) from the background \((I_0)\) with a linear dependence on the directional derivative\(^1\):

\[
I - I_0 = kS_\theta.
\]  

(8.4)

Here \( \theta \) is the angle of the HMC derivative direction relative to the \( x \)-axis, and

\[
S_\theta = D_\theta S = \cos(\theta) \frac{\partial}{\partial x} S + \sin(\theta) \frac{\partial}{\partial y} S 
\]  

(8.5)

is the derivative of the optical thickness in the HMC derivative direction (the slit width direction).

### 8.2.2 HMC Image of Homogeneous Sphere

Based on this model of the image formation, we can predict an in-focus model-image \( I_M \) of a homogeneous spherical deviation in the refractive index. The optical thickness profile \( S_M = (x^c=0, y^c=0, R, n) \) of an origo centered, homogeneous sphere with radius \( R \) and refractive index \( n \) is given by:

\[
S(x, y) = \begin{cases} 
2(n - 1) \sqrt{R^2 - (x^2 + y^2)} & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{if } x^2 + y^2 > R^2 
\end{cases}
\]  

(8.6)

\(^1\)This assumption is not essential for the following derivation of a model image. One could use a general intensity transformation \((I = I(S_\theta))\) to arrive at a model image in the general case. However, when inferring the model parameters, and in particular the internal deviation of the relative refractive index \((n - 1)\), the search for the MAP estimate of \(n - 1\) can be substituted by a direct calculation when the dependence is linear. Furthermore, when extending the model to more than one spherical object, we can argue for approximate linearity of the image formation, when the intensity deviations depend linearly on the directional derivative of the optical thickness.
This is illustrated in Figure 8.3 (LEFT). Please note that the optical thickness profile of a sphere is not differentiable on the contour \((x^2 + y^2 = R^2)\). The coordinates \(S_x = \frac{\partial S(x,y)}{\partial x}\) and \(S_y = \frac{\partial S(x,y)}{\partial y}\) of the gradient tends to (plus or minus) infinity as one approaches the contour from inside:

\[
S_x(x, y) = \begin{cases} 
-2(n-1)x(R^2 - (x^2 + y^2))^{-\frac{1}{2}} & \text{if } x^2 + y^2 < R^2 \\
0 & \text{if } x^2 + y^2 > R^2 
\end{cases}
\]

\[
S_y(x, y) = \begin{cases} 
-2(n-1)y(R^2 - (x^2 + y^2))^{-\frac{1}{2}} & \text{if } x^2 + y^2 < R^2 \\
0 & \text{if } x^2 + y^2 > R^2 
\end{cases}
\]

and of course the derivative outside the support of the object is zero. To cope with this, we have chosen to smooth the optical thickness profile slightly before calculating the derivative. One can actually think of this smoothing as a simple way of trying to model the diffraction effects on a light ray. The absolute differences in optical thickness for neighboring rays in a sufficiently thin bundle are arbitrarily small. We can think of averaging the amplitudes of the photons from each of these rays as an approximate way of modelling the diffraction of light. Since the phase differences of these photons are sufficiently small and they have approximately the same intensity, averaging the amplitudes would result in averaging the phases. These phases are just \(2\pi/\lambda\) times the optical thicknesses (“modulo \(2\pi\)”). So we can think of this smoothing as an approximation of diffraction. Surely diffraction effects would not allow the optical thickness profile to vary faster than 1 (length unit per length unit), (or the phases \(2\pi\) per wavelength). We thus calculate the derivative of the optical thickness profile at a scale \(\sigma_{\text{diff}} \approx \frac{\lambda}{2\pi} \approx 0.1\mu m\) of the same order of magnitude as the wavelength of light \(\lambda \approx 0.5\mu m\), and use the result as our model predicted image of a single spherical blastomer (see Figure 8.3 (Right)):

\[
I_M(x, y) - I_0 = I_{(x_c, y_c, R, n)}(x, y) - I_0 = kD_0(G_{\sigma_{\text{diff}}} * S)(x - x^c, y - y^c).
\]

Here * is the convolution operator,

\[
(G_{\sigma_{\text{diff}}} * S)(x, y) = \int \int G_{\sigma_{\text{diff}}}(x - x', y - y')S(x', y')dx'dy',
\]

and \(G_{\sigma_{\text{diff}}}\) is the smoothing kernel. We use a Gaussian\(^2\) smoothing kernel:

\[
G_{\sigma_{\text{diff}}}(x, y) = (2\pi\sigma_{\text{diff}}^2)^{-1}e^{-(x^2+y^2)/(2\sigma_{\text{diff}}^2)}.
\]

\(^2\)It would be more honest to use a kernel that is related to the diffraction phenomenon. To arrive at such a kernel, one could analyze for example the linear filtering model of coherent wave propagation in the Fourier Optics formulation of scalar diffraction theory. See e.g. J.W. Goodman “Introduction to Fourier Optics” p. 50-51 [21].
To simplify later notation, we introduce the smoothed optical thickness profile $T_{R,s_{\text{dir}}}$, and its directional derivative $T_{R,s_{\text{dir}},\theta}$, of a sphere with radius $R$ and unit deviation in refractive index ($n-1=1$):

\begin{align}
T_{R,s_{\text{dir}}} &= \frac{1}{n-1} G_{s_{\text{dir}}} * S \\
T_{R,s_{\text{dir}},\theta} &= D_\theta T_{R,s_{\text{dir}}}.
\end{align}

(8.12) (8.13)

Using this notation the image model can be written:

$$I_M(x,y) - I_0 = (n-1) k T_{R,s_{\text{dir}},\theta}(x-x^c, y-y^c).$$

(8.14)

### 8.2.3 Noise and Background Model

The deterministic model image $I_M(x,y)$ is now extended to a probabilistic model of the image formation, to handle pixel noise and a varying background partly due to other out-of-focus structures like other blastomers. Following [5], we do this by assuming that the observed deviations ($I(x,y) - I_M(x,y)$) from a true model image, can be modelled as a sum of independent Gaussianly distributed pixel noise with spread $\sigma_{\text{noise}}$ and a random background image drawn from a multivariate Gaussian distribution on $I$ with mean $\overline{b} \in I$ and covariance structure given by a symmetric and positive definite operator $C_b : I \to I$ (in case of pixel-ized images with a finite number of pixels, $C_b$ is simply the variance-covariance matrix between all the pixels in the images):

$$P(I|\{x^c, y^c\}, R, n)) \propto \exp\left( -\frac{1}{2} <I - I_{\{x^c, y^c\},R,n} - \overline{b}, (C_b + \sigma_{\text{noise}}^2)^{-1} (I - I_{\{x^c, y^c\},R,n}) - \overline{b}> \right).$$

(8.15)

Here the associated energy,

$$E(I|M = \{x^c, y^c\}, R, n)) = \frac{1}{2} ||I - (I_M + \overline{b})||^2_{(C_b + \sigma_{\text{noise}}^2)^{-1}} = \frac{1}{2} <I - I_M - \overline{b}, (C_b + \sigma_{\text{noise}}^2)^{-1} (I - I_M - \overline{b})>,$$

(8.16) (8.17)

is half the squared distance between the observed image $I$ and the model image $I_M + \overline{b}$ as measured by the norm arising from $C_b + \sigma_{\text{noise}}^2$ and the standard inner product $\langle \cdot, \cdot \rangle$ in the vector space of complex-valued images:

$$<I_1, I_2> = \int \int \overline{I_1(x,y)} I_2(x,y) dx dy.$$

(8.18)

Corresponding to an assumption of translational invariance of the random background distribution, we shall assume that $\overline{b}$ is constant and $C_b$ is diagonal in the Fourier domain ($C_b e^{i(k_x x + k_y y)} = c_k e^{i(k_x x + k_y y)}$) and thus simply a convolution operator $C_b I = K * I$ where $K \in I$ is some filter kernel. We shall also assume isotropy of $K$ corresponding to an assumption of rotational invariance of the random background. So $c_k$ can only depend on the magnitude of $k = (k_x, k_y)$: $c_k = e^{-\sqrt{k_x^2 + k_y^2}}$. Since in $I_M$ we already have a model parameter $I_0$ describing a constant background intensity we simply take $\overline{b} = 0$. Instead of calculating $(C_b + \sigma_{\text{noise}}^2)^{-1}$ from a model of the background $C_b$, we have used a simple high pass filter of the form $\sigma_{\text{noise}}^2 (1 - G_{\sigma_{bg}})$, where $1$ is the identity operator and $G_{\sigma_{bg}}$ is the convolution operator with a Gaussian kernel with spatial spread $\sigma_{bg}$. Here the parameter $\sigma_{bg}$ expresses the lower “limit” of spatial scales of image structures that should be regarded as part of the background. We used $\sigma_{bg} = 10\mu m$, when detecting blastomers in HMC-images of human embryos.
This assumption is mainly motivated as a computational simple approximate model of deviations from the deterministic model image $I_M$. It can be seen as a simple approximation to “slice reconstruction”, in the sense that out-of-focus structures are likely to be in the low-frequency domain while the high-frequency structures are believed to be “in-focus”. Of course, such a simple frequency filtering will wrongly delete the low frequency parts of in-focus structures and will not eliminate out-of-focus structures with a high spatial frequency.

8.2.4 The Final Image Formation Model - of a Single Blastomer

We conclude this section by repeating the full model of the image formation of a single blastomer:

$$P(I|M) \propto \exp(-E(I|M))$$  \hspace{1cm} (8.19)

$$E(I|M) = \frac{1}{2} ||I - I_M||_2^2$$  \hspace{1cm} (8.20)

$$I_{x^c,y^c,R,n,I_0}(x,y) = (n-1)kT_{R,\sigma_{\text{filt}},\theta}(x-x^c,y-y^c) + I_0$$  \hspace{1cm} (8.21)

$$T_{R,\sigma_{\text{filt}},\theta} = D_\theta T_{R,\sigma_{\text{filt}}}$$  \hspace{1cm} (8.22)

$$T_R(x,y) = \begin{cases} 2\sqrt{R^2 - (x^2 + y^2)} & \text{if } x^2 + y^2 \leq R^2 \\ 0 & \text{if } x^2 + y^2 > R^2 \end{cases}$$  \hspace{1cm} (8.23)

$$G_{\sigma} * I(x,y) = \int \int G_{\sigma}(x',y')I(x',y')dx'dy'$$  \hspace{1cm} (8.24)

$$G_{\sigma}(x,y) = \left(2\pi\sigma^2\right)^{-1}e^{-\left(x^2+y^2\right)/(2\sigma^2)}$$  \hspace{1cm} (8.25)

$$D_\theta = \cos(\theta)\frac{\partial}{\partial x} + \sin(\theta)\frac{\partial}{\partial y}$$  \hspace{1cm} (8.26)

$$||I||^2_{\sigma_{\text{noise}}(1-G_{\sigma_{\text{bg}}})} = <I, (\sigma_{\text{noise}}(1-G_{\sigma_{\text{bg}}}))I >$$  \hspace{1cm} (8.27)

$$< I_1, I_2 > = \int \int T_1(x,y)I_2(x,y)dxdy.$$  \hspace{1cm} (8.28)

8.3 Single Blastomer Detection and Inference

We now turn to the tasks of detecting the presence of a single blastomer and inferring the model parameters $M = (x^c, y^c, R, n)$ describing its position, radius and internal refractive index relative to the surroundings.

Above we have described a model of the image formation $P(I|M)$. In order to do Bayesian inference we also need a prior model of the blastomer parameters $P(M)$. However, since we want the result of the inference to have as much character of a measurement as possible, we simply use independent uniform priors of the model parameters within suitable limits. In that case the maximum a posteriori estimate is simply the maximum likelihood estimate based on the above image formation model.

8.3.1 Detection

When performing the detection, we shall assume that the likelihood $P(I|M)$ is sufficiently peaked, that we can approximate the marginal probability $P(\text{there is 1 blastomer}|I) = \int dMP(M|I)$ by the likelihood $P(M_{\text{max}}|I)$ of the maximum likelihood estimate. As noted in [5] we can then base the detection decision solely on the likelihood ratio of the likelihood of the MAP estimate $M_{\text{MAP}} = \arg \max_M P(I|M) = \arg \min_M ||I - I_M||^2_{\sigma_{\text{noise}}(1-G_{\sigma_{\text{bg}}})}$, and the likelihood of the alterna-
tive model of “there is no blastomer”:

\[ P(\text{there is no blastomer} | I) \propto \exp\left(-\frac{1}{2} ||I||^{2}_{\text{noise}}(1-G_{\text{noise}})\right). \]  

(8.30)

Since this last likelihood is constant, the decision of whether there is a blastomer or not can in theory be taken simply by thresholding \( P(I|M) \) or equivalently by thresholding \( E(I|M) \).

### 8.3.2 Inference

When inferring the model parameters by MAP estimation, we simply minimize the energy function with respect to the model parameters:

\[ ||I(x, y) - (n - 1)kTR_{\sigma_{\text{diff}}, \theta}^{}(x - x^{c}, y - y^{c}) - I_{0}||^{2}_{\text{noise}}(1-G_{\text{noise}})\]. \]  

(8.31)

We consider the parameters \( k, \sigma_{\text{diff}}, \theta, \sigma_{\text{noise}}, \) and \( \sigma_{\text{bg}} \) as given constants describing our model of the image formation, not to be inferred from this image.

### 8.3.3 Analyzing the Energy Term

We could consider minimizing the energy by a combinatoric search for the minimizing parameters \( I_{0}, n - 1, (x^{c}, y^{c}), \) and \( R \). However it is worth recognizing the energy as a distance in an Euclidean vector space between the observed image \( I \) and a model image lying in one of a family of 2 dimensional subspaces spanned by the constant image \( I_{0} \) and a model deviation image \( kTR_{\sigma_{\text{diff}}, \theta}^{}(x - x^{c}, y - y^{c}) - I_{0}||^{2}_{\text{noise}}(1-G_{\text{noise}})\). \]  

(8.32)

The computational structure of the energy as a norm given by an inner product allows us to readily calculate the maximum likelihood estimate of the refractive index difference \( n - 1 \) of each sphere model and the background image intensity \( I_{0}^{3} \), by calculating the projection of the observed image \( I \) on the subspaces spanned by the model deviation image \( kTR_{\sigma_{\text{diff}}, \theta}^{}(x - x^{c}, y - y^{c}) \) and the constant image respectively, since they are mutually orthogonal. These projections are given by the inner products with model images normalized to unit norm.

When minimizing the energy with respect to the other model parameters \( (x^{c}, y^{c}) \) and \( R \) we thus need to maximize (the squared modulus of) inner products of the form:

\[ < I(x, y)|\frac{T_{R,\sigma_{\text{diff}}, \theta}^{}(x - x^{c}, y - y^{c})}{||T_{R,\sigma_{\text{diff}}, \theta}^{}(x - x^{c}, y - y^{c})||^{2}_{\text{noise}}(1-G_{\text{noise}})}^{2}_{\text{noise}}(1-G_{\text{noise}}) \]. \]  

(8.33)

Below I shall formally derive that the introduction of a general covariance structure for modelling the background, is equivalent to our use of a high pass filtering of correlation images.
prior to the detection of local maxima:

\[ < I_1, I_2 > = \int dx \overline{I_1(x)} I_2(x) \]  
\[ (K * I)(x) = \int dy K(x - y) I(y) \]  
\[ < I_1, I_2 > = (\Pi(I_1) * I_2)(0) \]  
\[ \Pi(I)(x) = I(-x) \]  
\[ \tau_z(I)(y) = I(y - x) \]  
\[ (K * I)(x) = \int dy \Pi(K)(y - x) I(y) \]  
\[ = \int dy \tau_z(\Pi(K))(y) I(y) \]  
\[ = < \tau_z(\Pi(K)), I > \]  
\[ < I_1, K * I_2 > = \Pi(I_1) * (K * I_2)(0) \]  
\[ K1 * K2 = K2 * K1 \]  
\[ (\tau_z(K) * I)(y) = \int dz \tau_z(K)(y - z) I(z) \]  
\[ = \int dz K(y - z - x) I(z - x) \]  
\[ = \int dz K(y - (z - x) - x) I(z - x) \]  
\[ = K * (\tau_z(I))(y) \]  
\[ \tau_z(K) * I = K * (\tau_z(I)) \]  
\[ < \tau_z(I_1), K * I_2 > = \Pi(\tau_z I_1) * (K * I_2)(0) \]  
\[ = \tau_z \Pi(I_1) * (K * I_2)(0) \]  
\[ = \Pi(I_1) * \tau_z(K * I_2)(0) \]  
\[ = \Pi(I_1) * (K * I_2)(x) \]  
\[ K1 * (K2 * I) = (K1 * K2) * I \]  
\[ = (K2 * K1) * I \]  
\[ = K2 * (K1 * I) \]  
\[ < \tau_z(I_1), K * I_2 > = K * (\Pi(I_1) * I_2)(x). \]  

8.4 Extending to 3D Image Focus Sequence Model

Up to now we have only considered how to analyze a single image for the presence of an in-focus sphere. Our model of the image formation only predicts in-focus images. Extending it to a model of the focusing effects is not straightforward. In the limiting case where all rays, both incoming and deflected, are all parallel with only infinitesimal mutual angles, there would be no focusing effects.

We do the simplest possible, and extend the 2D model of an in-focus HMC image of a sphere to a model of a full focus sequence of images, by ignoring them. This can be seen as an assumption of independence between the images and use of a uniform distribution model for the “out of focus” images. This model is for sure wrong, but implies that we should simply study one image at a time.

However, since we now expect the image \( I_{x, c} \) to also contain “out-of-focus” images superimposed on the in-focus image, we try to remove the out-of-focus structures by a high-pass filtering.
We have just tried a simple high-pass filtering by $1 - G_{\sigma^bg^*}$.

$$P((I_{z_1}, \ldots, I_{zn_k})|(x^c, y^c, z^c, n)) \propto P(I_x - I_M*G_{\sigma^bg^*} | (x^c, y^c, R, n)). \quad (8.57)$$

### 8.4.1 Automatic Cell Detection - What do we Detect?

All in all we detect the global extrema of the correlation function:

$$(x_0, y_0, z_0, R) \mapsto Corr(I_{z_0}, I_M)(x_0, y_0). \quad (8.58)$$

### 8.5 Multiple Blastomer Detection - Model Match Measure

Above we described how to detect a single blastomer, by finding the maximum correlation with normalized model predicted images. We used uniform priors $P(M)$ of blastomer sizes, positions and refractive indices $M = (x^c, y^c, z^c, R, n)$. In order to detect multiple blastomers, we need to extend both the prior model $P(M)$ and the model of the image formation $P(I|M)$ to the case where the morphology model $M = (M_1, \ldots, M_N)$ consists of more than 1 blastomer.

**Interactions in the morphology prior - No overlapping**

Modelling the physical interactions between the different parts of the embryo is not easy. We can however relatively easily implement a simple overlap criteria based on the center positions and radii of the blastomers.

### 8.5.1 Approximate Linear Image Formation

The image formation is extended by an assumption of linearity. This is based on the following considerations using the extreme model of anomalous diffraction.

We now consider a combined object described by $n(x, y, z)$ and consisting of two sub-objects each described by a spatial refractive index $n_1(x, y, z)$ and $n_2(x, y, z)$. If the two sub objects have non-overlapping support (in the sense $n_1^{-1}(R \setminus \{1\}) \cap n_2^{-1}(R \setminus \{1\}) = \emptyset$) the combined model of deviations in refractive index will simply be the sum of the two objects’ deviations in refractive index $(n - 1 = (n_1 - 1) + (n_2 - 1))$. Now since $S(x, y)$ depends linearly on $n(x, y, z) - 1$, $S$ will simply be the sum of the two optical profiles of each sub-object.

If either we are in the setting where the image intensity deviations from the background depend linearly on $S$, or we can further assume about the sub-objects that their optical thickness profiles $S_1$ and $S_2$ have directional derivatives $D_b(S_1)$ and $D_b(S_2)$ with non-overlapping support (in the sense $(D_b(S_1))^{-1}(R \setminus \{0\}) \cap (D_b(S_2))^{-1}(R \setminus \{0\}) = \emptyset$, then the image intensity deviations $I - I_0$ from the background will also be the sum of the two sub-objects’ observed image intensity deviations from the background:

$$I - I_0 = I_1 - I_0 + I_2 - I_0. \quad (8.59)$$

### 8.5.2 Ignoring Blastomer Image Interactions

When arguing for finding the MAP estimate of a single blastomer by maximizing the modulus of the correlation value, we used an assumption of orthogonality between the constant image and the model image of a single blastomer. When inferring the positions of multiple blastomers this assumption is no longer valid exactly. Blastomers sharing projected contours have nonzero inner products between their model images. We should thus in theory compensate for these linear dependencies between the different model images. However when making the high pass filtering the model images essentially have only very thin circular support. So as long as the models
are not coincidentally sharing contours in an osculating fashion, we can ignore these mutual linearities. With the morphology prior of non-overlapping blastomers we hope to avoid these osculating simultaneous blastomer models, and thus expect this ignorance of blastomer image interactions to be non-fatal.

8.5.3 Conclusion

In the next chapter we shall describe, how we detect the strongest local extrema of the correlation functions, and present them to the user as suggested cells in an interactive selection program, using a greedy implementation of the heuristic of not allowing too overlapping cells.
Chapter 9

Semiautomatic Sphere Model Selection

In this chapter we describe an interactive program “SELECT”. It was developed for interactive 3D inspection of automatic detections of spherical cells allowing for semiautomatic selection of 3D spheres making up a consistent 3D spheres model of the blastomers. This work was done in close collaboration with Jon Sporring, but also Mads Nielsen and Peter Berg Larsen have contributed.

9.1 Design Philosophy

Perfect automation of model inference is difficult to achieve. Because of this, it is important that the results can be both easily checked and corrected by the user. As a consequence, in order for a fully automatic method to be useful, it should be encapsulated by an interactive interface allowing for inspection and correction, turning it into a semi automatic tool as a whole. Depending on the extend and kind of needed user corrections a more manual tool may be preferred.

Since the goal is full automation with a minimal interactive shell around the automatic inference of sphere models, the “SELECT” program was first designed to serve the following two purposes:

1. As an interactive user interface to the automatic detection of spheres in recorded HMC image sequences.
2. As an interactive tool for correcting errors in the automatic sphere model detections.

To serve these purposes, “SELECT” was designed with the following primary functionalities:

1. User specification of the inputs for the detection algorithm
2. Presentation of the inferred spheres model for user inspection.
3. Interactive rejection and acceptance of suggested spheres.
4. Inspection and selection of sphere detections.

9.2 Multiple Blastomers

Extending the model of a single blastomer to the case of multiple blastomers is not straight forward. The standard way of extending the prior model $P(M)$ of the size, position and refractive index of a single spherical blastomer to a model of $N$ blastomers by construction of the product distribution $P(M_1, \ldots, M_N) = \prod_{i=1}^{N} P(M_N)$ has an implicit assumption of independence which is very much violated by the physical interactions between the blastomers.
9.2.1 Blastomer Interactions - No Shape Interaction

The physical interactions between neighboring blastomers, and the consequent relations between their positions, sizes and shapes are difficult to model. Some minimal amount of modeling of these effects is however needed. Otherwise it would be entirely up to the image formation model to rule out multiple detections of the most visible blastomer. And so the modeling of their interactions in the image formation would be needed. In the model of spherical blastomer shapes it is only their positions and sizes that can interact.

Since the physical interactions is closely related to their geometry, implementing these interactions in the prior of their geometry, is likely to be easier, and hopefully it can completely eliminate the need1 for interactions in the image formation model.

I thus cannot simply again approximate the simultaneous prior distributions of blastomers $P(M|M.N)$ by the product distributions of the marginal distributions of each blastomer. Since my main concern for the moment is only to eliminate spatially conflicting blastomer models which correspond to overlapping blastomers, I will try with a simple model of this geometrical interaction based on the blastomer sizes $M.Sz = (M_1.Sz, \ldots, M_{M.N}.Sz) \in \mathbb{R}^{M.N}_+$ and 3D center positions $M.P = (M_1.P, \ldots, M_{M.N}.P) \in (\mathbb{R}^3)^{M.N}$, and as an interaction solely between these. I will consequently ignore any interaction between the different blastomer shapes and orientations2, and approximate:

$$P(M|M.N) \approx P(M.Sz, M.P|M.N) \prod_{n=1,\ldots,M.N} P(M_n.Shape|M.N).$$

(9.1)

With these assumptions and approximations of the prior model of blastomers, the prior consists of the following two parts:

- A prior on the number of blastomers $P(M.N)$.
- For each $M.N \in \mathbb{R}$, a prior on the simultaneous distribution of blastomer sizes and positions $P(M.Sz, M.P|M.N)$.

9.2.2 Prior of Blastomer Sizes and Positions

The model $P(M.Sz, M.P|M.N) = P(S, P|N)$ must fulfill some symmetries. Because all blastomers are considered equal, and numbered arbitrarily, it should be symmetric with respect to reorderings of the blastomers. Furthermore, it is likely that there will be approximate invariance with respect to the absolute position and orientation since these are measured relatively to the arbitrary choice of camera coordinate system, which is assumed to be independent of the detailed object position and orientation. This assumption of invariance is violated by the fact that the object was placed in the field of view of the camera, and the microscope was initially focused on the object. Still the absolute position and orientation as measured by the mean position

$$P = \frac{1}{N} \sum_{n=1}^{N} P_n$$

(9.2)

1I believe that there are also visible effects in the images from blastomer interactions in the image formation, that cannot be modeled in the prior. However, for the application of detecting near spherical blastomers, I hope these effects can be ignored. I am thinking of the effects when two blastomers touch and the usually occurring refractions more or less cancel out.

2This is in obvious contradiction to the clearly visible effects on the blastomer shapes and their mutual alignment from blastomers interacting by touching and pressing against each other. I believe that blastomers would be spherical if left alone, and any interaction in the form of external forces e.g. from other blastomers would result in deviations from the spherical shape. The motivation here is thus only to make a simplifying model approximation.
is likely to have larger variations which are independent of the relative positions of the blastomers. I thus approximate by:

$$P(S, P|N) = P(S, P|N, \overline{P})P(\overline{P}|N),$$

(9.3)

where $P(\overline{P}|N)$ allows for modelling of the object positioning and focusing, and to obtain translational invariance of $P(S, P|N, \overline{P})$, I will use:

$$P(S, P|N, \overline{P}) = P(S, P - \overline{P}|N).$$

(9.4)

Since I am here only interested in a prior that can be used for inference of among other things the statistics of simultaneous blastomer positions and sizes, it is preferable to use a least committed prior. I will thus not continue the analysis of the structure of $P(S, P - \overline{P}|N)$, such that it both obeys the symmetry constraints and is a good approximation of the true distribution. I will rather concentrate on how to implement the blastomer interactions avoiding non-physical overlapping detections. A simple way of doing this is simply to require that all distances between any two blastomers be larger than some fraction of the sum of their linear sizes like e.g. their mean radius.

9.2.3 Interactive Inspection

To allow for interactive inspection and correction of the sphere detections, we have developed a program with user interface as illustrated below:

9.3 Implementation Considerations

9.3.1 Automatic Detection of the HMC Analysis Direction

Since we do not have direct measurements of the HMC analysis direction, we have developed a procedure which estimates this from the equator image of an egg. This can easily be computed off-line and stored in the sequence info file.

As also described in [37], the HMC derivative direction may be estimated as the dominating direction of structure tensor [52] at global scale. In the present algorithm a variant of the structure
tensor is used \[54\], where the direction of light is estimated as the largest eigenvector of the \(2 \times 2\) matrix
\[T_{ij}(x, y, \sigma) = \int_{\Omega} P(x, y, \sigma)(S_{\theta, x, \sigma} S_{\theta, y, \sigma}) \, dx \, dy\]
with \(P(x, y, \sigma) = k(S_{\theta, x, \sigma}^2 + S_{\theta, y, \sigma}^2)\), where \(\ast\) is the convolution operator, \(k\) is a normalizing constant such that \(P(x, y, \sigma)\) integrates to unity, \(S_{\theta, x, \sigma} = \frac{\partial}{\partial x} G_x \ast S_\theta\), \(S_{\theta, y, \sigma} = \frac{\partial}{\partial y} G_y \ast S_\theta\), and \(G_\sigma(x, y) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)\).

Unless we know beforehand the sign of the deviation in refractive index, we can only find the analysis direction up to a 180 degrees change of direction.

### 9.3.2 Automatic Registration of the Images in the Sequence

Unfortunately, the embryos are not stationary during image acquisition. They move slightly from image to image because of the finite time between acquisition of the different images.

Since we do not have direct measurements of how the egg moved during the acquisition of the image sequence, we have developed a registration algorithm, which estimates the translations of the egg between the images. This could easily be calculated off-line and stored in the sequence information file. Even if it were not computed off-line it would still be necessary to somehow communicate the registration used when making a 3D-model. Deciding how and where to communicate this information, should be done after analysing different strategies for choosing coordinate systems.

Storing the registration information in the sequence information file corresponds to considering the registration, part of the calibration of the images. This way subsequent analysis of the image sequence can assume that the images are registered. If however, the registration is thought of as part of the analysis of the image sequence, the results of the registration should be communicated along with the results of the analysis, or in a separate registration file.

Input to the registration algorithm is a sequence information file, specifying the filenames of the individual images.

### 9.3.3 Automatic 3D Detection of Spherical Blastomers

The automatic detection of spherical blastomeres is done by linear filtering, followed by a 3D local extremum detection. The filtering consists essentially of correlation of the individual images in the measured HMC sequence with model HMC images of spheres with varying radius.

#### Position Detection by Maximum Correlation

Position detection is achieved by global search for the maximum correlation modulus in the high pass filtered correlation images (see Figure 9.1) of the periodically extended Images \(\tilde{I}, \tilde{M}\):

\[
\text{Corr}(\tilde{I}, \tilde{M}_r)(x_0, y_0) = \int \int \tilde{I}(x,y)\tilde{M}(x-x_0,y-y_0)dx dy
\]  

Figure 9.1: High pass filtered correlation image.
has complexity $O(N^2 \times N^2) = O(N^4)$, if not implemented using Fast Fourier Transformation and convolution

$$\text{Corr}(\hat{I}, \hat{M}_r)(x_0, y_0) = \int \int \hat{I}(x, y) \hat{M}_r(-(x_0 - x), -(y_0 - y)) dxdy \quad (9.6)$$

$$= \int \int \hat{I}(x, y) \hat{M}_r(x_0 - x, y_0 - y) dxdy \quad (9.7)$$

$$= (\hat{I} \ast \hat{M}_r)(x_0, y_0) \quad (9.8)$$

by multiplication in the Fourier domain (Convolution Theorem):

$$\text{FFT}(\hat{I} \ast \hat{M}_r)(k_x, k_y) = \text{FFT}(\hat{I})(k_x, k_y) \text{FFT}(\hat{M}_r)(k_x, k_y) \quad (9.9)$$

giving a complexity of $O(N^2 \log(N))$.

### 9.3.4 Inferring the Size

When both position and size is unknown, we detect the global extrema of the correlation function:

$$(x_0, y_0, r) \mapsto \text{Corr}(I_{z_0}, I_{M_R})(x_0, y_0), \quad (9.10)$$

which we calculate as many correlation images with model images of spheres with varying radius. Such a sequence of correlation images is illustrated in Figure 9.2

![Figure 9.2: Sequence of correlation images of the same observed image with varying radius sphere model images. Notice how the correlations peak in the last correlation image, where many parts of a cell contour agree on the same center position for this radius. This effect makes the correlation filtering similar to a Hough transform.](image)

Unfortunately it is not easy to separate the detection of the local extrema from the filtering, since the 4D volume (3D space and 1D radius) of correlation values easily gets too big, when a suitable range of sphere model sizes is used. In order to cope with this problem, we do the local extremum detection in-between the different filterings, and keep only (the largest of) the local extrema.

### 9.3.5 Filtering Part

The filtering part needs of course access to the images, but also to the filters that should be used, or at least to a specification of them in terms of what size blastomers to look for. Since these filters depend on the HMC analysis direction, this has to be known or found as well.

As it is implemented today, the filters are computed individually for each image sequence after the used HMC analysis direction has been detected. Since creating these filters’ Fourier transforms takes quite some time, the use of precomputed filters could be useful in online filtering to help speed up the processing. As far as I can see, the problem, that the filters have to depend
on the HMC analysis direction, can be solved by using only two sets of steerable filters. The non-isotropy of the filter originates from a directional derivative, and this family of filters is spanned by just the two filters differentiating along the two coordinate directions. I thus believe that one could do with only two sets of pre-calculated filters. However for this to work, the images must have known sizes, or I must figure out how to re-sample the Fourier transforms of the filters.

Another way of achieving speedups, is by down-sampling the images. We have found that down-sampling to 256x256 or even 128x128, achieves processing times worth waiting for.

We choose to solve the problem of long processing times in the interactive interface to the automatic detection of spherical blastomers, by allowing the user to work on down-sampled images. This complicates a little the communication of the detection results since the coordinates have to be relative to a standard coordinate system. This problem of choice of coordinate system has to be solved anyway in relation to a registration of the images (see "local extrema detection").

For an off-line automatic detection, the processing time will not be critical, and so we can compute the filters from a simple specification of the search. That is, which size blastomers to look for, and too other parameters, which we don’t know the optimal values for yet. Such search specifications might also be useful in the interactive interface, since they could be offered as alternative standard choices, and maybe later augmented with precomputed filters.

High Pass Filtering

In the current implementation of “SELECT”, we perform the high pass filtering after having calculated correlation images with the model images. From the theoretical understanding of this filtering, we could implement this high pass filtering as a high pass filtering of the model images, thus saving a filtering step.

9.3.6 Local Extrema Detection

Unfortunately we are not always able to separate the true blastomers from false detections, simply from a thresholding of the above mentioned local extrema of filter outputs. This is largely due to the problem that each blastomer creates multiple detections, but also because image evidence from combinations of blastomers is wrongly interpreted as a blastomer. This difficulty of setting one threshold value which is valid for many different embryos and blastomers is one of the main motivations for introducing user interaction.

9.3.7 The No-Overlap Criteria

To partly overcome this problem we can postprocess the local extrema, by demanding that the spheres do not overlap (too much). This solves the problem with most of the false detections. However, since there are still some left, we have to allow for manual interactive 3D inspection and correction of the sphere detections. And thus the consistency analysis has been moved there as well, such that only accepted spheres are allowed to rule out other detections.

Maybe we should consider attacking these two problems separately. The demand for consistency should still be met by only considering (almost) non overlapping spheres. Whereas the problem of multiple detections of the same blastomer could be handled by a quantization of the space of spheres, essentially only keeping one local extremum as a representative for a group of local extrema. This could be implemented by "throwing away" local extrema that are similar to a stronger local extremum, in the sense of having the center and radius within some limits. This would be very similar to the consistency postprocessing, only with a stronger degree of similarity.
9.4 Conclusion

- The simple understanding of in-focus HMC-images led to a model image useful in automating cell detection.

- Global search by correlation was only possible because of the convolution theorem and efficient implementation using the Fast Fourier Transform.

- The focusing effects in the images are only modelled indirectly in terms of the general low frequency background model, and so they are not used in the detection.

In the next chapter we evaluate this “SELECT” program, and identify problems and their possible origins.
Chapter 10

Evaluation of Semiautomatic Blastomer Detection

This chapter documents an evaluation of the program “SELECT”, for detection of blastomers in HMC focus sequences of human embryos. Both an automatic procedure and the results of a minimal interactive selection of the number of blastomers is evaluated, on 21 HMC-image sequences with annotations of the true blastomer contours\(^1\).

10.1 What and How To Evaluate

This subsection contains a discussion of the chosen way of evaluating the program “SELECT”.

Ideally we should evaluate both the “validity” and the “quality” of the results. And in case of non-satisfactory performance, we also want to “isolate the errors” and “locate their origin”. Since “SELECT” is an interactive program, we could also evaluate the “functionality” or “ease of use” by measuring the “complexity of the interactions” needed to reach a certain result. We believe that this last type of evaluation of "functionality" should be done by the end users, and only after an evaluation of the validity and quality.

10.1.1 Evaluating Validity and Quality

To investigate the validity we want to answer the following questions:

- Does it detect all cells?
- Does it only detect true cells?
- Do the detected cells have the correct size and position?

They have all been phrased as “Yes or No”-questions, which should all be answered with a “yes” for the program to be strictly valid. Since it is likely that there will be some miss-detections, we would also like to measure the degree of validity.

We can quantify the degree of validity by measuring the quality of the detection results in terms of the following questions:

- For each detection - is it true or false? The criteria for being a true detection is of a qualitative character. The decision will be based on my visual judgment of whether the detection is "correctly triggered by a true cell".

\(^1\)The images and annotations were supplied by Christina Hnida and Søren Ziebe, and the evaluation was performed in collaboration with Jon Sporring.
• For the false detections I should make a note of my immediate understanding of the cause. How to quantify this I don’t know, before I know the possible causes.

• For the missing detections I should also make a note of the cause.

• For the correctly detected cells I should quantify the deviation in position and size from the ground truth. Specifically I will check the z-position, to check if the equator plane image was correctly detected. And last I will compare the found radius, with the mean radius of the ground truth equator contour.

10.1.2 Evaluating User Interaction

We will run the program on a "representative set of embryos", and compare the results to manual markings and/or visual inspection.

Since the program is interactive, it can produce many different results depending on the interactive inputs. Also since the program uses some adjustable parameters (like down-sampling factor, diffraction length, background scale and what size cells to search for), the results may depend on these.

The evaluation of “validity” and “quality” is thus not a simple single evaluation, but potentially an evaluation of a functional relationship between parameter settings and user interaction on one side and the "validity and quality" of the resulting detections/selections on the other side. Last but not least the results of course depend on the image data that is analyzed. If I denote image data by $I$, the parameter settings by $P$, and the user interaction by $U$, I can formally express the functional dependence, of the quality denoted by $Q$, on $I$, $P$ and $U$ as

$$Q(P, I, U) = Q(SELECT(P, I, U), M_{true}(I)).$$ (10.1)

I can now discuss how the evaluation should be performed in terms of a strategy for varying $I$, $P$ and $U$, while evaluating $Q$. To keep things realistic we will keep $P$ fixed to the same setting for all data sets $I$ from a fixed “representative ensemble of image sequences”. We could however consider varying $P$ and then repeating the evaluation for the full set of embryo Images, in order to find the “optimal” values for $P$. For example by maximizing the “worst case quality, with optimal interaction”:

$$\max_P \min_I \max_U Q(P, I, U).$$ (10.2)

This however implies trying out many $P$ and $U$, and is thus very laborious. The other extreme of fixing $U$, effectively evaluates “SELECT” as a fully automatic procedure. As already mentioned, this is what we will do first. Regarding the choice of $P$ or its systematic variation, we will start by using one fixed setting, and in a later evaluation address the optimality of this $P$, by trying other choices of parameter settings $P$

An evaluation of the interactive part of “SELECT” and its "ease of use" should be done in an environment and setting where it is intended to be used, and with a quality measure designed for the context in which it is used as a tool.

10.1.3 3 Evaluation Iterations

Based on the above analysis of different evaluation strategies, we decided to perform 3 iterations of evaluation, where the conclusions of the previous evaluations, influenced the design of the following evaluations:

1. We first focus on the "automatic parts", and use a fixed parameter setting $P$, and user interaction $U$. 

86
2. Based on the results of the evaluation of the automatic parts, we try varying the parameter settings in 3 different ways, and evaluate these using a very small amount of user interaction $U$, where the user can set an upper limit of how many cells to detect.

3. Finally we implemented a non-greedy combinatoric search for the optimal combined model, using the best parameter settings $P$ found in the previous evaluation and a user supplied number of true cells.

These 3 evaluations are described below.

### 10.2 Evaluation I - Automated Cell Detection

We present the results of a manual evaluation of the automated cell detections in 21 image sequences of human embryo, done by the “SELECT” program. The detections have been evaluated manually by classifying the detections as true positive (accepted detections) or false positive (rejected detections).

We used the “sif2mat” program to batch generate the raw list of cell detections sorted according to their measured evidence. It is this list from which the “SELECT” program draws the at most 200 most evident cells, when it presents the user for suggested cells. We used the “SELECT” program to suggest as many non-overlapping\(^2\) cells as possible. We manually accepted the true positive detections and rejected the false positive detections.

The “result” of the automatic detections and the manual evaluation is presented using the “documentmatfile” program.

#### 10.2.1 Analysis

We are at first interested in an evaluation of the automatic detections just as a detection algorithm and not as an algorithm for measurement of position and size of the individual cells. We thus count, for each of the image sequences:

- The true number of cells
- The number of detected cells.
- The number of true positive detections.
- The number of false positive detections.
- The number of false negative detections.

For each of the false detections, we will make a note of “type” and “cause”.

#### 10.2.2 Input

We used 21 image sequences for the evaluation. Appendix A presents circular approximations to the equator contours based on manual markings by Søren Ziebe of the cell-contours.

For the automatic cell detection we used a search specification file

```matlab
function maked2ssfile(ssffilename)
    
    % This script creates a default Search Specification File with the name ssffilename
```

\(^2\)We use a criteria of the distance between the centers being greater than 75 percent of the sum of the two radii.
10.2.3 Manual Evaluation

The raw automatic detections were analyzed using the interactive “SELECT” program. The interactive input was kept as “automatic” as possible. We thus interactively asked for display of as many non-overlapping cells as possible. So this part of the interaction could just as well have been implemented as an automatic procedure.

However the input during the manual evaluation is difficult to describe as a precisely defined criteria for accepting or rejecting a shown detection. The goal was to accept those suggested detections that were

“rightly triggered by part of a true cell”.

We have thus accepted some suggested cell detections even if they were of a measured size or position much different from the true size or position. Similarly we have rejected some that were pretty close to a true cell, but were triggered (mainly) by some other image feature than the true cell itself.

10.2.4 Output

In this section we present the results of the automatic detections and the manual evaluation.

10.2.5 Manual Evaluation of Detections

Below we show a full detail example output for sequence 2 (200100013)\(^3\), including also all suggested cells. The figures display the equator-contour of each of the detected cells (accepted in green, rejected in red and suggested in blue) in the image which is the equator plane of the cell. In appendix B we present similar figures for all the sequences, except that suggested cells ruled out by the overlap criteria are not shown, to save space.

Example Output for Sequence 2 (200100013)

\[^3\]The reader may recognize this as the example sequence from chapter 1
10.2.6 Detection Statistics

Based on appendix A and B, we made a table, which for each of the eggs lists:

- The true number of cells
- The number of detected cells (number of accepted and rejected).
- The number of true positive detections (number of accepted).
- The number of false positive detections (number of rejected).
- The number of false negative detections (number of cells not detected).
<table>
<thead>
<tr>
<th>Sequence</th>
<th>True Number of cells</th>
<th>Number of Detected Cells</th>
<th>True Positive</th>
<th>False Positive</th>
<th>False negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (12)</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2 (13)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3 (18)</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>4 (19)</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5 (22)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6 (23)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 (24)</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>8 (25)</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9 (26)</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10 (29)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11 (30)</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>12 (31)</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>13 (32)</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>14 (33)</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15 (34)</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>16 (37)</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>17 (38)</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>18 (42)</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>19 (43)</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>20 (45)</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>21 (47)</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>total</td>
<td>83</td>
<td>68</td>
<td>43</td>
<td>25</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 10.1: Detection Statistics

10.2.7 Types and Causes of False Detections

In order to get a better understanding of the shortcomings of the automatic detection algorithm, we have tried to classify the false detections by “type” and “cause”. The hope being that the noted types can be used as categories in a future automatic evaluation against the ground truth, and the causes will point out possible improvements to the algorithm.

We have thus visually inspected the automatic detections (appendix B) in each of the image-sequences comparing them with the true cells (appendix A), classifying the “type” and “cause” of each of the false detections.

In order to be able to decide, if non-suggested cells were detected, but just didn’t pass the thresholding for the suggested cells, we prepared figures with the equator contours of the 1000 most evident cells. These can be seen in appendix C, along with a graph of the correlation based evidence measure used.

The systematic results of the visual inspection of each image sequence is presented in appendix D. Here we just repeat all the types and causes of false detections that we came across:

False Detection Types

1. False Positive Detection Types:
   
   1. Multiple detection by out of focus cell.
   2. Wrong sign of extrema.
   3. Coincidence between two (or more) structures/cells.
4. Fragment.
5. Wrong size, partly triggered by coincident structure.
6. Wrong z-position.

2. False Negative Detection Types:
   1. Not suggested.
   2. Suggested but overlapping with true positive.
   3. Suggested but overlapping with false negative.
   4. Not detected as 1 of the 1000 strongest.

**False Detection Causes**

1. False Positive Detection Causes:
   1. Wrong cell image model.
   2. Bad overlap criteria.
   3. False negative didn’t rule out this false positive.
   4. Wrong model match measure?
   5. Wrong model size (too small/large).
   6. Too low threshold.
   7. True detections not strong enough.

2. False Negative Detection Causes:
   1. (Not suggested) Too high threshold for suggested.
   2. (Not suggested) Cell image lacks (white/dark) edges, (due to influence from other cells?).
   3. (Not suggested) Model image lacks out of focus dark shadow halo.
   4. (Not suggested) model image is not elliptical.
   5. (Suggested but overlapping) Non flexible sphere model. (Strong elliptic rules out too much)
   6. (Suggested but overlapping) Wrong z-positions from Greedy algorithm.
   7. (Suggested but overlapping) Wrong number of cells from Greedy algorithm.
   8. (Suggested but overlapping) Lack of integrated match measure for many cells.
   9. (Suggested but overlapping) Too strict overlapping criteria.
   10. Wrong image model.
   11. (Suggested but overlapping) Model is not elliptic. (week because elliptic)

### 10.2.8 Statistics of Detection Error Types and Causes

Below we list how many times each of the false detection types and causes have been encountered. Since some of the false detections have been classified as belonging to more than one type or cause\(^4\), the numbers don’t add up. The numbers may anyway give an idea of which types and causes of false detections are most often encountered.

\(^4\)And partly because the bookkeeping has been done manually.

91
Table 10.2: False Positive

<table>
<thead>
<tr>
<th>Cause</th>
<th>Type</th>
<th>1.1 unfocused cell</th>
<th>1.2 sign</th>
<th>1.3 coincidence</th>
<th>1.4 fragment</th>
<th>1.5 size</th>
<th>1.6 z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 model</td>
<td></td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.2 overlap</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.3 false negative</td>
<td></td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1.4 match measure</td>
<td></td>
<td>0</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1.5 size</td>
<td></td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.6 low threshold</td>
<td></td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 10.3: False Negative

<table>
<thead>
<tr>
<th>Cause</th>
<th>Type</th>
<th>2.1 not suggested</th>
<th>2.2 overlapping true pos.</th>
<th>2.3 overlapping false neg.</th>
<th>2.4 not detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 (NS) high threshold</td>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.2 (NS) lack of edges</td>
<td></td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2.3 (NS) lack of focusing</td>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.4 (NS) non elliptical model</td>
<td></td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2.5 (SO) strong elliptic</td>
<td></td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2.6 (SO) wrong z from greed</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.7 (SO) wrong N from greed</td>
<td></td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>2.8 (SO) no integrated measure</td>
<td></td>
<td>0</td>
<td>0</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>2.9 (SO) strict overlap</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.10 wrong image model</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2.11 (SO) week because elliptic</td>
<td></td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

### 10.2.9 Observations

Based on the above evaluation, we observe that:

- In many cases the False Positive detections and the False Negative detections were closely related. 10 times did we observe a False Positive detection triggered by a coincidence of 2 or more structures/cells, which wasn’t ruled out by False Negative detections. Similarly we observed 22 False Negative detections which were ruled out by a False Positive detection.

- The second most occurring type of False Positive detections were spurious detections of structures matching the inverted image model. These can be eliminated by fixing the sign of the HMC direction.

- The second most occurring type of False Negative detections were those that were “not suggested”, but which were detected. They just didn’t pass the thresholding which only allowed the 100-200 strongest detections. Most of these (8) had an elliptical shape which resulted in weak detections by the circular model. Also 7 of the detected but not suggested cells lacked some of their contour edges (probably due to influence from other cells).
10.3 Evaluation II - First Improvements

Based on the observations from the first evaluation of the automated cell detections, we decided to:

1. Adjust the interval of searched radii, in order to eliminate the very big or very small false positive coincidental detections. This is based on an assumption, that we know the expected number of cells, and their typical radii. We thus restricted the search from $r \in \{10, 11, ..., 70 \mu m\}$ to $r \in \{20, 21, ..., 40 \mu m\}^5$.

2. Fix the sign of the HMC direction, and only consider extrema in correlation values corresponding to spheres with higher refractive index than the surroundings.

3. Change the minimal distance allowed between two cell centers from $3/4$ to $5/8$ of the sum of their radii in the overlap criteria, allowing a slightly bigger overlap$^6$.

We thus performed the following 3 automatic detections:

1. • Signed HMC direction  
   • $r \in \{20, 21, ..., 40 \mu m\}$,  
   • minimal cell distance $3/4$

2. • Signed HMC direction  
   • $r \in \{10, 11, ..., 40 \mu m\}$,  
   • minimal cell distance $5/8$

3. • Signed HMC direction  
   • $r \in \{20, 21, ..., 40 \mu m\}$,  
   • minimal cell distance $5/8$

The resulting detections were evaluated manually using the “SELECT” program by showing enough suggestions to see all the true suggestions$^7$. We marked the true positive detections as accepted and the false positive as rejected. The results of the detections and their evaluations can be seen in appendices E, F and G. The statistics of these evaluations have been summarized in the table below.

10.3.1 Improved Detection Statistics

We have made a table which, for each of the image sequences, lists:

- The true number of cells
- The number of detected cells for each of the 3 evaluations.
- The number of true positive detections for each of the 3 evaluations.
- The number of false positive detections for each of the 3 evaluations.

$^5$This has the obvious drawback of not finding any cells outside the searched size interval.

$^6$This has the drawback of possibly allowing false positive multiple detections.

$^7$We are thus sure to see all the true positive suggestions, but will not see any false positive detections that have less evidence than the least evident true positive. The motivation for this change in evaluation strategy is to mimic the incorporation of the knowledge of the true number of cells. But rather than always stopping at the true number of cells, even when some true suggestions are not shown, we check if they are there and show enough to see them.
• The number of false negative detections for each of the 3 evaluations.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>True Number of Cells</th>
<th>Number of Detected Cells</th>
<th>True Positive</th>
<th>False Positive</th>
<th>False negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>II1</td>
<td>II2</td>
<td>II3</td>
<td>II1</td>
<td>II2</td>
</tr>
<tr>
<td>1 (12)</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2 (13)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3 (18)</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4 (19)</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5 (22)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6 (23)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7 (24)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>8 (25)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>9 (26)</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>10 (29)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>11 (30)</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>12 (31)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>13 (32)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>14 (33)</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>15 (34)</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>16 (37)</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>17 (38)</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>18 (42)</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>19 (43)</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>20 (45)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>21 (47)</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>total</td>
<td>83</td>
<td>82</td>
<td>99</td>
<td>91</td>
<td>76</td>
</tr>
</tbody>
</table>

10.4 Evaluation III - Combinatoric Search

Based on the observations from the second evaluation of the automated cell detections, we decided to:

1. Given (by the user) the true number of cells, do a combinatoric search for the non overlapping detections with maximum simultaneous evidence, as measured by the sum of each of the detection correlation values.

We thus performed a combinatoric search using the same search and overlap specification as in the 3'rd of the first improvements evaluation:

• Signed HMC direction
• \( r \in \{20, 21, \ldots, 40 \mu m\} \),
• minimal cell distance 5/8

The resulting detections are presented in appendix H.

10.4.1 Evaluation

Since we this time search for the number of true cells, and thus always detect this number of cells, we will always find the same number of false positive detections as false negative detections. What we are interested in is whether or not the combinatoric search will
- find the weekly detected or overlapped true cells, which would otherwise have been ruled out by a greedy algorithm
- without finding any new wrong detections.

Of course, the combinatoric search will not help the problem of non-suggested cells. We thus still expect at least 1 false negative detection in sequence 1(12).

By visual inspection of the presented cell detections in appendix H, I found the following false positive detections:

12 Sphere 3 and 4 are multiple detections of the same cell. I thus say sphere 4 is a false positive.
   Cause: the 4’th true cell is not suggested:

25 Sphere 4 is partly falsely triggered. Cause: The filter used is triggered by the other out-of-focus cell:

26 I think sphere 5 is a false positive? Cause: Too simple filtering:

30 Sphere 1 and 3 are multiple detections of the same elliptical cell. I thus say sphere 3 is a false positive. Cause: We only have a spherical model:
Sphere 4 is partly falsely triggered by a small fragment. Cause: The true cell that is partly detected, is elliptical. The fragment is not recognized and explained as a fragment:

We thus found 5 errors, giving a total of 10 false positive or false negative detections. In each of the 3 “first improvements” to the greedy algorithm we had the following total number of false detections:

1. $6+7 = 13$
2. $21+5 = 26$
3. $12+4 = 16$

10.5 Conclusions

- We can detect most spherical cells.
- Non-spherical cells are not always detected.
- Many wrong detections are triggered by matching to more than one cell.
- When faced with the problem of detecting multiple cells, the greedy optimization algorithm for multiple blastomers does not always give the optimum.
- Combinatoric search for non-overlapping cell detections based on the true number of cells, still resulted in 5 pairs of false detections.
- Using the combinatoric search in conjunction with the information of the true number of cells, we are thus able to further reduce the total number of false detections to a total of 10 (5 false positive and 5 false negative) when faced with 83 cells in 21 eggs.
- More accurate models of blastomer interactions are needed.
- Deformable models of cells are needed and should be integrated with the models of blastomer interactions.
- More accurate models of the focusing effects in the images are needed.
Part IV

Shape Models
Chapter 11

Lie Group Models of Point Set Shape Variability

This chapter\(^1\) introduces a Lie group model of point set shape variability, and analyzes it in the setting of point sets in the plane. Modeling point set variability with groups of transformations generated by linear vector fields constitute an algebraic frame for modeling simple nonlinear point set variability suitable for the modeling of shape variability. The main result is that, for centered point sets in the complex plane the complex linear vector fields also induce well defined shape variabilities.

11.1 Introduction

Linear statistical models of variability of identifiable point sets have previously been described and applied successfully to the empirical modeling of appearance variability in natural images. One of the limitations of these linear models has been demonstrated in the nonlinear “bending” shape variability of point sets where a length ratio is constant. As we shall see, this very simple “bending” shape variability of three points in the complex plane, is in fact an example of a Lie group which is generated by a linear vector field described by a complex \(3 \times 3\) matrix.

11.1.1 Shape

Shape is often defined as whatever is left when position, size and orientation are ignored [9, 29]. Abstractly the set of shapes can be defined as a set of equivalence classes, where the equivalence is “equal except for translation, scaling and rotation”. Often these equivalence classes are modeled by a choice of a canonical element from each class.

When one considers a finite set of uniquely identifiable points in the plane, space or generally \(\mathbb{R}^m\), an explicit assumption of invariance of variability under centroid translation, scaling and rotation leads to the statistical theory of shape as introduced by David G. Kendall in 1977 [29]. In this theory the space of shapes is modeled as a Riemannian manifold. For an in-depth coverage please consult one of the two books “The statistical Theory of shape” [42] by Christopher G. Small and “Statistical Shape Analysis” [17] by Ian L. Dryden and Kanti V. Mardia.

To keep things simple only point sets in the plane will be considered. Though this restriction avoids the non-trivial generalizations of rotations and the complications of singular point set

\(^{1}\)This chapter contains a slightly modified version of an article[35] written by Niels Holm Olsen and Mads Nielsen and presented at the second international workshop on Algebraic Frames for the Perception-Action Cycle, in Kiel September 2000.
configurations invariant under non-trivial subgroups of rotations [42, p. 84], it is not essential for the modeling by 1-parameter groups of transformations.2

The set of n-point sets in the plane \((\mathbb{R}^2)^n\) is conveniently modeled by the complex vector space \(\mathbb{C}^n\), since the rotation of an n-point set is then simply given by multiplication with a unit complex number. This identification is only introduced to get a short and convenient notation, and we will still consider \(\mathbb{C}^n\) as a 2 \(n\)-dimensional real vector space.

11.1.2 Kendall’s Manifold of Shapes of Point Sets in the Plane

To fix terminology and notation we shortly review Kendall’s shape space of n-point sets in the plane. First position normalization is done by orthogonal projection onto the subspace of “centered point sets” \(C_{2n-2} = \{(z_1, \ldots, z_n) \in \mathbb{C}^n | \sum_{i=1}^n z_i = 0\}\) with real dimension \(2n - 2\). Size normalization is done by scaling onto the manifold of “pre-shapes” \(S_{2n-3} = \{p_0 \in C_{2n-2} | \|p_0\|_2 = 1\}\). \(S_{2n-3}\) is thus a sphere of real dimension \(2n - 3\) in the linear subspace \(C_{2n-2}^2\) of \(\mathbb{C}^n\).

Now the manifold \(\Sigma_2^n\) of shapes of n-point sets in the plane can be identified with the complex projective space \(\mathbb{CP}^{n-2} = \{\{zp | z \in \mathbb{C}\} | p_0 \in C_{2n-2}^2 \setminus \{0\}\}\). Kendall’s “Shape Space” \(\Sigma_2^n\) is a Riemannian manifold with the Procrustes metric:

\[
d(\Sigma(p), \Sigma(q)) = \cos^{-1}(\sqrt{\sum_{k=1}^n p_k q_k^*}), \quad p, q \in S_{2n-3},
\]

where \(\Sigma(p) = \{\{zp | z \in \mathbb{C}\} \in \Sigma_2^n\) is the equivalence class representing the shape of \(p\) [42, p. 13]. In the following we shall simply refer to \(C_{2n-2}^2\) as \(C_0\) without explicitly noting the dimension.

11.1.3 Linear Point Set Models of Local Shape Variation

In “Active Shape Models” [16] Cootes et al. use a linear model of local point set variation. The example point sets of which they model the variation have all been translated, scaled and rotated to match as well as possible4 a mean point set, which has been given a standard position, size and orientation and is found iteratively. They use a local linear model of the variation of the matched point sets:

\[
p = \overline{p} + P b,
\]

where \(p \in (\mathbb{R}^2)^n\) is a point set and \(P\) is a matrix, which consists of the eigenvectors of the covariance matrix describing the example point sets deviations from the mean point set \(\overline{p}\). The variable \(b\) is a column vector of parameters describing the point set deviation from the mean. It is a linear model of local shape variation in the sense that the point set deviation from the mean depends linearly on the parameters \(b\).

They give an example of shape variability where this local linear model is inappropriate. The example is point sets along the outline of worms, and the problem is that they can bend at the middle. This bending is intuitively a 1-dimensional shape variability, so even though translation, scaling and rotation have been explicitly removed, there is one significant variation left.

The problem is that it is not possible to model the bending exactly by a 1-dimensional local linear model. When the bending is large, 2 eigenvectors are needed to span the example point set configurations. This problem of bending can be studied in the very simple setting of only three points in the plane allowed to vary freely under the constraint that a length ratio is preserved.

---

2Generalizations to 3D and higher dimensions is analyzed in the next chapter, where we do find non-trivial 1-parameter groups generated by linear vector fields commuting with the general rotation groups.

3When all points are coincident this is not possible. Thus these point sets are excluded.

4They use a non-trivial weighting of the distances of corresponding points. This is not essential for the point made here.
This is an inherent problem of modeling the variation of point sets on the unit sphere of pre-shapes $S_{0}^{2n-3}$. This sphere is curved and thus any large variation will be of a nonlinear nature. To overcome this difficulty Kent [30] uses a tangent space approximation to the pre-shape manifold.

However it is still possible to think of a 1-dimensional shape variability which after projection on a tangent space has a nonlinear image. In the section on the bending of three points in the plane we shall see an example where the preserved length ratio is different from 1 and where the projected shape variation is nonlinear.

**11.1.4 Nonlinear Models of Point Set Shape Variability**

We have thus been inspired to study differential geometric modeling of nonlinear point set and shape variability. The theory of Lie groups and their Lie algebras provide a framework for modeling nonlinear continuous 1-dimensional modes of variation (variability) by 1-parameter groups of transformations. These 1-parameter groups are generated by vector fields describing the modeled variability. The linear point set models are in this framework generated by the constant vector fields. However, in the context of modeling shape variability the requirement of commutativity with scaling naturally leads to the study of point set variability generated by linear vector fields.

Lie group models of point set variability not only deal with the subject of modeling nonlinear point set variability, but also allows us to analyse the modeling of point set variability in the context of the inherently nonlinear shape space of point sets.

**11.2 Variability Modeled by Group of Transformations**

A 1-dimensional point set variability may be modeled by a 1-parameter group $(T_t)_{t \in \mathbb{R}}$ of transformations of $(\mathbb{R}^2)^n$. That is, it is assumed that all variations originates from the same variability and are additively parameterized:

$$T_t \circ T_s = T_{t+s}.$$  \hspace{1cm} (11.3)

Thus, a variation corresponding to a finite change is modeled by a transformation, while a variability is given by a continuous family of transformations corresponding to the continuum of degrees of variation describing a continuous course of change [36].

In this frame the linear model is written as a 1-parameter group $(T_t)_{t \in \mathbb{R}}$ of translation transformations of $(\mathbb{R}^2)^n$

$$p = T_t(\mathbf{p}) = \mathbf{p} + t \Delta p,$$  \hspace{1cm} (11.4)

where $\Delta p \in (\mathbb{R}^2)^n$ specifies the direction of the linear variability.

Higher dimensional variabilities may be modeled by independent 1-dimensional variabilities.\textsuperscript{5} Two 1-parameter groups $(X_t)_{t \in \mathbb{R}}$ and $(Y_s)_{s \in \mathbb{R}}$ are capable of describing independent variabilities if they commute:

$$\forall t \in \mathbb{R}, \forall s \in \mathbb{R} : X_t \circ Y_s = Y_s \circ X_t.$$  \hspace{1cm} (11.5)

**11.2.1 Continuous Variability Generated by a Vector Field**

Just as the linear variability above was described by $\Delta p \in (\mathbb{R}^2)^n$, a nonlinear variability may intuitively be described by a vector field on $(\mathbb{R}^2)^n$, which everywhere points in the direction of change under the studied variability. Consider the integral curves $(\gamma_p)_{p \in (\mathbb{R}^2)^n}$ defined by the action of a 1-parameter group $(T_t)_{t \in \mathbb{R}}$:

$$\gamma_p(t) = T_t(p), \hspace{0.5cm} t \in \mathbb{R}, p \in (\mathbb{R}^2)^n.$$  \hspace{1cm} (11.6)

\textsuperscript{5}For simplicity we have here excluded variabilities with other topologies and inherent non-commutativity. Such a variability can be modeled by a non-commutative Lie group of dimension higher than 1.
The derivative of these curves define a vector field \( X : (\mathbb{R}^2)^n \to (\mathbb{R}^2)^n \) satisfying
\[
\gamma_p'(t) = X(\gamma_p(t)), \quad \forall t \in R. \tag{11.7}
\]
On the other hand, such a vector field uniquely identifies a 1-parameter group [51, p. 37]. In this way the vector field \( X \) is said to generate a 1-parameter group of transformations \((X_t)_{t \in R}\) where the parameter indicates how far along the variability the point set should be transformed.

We now observe that when logarithmically parameterized the group of point set scalings \( s(p) = e^r p \) is generated by the linear vector field given by the identity transformation \( I \) of \((\mathbb{R}^2)^n\):
\[
\frac{\partial}{\partial s}\bigg|_{s=0} s(p) = p = I(s_0(p)). \tag{11.8}
\]
Similarly using complex notation it is seen that rotation \( \rho_\theta(p) = e^{i\theta} p \) is generated by the linear vector field given by \( iI \):
\[
\frac{\partial}{\partial \theta}\bigg|_{\theta=0} \rho(p) = ip = iI(\rho_0(p)). \tag{11.9}
\]
It is well known that point set centroid translations \( \tau(p) = (p_1 + a, \ldots, p_n + a) \) do not commute with scaling and rotation.\(^6\) We observe that this holds in general for the (non-trivial) linear variabilities generated by the constant vector fields.

### 11.2.2 Well Defined Shape Variability Generated by Linear Vector Fields

In the context of point set shape variability it is natural to consider only point set variabilities commuting with the shape similarity group of point set translations, scalings, and rotations. Such point set variabilities induce well defined shape variabilities. The induced shape transformations are defined by applying the corresponding point set transformations on all the point sets in a shape equivalence class. Because of commutativity with the shape similarity group the resulting point sets will all belong to one and the same shape equivalence class, thus providing a well defined shape transformation.

The problem of non-commutativity between the linear variabilities and scaling and the wish to model nonlinear bending naturally lead to higher order modeling of the generating vector fields. As a first simple step towards the general case it is natural to consider 1. order vector fields.

Since commutativity between 1-parameter groups corresponds to commutativity of the generating vector fields\(^7\) [44, lemma 13 p. 5-35] and the Lie bracket between linear vector fields \( X, Y \) is given by \([X, Y] = X \circ Y - Y \circ X \) [51, p. 87], it is seen that the requirement of commutation with scaling (generated by the identity vector field \( I \)) is automatically fulfilled when considering real linear vector fields on \((\mathbb{R}^2)^n\). In order to secure commutation with centroid translations \( s = (s_1, \ldots, s_n) \) and because we only want variabilities which do not change the centroid position\(^8\) we only consider real linear vector fields on the subspace of centered point sets \( C_0 \) and extended to zero on the orthogonal subspace \( C_0^\perp \) modeling the centroid position.

It remains to analyze the requirement of commutativity with point set rotation. Not all real linear vector fields on \( C_0 \subset (\mathbb{R}^2)^n \) commute with the vector field generating point set rotation which is most easily expressed as \( iI \) using the complex representation \((\mathbb{R}^2)^n = \mathbb{C}^n\). From this it is seen that the complex linear vector fields do commute with rotation.

To summarize [36]:

\(^6\)One may without changing the defined shape space choose to consider centroid centered scaling and rotation which do commute with centroid translation. These centered scalings and rotations are generated by the linear vector fields given by the identity \( I_{C_0} \) on the subspace of centered point sets and \( iI_{C_0} \) [36].

\(^7\)Two vector fields commute when their Lie bracket is zero.

\(^8\)This is an orthogonality constraint stating that the vector field should be everywhere orthogonal to the two constant vector fields \((1, \ldots, 1)\) and \((i, \ldots, i)\) generating centroid translations.
• Commutativity with point set translations - Is obtained by working in the subspace of centered point sets.

• Commutativity with centered scaling - Is obtained by real linearity of the vector field on the subspace $C_0 \subset (\mathbb{R}^2)^n$ of centered point sets.

• Commutativity with centered rotation - Is obtained by complex linearity of the vector field on the subspace $C_0 \subset \mathbb{C}^n$ of centered point sets.

The above arguments have only resulted in sufficient conditions for a vector field to describe a well defined shape variability. They are not necessary.

11.2.3 1-parameter Groups Generated by Complex Linear Vector Fields

Consider the Lie group $GL(n,C)$ of nonsingular $n \times n$ complex matrices. This has Lie algebra $gl(n,C)$ which consists of all $n \times n$ complex matrices. A complex linear vector field $X : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is represented by a complex $n \times n$ matrix $X \in gl(n,C)$. The corresponding 1-parameter group of transformations $(X_t)_{t \in \mathbb{R}}$ is obtained by considering the usual matrix exponential map $\exp : gl(n,C) \rightarrow GL(n,C)$ [3, p. 283]:

$$X_t = \exp(tX) = I + tX + (1/2!)(tX)^2 + (1/3!)(tX)^3 + \cdots.$$  \hspace{1cm} (11.10)

11.3 Preserved Length Ratio - Nonlinear “Bending”

Complex linear vector fields commute with both scaling and rotation, and they can describe the nonlinear bending of three points in the plane as a one dimensional variability.

The simplest non-trivial example of a shape space is the shape space of three points in the plane, $\Sigma^3$ and is a 2-dimensional Riemannian manifold isometric to the sphere in $\mathbb{R}^4$ with radius $1/2$, $S^2(1/2)$ [42, p. 70, 73]. It can thus be visualized by orthogonal projections on three orthogonal planes. In the following we will describe how the nonlinear bending shape variability of three points is generated by a complex linear vector field.

11.3.1 Centered Bending of 3 Points in the Plane

Centered bending of 3 points in the plane $p_1, p_2, p_3 \in \mathbb{R}^2$ with centroid position $p_c \in \mathbb{R}^2$ and the constant distances $|p_1 - p_2| = l_1$ and $|p_2 - p_3| = l_2$ can be parameterized by the angle $\theta$:

$$p_1 = \frac{-2l_1 e^{i\theta} - l_2 e^{-i\theta}}{3} + p_c,$$

$$p_2 = \frac{l_1 e^{i\theta} - l_2 e^{-i\theta}}{3} + p_c,$$

$$p_3 = \frac{l_1 e^{i\theta} + 2l_2 e^{-i\theta}}{3} + p_c.$$  \hspace{1cm} (11.11) (11.12) (11.13)

The above family of point sets can be considered as the result of “centered bending” of $(p_1(0), p_2(0), p_3(0))$ by the angle $\theta$. But it can also be considered as the result of centered bending of $(p_1(\theta_0), p_2(\theta_0), p_3(\theta_0))$ by the angle $\theta - \theta_0$. In this way a 1-parameter group $(B_\theta)_{\theta \in \mathbb{R}}$ of “centered bending” transformations is defined. The vector field $B : (\mathbb{R}^2)^3 \rightarrow (\mathbb{R}^2)^3$, generating the 1-parameter group of
centered bending of a 3-point set is seen to be given by \( \left( \frac{\partial}{\partial \theta} \bigg|_{\theta=0} \right) p(\theta) \):

\[
\frac{\partial}{\partial \theta} p_1(\theta) = \frac{-2l_1e^{i\theta} + l_2e^{-i\theta}}{3} = \frac{i}{3}(2p_1(\theta) - 3p_2(\theta) + p_3(\theta)),
\]

(11.14)

\[
\frac{\partial}{\partial \theta} p_2(\theta) = \frac{l_1e^{i\theta} + l_2e^{-i\theta}}{3} = \frac{i}{3}(-p_1(\theta) + p_3(\theta)),
\]

(11.15)

\[
\frac{\partial}{\partial \theta} p_3(\theta) = \frac{l_1e^{i\theta} - 2l_2e^{-i\theta}}{3} = \frac{i}{3}(-p_1(\theta) + 3p_2(\theta) - 2p_3(\theta)).
\]

(11.16)

This is seen to be a linear vector field, given by the matrix

\[
B = \frac{i}{3} \begin{bmatrix}
2 & -3 & 1 \\
-1 & 0 & 1 \\
-1 & 3 & -2
\end{bmatrix} = \frac{1}{3} \begin{bmatrix}
0 & -2 & 0 & 3 & 0 & -1 \\
2 & 0 & -3 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 \\
-1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & -3 & 0 & 2 \\
-1 & 0 & 3 & 0 & -2 & 0
\end{bmatrix}.
\]

(11.17)

The above equations use first complex and then real notation corresponding to the discrimination between \( \mathbb{C}^3 \) and \( (\mathbb{R}^2)^3 = \mathbb{R}^6 \).

Both the odd and the even columns of \( B \) add to \( 0 \in \mathbb{R}^6 \) in agreement with the fact that \( B \) defines a vector field which is zero on the subspace modeling the position of the point set centroid. Similarly both the even and the odd rows add to \( 0 \in \mathbb{R}^6 \), proving that this vector field is orthogonal to the directions for translating the centroid. The vector field \( B \) thus commutes with and is orthogonal to the vector fields for translation in the x and y directions. Since \( B \) is linear and represented by a complex matrix, it also commutes with the vector fields generating scaling and rotation.

As an indirect illustration of the vector field represented by \( B \), the generated 1-parameter group \( (B_\theta)_{\theta \in \mathbb{R}} \) has been evaluated for a few different angles using the matrix exponential. These have been applied to two different 3-point configurations (see Figure 11.1). Since the linear vector field \( B \) commutes with point set translation, scaling and rotation, it generates a 1-parameter group of point set transformations inducing a well defined shape variability by acting on the point sets in the equivalence class representing the shape. This shape variability is illustrated in Figure 11.2. The figure shows orthogonal projections of \( S^2(1/2) \) on three orthogonal planes. The two coordinate axes in the projection plane and the axis coming up from the paper have been illustrated with a small corresponding point set configuration. In the left column a close sampling
Figure 11.2: “Bending” of the shape of a 3-point set in the plane
of the shape of $B_0(-1,1,\sqrt{3}) = \exp(\theta B)(-1,1,\sqrt{3}) \in \mathbb{C}^3 = (\mathbb{R}^2)^3$, for $\theta = 0, \pi/60, \ldots, \pi$ has been marked by a “+”. In the right column 7 samples ($\theta = 0, \pi/60, \ldots, \pi$) along the bending have been marked by small point-set configurations. The first and the last of these 7 shapes are the same, but the point set figures have been rotated 180 degrees relatively to each other.

The starting point set configuration has been chosen so that the bending passes through the point $(x, y, z) = (0, 1/2, 0)$. This is most easily seen in the bottom left projection, which can be considered a projection on Kent’s tangent subspace approximation to $\Sigma_3^2$ at the shape of $(0,1/2,0)$. It is thus seen that the bending shape variability of a shape with a preserved ratio of $2/(1 - \sqrt{3})$ does not have a linear image in this subspace. You may also note that the non-equidistant spacing of the bended shapes is not in harmony with the Procrustes distance, but corresponds to the choice of additive parameterization.

11.4 Conclusions

We have described how Lie groups provide a framework for modeling general variability. The classical linear models of point set variability are in this framework modeled by the Lie groups generated by the constant (0-order) vector fields.

We have found that in the context of shape variability, the natural constraint of commutativity with the similarity group provides an inherent need for nonlinear modeling of point set variability. By considering Lie groups generated by 1-order vector fields, we are able to model:

- Nonlinear point set variabilities commuting with scaling and rotation, thus inducing well defined shape variabilities.
- A larger class of variabilities which include the nonlinear bending of 3 points in the plane.

Future research will focus on methods of inferring these variabilities.
Chapter 12

Linear Transformation Groups and Shape Space

This chapter investigates the mathematics of the Lie group model of point set shape variability further for general dimensions and determine its expressibility. This is a necessary foundation for any future work on inference techniques in this model.

Let \( \Sigma^k_m \) denote Kendall’s shape space of sets of \( k \) points in \( m \)-dimensional Euclidean space \((k > m)\): this consists of point sets up to equivalence under rotation, scaling and translation. Not all linear transformations on point sets give well-defined transformations of shapes. However, we show that a subgroup of transformations determined by invertible real matrices of size \( k - 1 \) does act on \( \Sigma^k_m \). For \( m > 2 \), this group is maximal, whereas for \( m = 2 \), the maximal group consists of the invertible complex matrices. It is proved that these groups are able to transform any generic shape to any other. Moreover, we establish that for \( k > m + 1 \) this may be done via one-parameter subgroups. Each one-parameter subgroup is given by exponentiation of an arbitrary \((k - 1) \times (k - 1)\) matrix. Shape variabilities may thus be modelled by elements of a \((k - 1)^2\)-dimensional vector space.

12.1 Introduction

Natural ensembles of geometrical objects are often complex. They have both a richness in geometry shape variation and yet a striking similarity within the class. Models of such complex ensembles thus need to capture both the allowed variabilities, as well as the restrictions on variability [16, 7]. Often the shape of an object may be captured by the configuration of a finite number of uniquely identifiable points on the object [8, 40]. The ‘shape’ of such a point set is then understood to be what is not changed under the action of the similarity group of translations, rotations and scalings.

The problem of modelling variabilities of ensembles of objects thus leads to the study of the ‘shape’ of point sets and of transformations between point sets. Linear models are only able to amply describe variabilities between shapes that are sufficiently close and require the computation of a mean configuration. The Lie group model based on linear vector fields describe a wider range of variabilities than linear models.

\footnote{This chapter contains a slightly modified version of an article written by Andrew Swann and Niels Holm Olsen, accepted for publication in Journal of Mathematical Imaging and Vision. It documents joint research performed in the Natural Shape Project, Danish Research Agency Project 9900995. We thank the members of that project for their support and encouragement. Andrew Swann is a member of EDGE, Research Training Network HPRN-CT-2000-00101, supported by The European Human Potential Programme. Thanks are also due to Kristian Kvist and the referees for comments on the manuscript of this article.}
12.1.1 Models of Point Set Shape Variabilities

If one studies ensembles of shapes of point sets, there are two problems one encounters. Firstly the shape spaces of point sets are not vector spaces, but at best curved Riemannian manifolds [42, 28], so the linear deviations of shapes are not defined. Secondly, if one tries to model the underlying point set variabilities, then the linear deviations from a mean point set as described in [16] do not commute with rotations and scalings of the mean point set, and thus do not induce well defined shape variabilities.

In [35], a Lie group model for shape variability was proposed and using linear vector fields it was shown how this model captures the phenomenon of non-linear bending for three-point configurations in two dimensions (Figure 12.6, page 117). Another main motivation for studying Lie group models of variabilities, is that it allows us to express the shape variabilities as originating from point set variabilities in a globally well-defined way.

Models of shape distributions are often phrased either in terms of the original Cartesian coordinates of the point sets, the size- and position-normalised point sets (also known as pre-shapes), directly as a distribution on shape space or on a tangent space at a mean shape.

The Lie group model is a contribution to further understanding of where and how to model shape variability. Because the studied variabilities act both on the original Cartesian point set coordinates, see equation (12.3), and their shapes as points in Kendall’s shape space, they can be used for modelling both point set distributions and shape distributions, and this may be done in a consistent way.

We thus emphasise that the studied models of shape variability are formulated and represented as models of point set variabilities, which induce well-defined shape variabilities. The main results presented show that such variabilities do indeed exist and that they have sufficient generality for use in applications.

12.1.2 Non-Linear Shape Ensembles

Multivariate Gaussian models have proved useful in situations where the ensemble can be described as independent Gaussian distributed orthogonal linear deviations from a mean, see e.g. [16, 17]. The strength of these models is their ability to capture the complexity of the ensemble, by fitting the allowed variabilities based on few examples from the ensemble, while still obeying the ensemble restrictions on variability. However, when faced with ensembles with large variations, the local linear models of the deviations may be inadequate to describe the variabilities accurately.

To illustrate this we consider the example of bending shape variability noted as a problematic case in [16], and further analysed in [35] for the most simple case of bending of three points in the plane (see Figure 12.1). This ensemble is obviously generated by a one-dimensional shape variability, but clearly does not belong to a one-dimensional linear or affine subspace, since the three vertices do not move on straight lines in any of the representations shown in Figure 12.1. When this example set is further analysed by projection on the tangent space to (pre-)shape space (see e.g., [17, p. 77]) at the mean shape, the projection of the 10 examples onto the principal plane spanned by the two first principal directions scatter as illustrated in Figure 12.2. This plot clearly shows that the one-dimensional bending variability spans at least a two-dimensional subspace. In this sense length-ratio-preserving bending is non-linear.

Modelling such non-linear shape variabilities is one of the main motivations for studying the Lie group model for shape variability introduced in [35]. There it was shown that bending of three points in the plane is described exactly by a Lie group of linear point set transformations. Although the presented example of bending of only three points in the plane is an artificial dataset, it illustrates the problem of ensemble curvature and consequent non-linearity which may also be observed in real datasets; see, for example, the Vilman rat dataset of [8, p. 68],
Figure 12.1: 10 samples from an ensemble generated by a bending variability of a 3 point configuration. The data is shown in its original form, and subsequently position-, size-, and orientation-normalised. The orientation-normalisation is relative to the mean shape (shown thickened) given by Procrustes analysis [17, p. 77].

Figure 12.2: Scatter plot for the examples of Figure 12.1 showing the first two principal components of the tangent space projection at the Procrustes mean shape.

which has a similar scatter plot. We thus expect the Lie group models of variability to be instrumental in improving model specificity without sacrificing model variability in such cases.
12.1.3 Linear Transformation Groups and Shape Space

A point set consisting of \( k \) points in Euclidean \( m \)-space, may itself be represented as a vector \( \mathbf{x} \) in a Euclidean space of dimension \( km \). By translation, the shape \( [\mathbf{x}] \) of \( \mathbf{x} \) may be represented by a point set \( \tilde{\mathbf{x}} \) with centre of mass at \( \mathbf{0} \). As the last point in the centred configuration is uniquely determined by the first \( k - 1 \) points, we may represent \( \tilde{\mathbf{x}} \), and hence \( \mathbf{x} \), by a vector \( \mathbf{y} \) in a Euclidean space of dimension \( N = m(k - 1) \), corresponding to the reduced point set.

Given an invertible \((N \times N)\)-matrix \( A \), we have the following linear transformation of reduced point sets:

\[
T_A(\mathbf{y}) = A\mathbf{y}.
\]

However, in general if \( \mathbf{y} \) and \( \mathbf{y}' \) describe the same shape, there is no guarantee that this is still true of \( A\mathbf{y} \) and \( A\mathbf{y}' \). Our first main result is a determination of the \( A \)'s that do give well-defined shape transformations. It turns out that these form a group of dimension \((k - 1)^2\), for \( m > 2 \), or \( 2(k - 1)^2 \), so there are many ‘linear’ shape transformations. For comparison, the set of shapes \([\mathbf{y}]\) has dimension that grows as \( m(k - 1) \), and in typical applications \( k \) is substantially larger than \( m \). Thus one would expect to be able to transform a given shape \([\mathbf{y}]\) in an arbitrary initial direction. We show more than this is true: given any two shapes \([\mathbf{y}]\) and \([\mathbf{y}']\) whose point-configurations do not lie in an \((m - 1)\)-dimensional subspace, there is a shape transformation \( T_A \) with \( T_A[\mathbf{y}] = [\mathbf{y}'] \).

The Lie group model proposes that variations of a shape \([\mathbf{y}]\) may be modelled by one-parameter groups \([\mathbf{y}] \rightarrow \varphi_t[\mathbf{y}] \), \( t \in \mathbb{R} \), i.e., \( \varphi_1 \) is a transformation of shape space, such that \( \varphi_{t+s} = \varphi_t \circ \varphi_s \). Given an \((N \times N)\)-matrix \( B \), one has a one-parameter group

\[
\varphi^B_t \mathbf{y} = \exp(tB)\mathbf{y} = \mathbf{y} + tB\mathbf{y} + \frac{1}{2}t^2B^2\mathbf{y} + \frac{1}{6!}t^3B^3\mathbf{y} + \cdots
\]

This is well-defined on shapes if \( B \) lies in the Lie algebra of the group of shape transformations \( T_A \).

We show that given \([\mathbf{y}]\) and \([\mathbf{y}']\) as above, there is a one-parameter group \( \varphi^B_t \) which takes \([\mathbf{y}]\) to \([\mathbf{y}']\) in time 1, provided \( k > m + 1 \). This result does not follow from the structure of the group of shape transformations.

Note that \( \varphi^B_t \) describes a variability globally, without reference to a fixed mean configuration: given any \([\mathbf{y}]\), \( \varphi^B_t[\mathbf{y}] \) is a variability about \([\mathbf{y}]\). Thus for example, bending of three-point configurations is described by a single \( \varphi^B_t \), regardless of the start configuration. This gives the possibility of comparing variabilities of different mean configurations. Such a comparison is not possible in for example Kent’s tangent plane analysis (as described in [17]), since there is no canonical way to relate tangent vectors at different points. Even if one chooses a connection on the tangent bundle, and uses parallel transport from one tangent space to the other, the result is dependent on the choice of joining path: this is the phenomenon of ‘holonomy’.

12.1.4 Outline

After establishing basic definitions in §12.2, we demonstrate the existence of a large symmetry group in §12.3 and show that this group has an open dense orbit on shape space. In particular, generic shapes may be transformed to generic shapes. Simple example symmetries are presented in §12.4. There is a crucial difference between point sets in two-dimensional space and those in higher dimensions. This is reflected in the symmetry groups, which in general consists of invertible real matrices, but in the two-dimensional case includes all invertible complex matrices. In §12.5 we show that these groups of symmetries are maximal. Having proved this, in §12.6 we compute stabilisers of generic shapes and thereby determine the structure of the open dense orbit.

For the study of variabilities, one wishes to understand the one-parameter groups of symmetries. Unfortunately there are invertible matrices that do not lie in any one-parameter subgroup. However, apart from one borderline case, we demonstrate that one-parameter groups may be
found joining arbitrary points of the open dense orbit of shapes. In the two-dimensional case analyzed in §12.7, there are many such choices, some picked out by isometry invariant models for the tangent space. In higher dimensions, we establish existence of one-parameter groups between given shapes in §12.8.

12.2 Shape Space

Let us begin by defining shape space and fixing our notation. General references for this material include [17, 42].

Consider $k$ points $(x_1, \ldots, x_k)$ in $\mathbb{R}^m$ not all equal, where $x_i$ are column vectors and $k > m$. We take two such point sets to be equivalent if one can be obtained from the other by translation, scaling and rotation. Shape space $\Sigma^k_m$ is then defined to be the set of these equivalence classes.

A common way to make this definition more concrete is to remove the translation and scaling invariance as follows. Let $x^0 = \frac{1}{k} \sum_{i=1}^k x_i$ denote the centre of mass. Then up to translation the point set $(x_1, \ldots, x_k)$ is determined by $(\tilde{x}_1, \ldots, \tilde{x}_k)$, where $\tilde{x}_i = x_i - x^0$. Scaling may now be removed by dividing by the Euclidean norm $\| \sum_{i=1}^k |\tilde{x}_i|^2 \|$. These centred and scaled point sets lie in preshape space $S^k_m$ which is a sphere of dimension $(k-1)m - 1$. Shape space is now the quotient $\Sigma^k_m = S^k_m / SO(m)$ of preshape space under the action of the special orthogonal group $SO(m) = \{ A \in M_m(\mathbb{R}) : A^T A = I, \det A = 1 \}$ given by $\tilde{x}_i \mapsto A\tilde{x}_i$. Here $M_m(\mathbb{R})$ denotes the set of real $m \times m$ matrices.

For $m = 1$, shape space and preshape space are the same, in particular shape space is a sphere. For $m = 2$, we may identify $\mathbb{R}^2$ with $\mathbb{C}$ and get that $\Sigma^2_2 \cong \mathbb{CP}(k-2)$. Here $\mathbb{CP}(n)$ denotes the complex projective space of real dimension $2n$; it consists of the complex one-dimensional subspaces of $\mathbb{C}^{n+1}$. For $m \geq 3$ and $k > m$, things are less pleasant: shape space is a singular manifold of dimension $(k-1)m - \frac{k}{2}m(m-1) - 1$ with singularities in codimension $2(k-m) + 1$.

It is traditional to equip shape space with a metric structure as follows. Preshape space $S^k_m$ is a sphere, so give it the usual round metric. The orbits of the $SO(m)$-action are compact and thus the distance between orbits is well-defined and descends to a metric, the Procrustes metric, on the quotient space $\Sigma^k_m$. Away from the singularities of shape space the Procrustes metric is a smooth Riemannian metric and the fibration $S^k_m \to \Sigma^k_m$ is a Riemannian submersion.

In this metric picture, we map a point set $(x_1, \ldots, x_k)$ to the $k-1$ points $y_1, \ldots, y_{k-1}$ in $\mathbb{R}^m$ by $Y = XH$, where $X$ and $Y$ have columns $x_i$ and $y_j$, respectively, and $H = (h_{ij})$, $i = 1, \ldots, k$, $j = 1, \ldots, k-1$, is the Helmert sub-matrix, defined by

$$h_{ij} = \begin{cases} -1/\sqrt{j(j+1)}, & \text{if } 1 \leq i \leq j, \\ \sqrt{j/(j+1)}, & \text{if } i = j + 1, \\ 0, & \text{if } i > j + 1. \end{cases} \quad (12.1)$$

(The columns of $H$ form a complete orthonormal set orthogonal to $1 = (1,1,\ldots,1)^T$.) The $(k-1)$-tuple $y_1, \ldots, y_{k-1}$ determines the original point set up to translation. The vectors $y_j$ may be used instead of $\tilde{x}_j$ to give the concrete realisation of preshape space and to define shape space itself. We call $(y_1, \ldots, y_{k-1})$ a reduced point set.

12.3 Existence of Symmetries

In this section we provide linear transformations on (reduced) point sets give rise to non-trivial transformation groups acting on shape space. In the next section we look at whether the groups we have found are maximal.
Let $G$ be a group acting on the space of (reduced) point sets. If $G$ is to descend to a well-defined action on shape space $\Sigma^k_m$, then it must send equivalent point sets to equivalent point sets. This is clearly the case if $G$ commutes with a group defining the equivalence of point sets. We thus begin by looking at groups with this property.

In the discussion, the case $m = 2$, i.e., points in $\mathbb{R}^2$, has extra properties which will be discussed after the general case.

Let $f_1, \ldots, f_{k-1}$ be the standard basis for $\mathbb{R}^{k-1}$. Write a reduced point set $(y_1, \ldots, y_{k-1})$ as $y = \sum_{j=1}^{k-1} y_j \otimes f_j \in \mathbb{R}^m \otimes \mathbb{R}^{k-1} \cong \mathbb{R}^{m(k-1)}$. The rotation group $SO(m)$ acts on $y$ by

$$ A \cdot y = \sum_{j=1}^{k-1} Ay_j \otimes f_j. $$

Let $GL(k-1, \mathbb{R})$ denote the group of invertible $(k-1) \times (k-1)$-matrices. This acts on a reduced point set by

$$ B \cdot y = \sum_{j=1}^{k-1} y_j \otimes B f_j = \sum_{j, \ell} b_{j\ell} y_j \otimes f_\ell. $$

Putting $Y = (y_1 y_2 \ldots y_{k-1})$, this action is

$$ B \cdot Y = YB. $$

This may be concretely extended to an action on point sets $X = (x_1 x_2 \ldots x_k)$ commuting with translations, scaling and rotation, as follows. Let $\tilde{H}$ be the full Helmert $k \times k$ matrix, whose first column is $\frac{1}{\sqrt{k}} (1, 1, \ldots, 1)^T$ and whose remaining columns are those of $H$ (equation (12.1)), so $\tilde{h}_{i,j+1} = h_{ij}$. Then

$$ B \cdot X = X \tilde{H} \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} \tilde{H}^T. $$

Proposition 12.3.1 The group $GL(k-1, \mathbb{R})$ acts on shape space $\Sigma^k_m$. Moreover, there is an open dense set $\Omega^k_m$ of shapes which may be transformed into each other by elements of this group. The complement of $\Omega^k_m$ has codimension $k - m$ in $\Sigma^k_m$.

Proof. The action of $GL(k-1, \mathbb{R})$ commutes both with the action of $SO(m)$ and with scaling, so we get a well-defined action on shape space $\Sigma^k_m$.

Let $e_1, \ldots, e_m$ be the standard basis for $\mathbb{R}^m$. Then

$$ y = \sum_{i=1}^{m} e_i \otimes v_i, $$

where $v_i = \sum_{j=1}^{k-1} (y_j) f_j \in \mathbb{R}^{k-1}$ for $(y_j) = \sum_{i=1}^{m} (y_j) e_i$. Let $\Omega^k_m$ denote the set of shapes $[y]$ such that in the representation (12.4) of $y$ the vectors $v_1, \ldots, v_m$ are linearly independent in $\mathbb{R}^{k-1}$. As $k > m$ this set is open and dense in $\Sigma^k_m$. For such a reduced point set, extend $v_1, \ldots, v_m$ to a basis $v_1, \ldots, v_{k-1}$ of $\mathbb{R}^{k-1}$. Then the matrix $V$ whose columns are $v_i, i = 1, \ldots, k-1$, is invertible and

$$ V \cdot \sum_{i=1}^{m} e_i \otimes f_i = y, $$

i.e., $V$ is an element of $GL(k-1, \mathbb{R})$ taking the reduced point set $(e_1, \ldots, e_m, 0, \ldots, 0)$ to $y$.

Now given two shapes $[y], [y'] \in \Omega^k_m$, find $V$ and $V'$ as above. Then $[y'] = W \cdot [y]$, where $W = V'V^{-1} \in GL(k-1, \mathbb{R})$, as required.

112
Figure 12.3: One-variable scaling of a regular heptagon.

Figure 12.4: Two-variable rotation of a regular heptagon.

Note that the proof also shows that $GL(k - 1, \mathbb{R})$ acts with an open dense orbit on the space of reduced point sets. The point sets that lie outside the open orbit are those that lie in an $(m - 1)$-dimensional subspace.

When $m = 2$, we may identify $\mathbb{R}^2$ with $\mathbb{C}$. The reduced point set may now be viewed as a complex vector $w \in \mathbb{C}^{k-1}$. Rotation and scaling are given by multiplying by a non-zero complex number. The group $GL(k - 1, \mathbb{C})$ takes any non-zero reduced point set to any other and so acts transitively on the shape space $\Sigma_k^2 = \mathbb{CP}(k - 2)$.

12.4 Example Symmetries

In Figures 12.3–12.5 we illustrate the effects of three elements of $GL(6, \mathbb{R})$ acting on a regular heptagon centred on the origin. In each case we consider an element of block-diagonal form $\text{diag}(1, 1, A, 1, 1)$, where $A$ is a $2 \times 2$ matrix. We have scaled the elements to lie on the preshape sphere.

In Figure 12.3, $A$ is of the form $\text{diag}(e^t, 1)$. Thus only one coordinate is scaled.
In Figure 12.4, we consider \( A \) to be a rotation matrix
\[
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}.
\]
Figure 12.5 demonstrates the action of a sheer \( A = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \). Thus Figures 12.3–12.5 show the behaviour of three types of one-parameter groups that together provide generators for \( GL(k - 1, \mathbb{R}) \).

In dimension 2, there are further examples provided by complex matrices. We will discuss the case of bending in §12.7.

12.5 Maximality of Symmetry Groups

In §12.3, we saw that the group \( GL(k - 1, \mathbb{R}) \) acts on shape space \( \Sigma^k \) via equations (12.2). However, the space of reduced point sets is simply \( \mathbb{R}^{m(k-1)} \) and one can ask whether there are other linear transformations \( T : \mathbb{R}^{m(k-1)} \to \mathbb{R}^{m(k-1)} \) that induce well-defined transformations on \( \Sigma^k \). The following theorem answers this question in the negative.

**Theorem 12.5.1** For \( m > 2 \), the maximal connected group of linear transformations on reduced point sets commuting with rotations and scaling is the identity component of \( GL(k - 1, \mathbb{R}) \).

**Proof.** Suppose \( G \) is a connected Lie group of linear transformations on the space \( \mathbb{R}^m \otimes \mathbb{R}^{k-1} \cong \mathbb{R}^{m(k-1)} \) of reduced point sets. Then \( G \) is determined by its Lie algebra \( \mathfrak{g} \), which is a linear subspace of the space \( \mathfrak{M}_{m}^{k-1}(\mathbb{R}) \): the group \( G \) is generated by the matrix exponentials of the elements of \( \mathfrak{g} \).

Let \( V \) denote \( \mathbb{R}^m \) with the standard action of \( SO(m) \). The space of \( m \times m \) matrices splits as
\[
M_m(\mathbb{R}) = V \otimes V = \mathbb{R} + \mathfrak{so}(m) + S^2_0 V,
\]
where the first factor consists of multiples of the identity \( Id \), \( \mathfrak{so}(m) = \{ A \in M_m(\mathbb{R}) : A^T = -A \} \) is the Lie algebra of \( SO(m) \) and \( S^2_0 V \) consists of the trace-free symmetric matrices. Let \( R_{ij}, 1 \leq i < j \leq m \), be the basis for \( \mathfrak{so}(m) \) consisting of the skew-symmetric matrices with 1 in the \( (i,j) \) position, \(-1 \) in the \( (j,i) \) place and all other entries 0. Let \( E_1, \ldots, E_\ell \) be any basis for \( S^2_0 V \).

The matrix of an arbitrary element of \( M_{m(k-1)}(\mathbb{R}) \) may be written as
\[
A = Id \otimes A_0 + \sum_{1 \leq i < j \leq m} R_{ij} \otimes A_{ij} + \sum_{i=1}^{\ell} E_i \otimes A_i,
\]

114
for some \((k - 1) \times (k - 1)\) matrices \(A_0, A_{ij}\) and \(A_i\).

The condition that \(G\) commutes with rotations translates into the requirement that the elements \(A \in g\) commute with the elements of \(so(m)\). Write \([B, C] = BC - CB\) for the usual Lie bracket on matrices and note that matrix product in \(M_{m(k-1)}(\mathbb{R}) \cong M_m(\mathbb{R}) \otimes M_{k-1}(\mathbb{R})\) is given by \((a \otimes b)(c \otimes d) = ac \otimes bd\). The image of \(so(m)\) in \(M_{m(k-1)}(\mathbb{R})\) is generated by the matrices \(R_{ij} \otimes \text{Id}\) and

\[ [R_{ij} \otimes \text{Id}, a \otimes b] = R_{ij}a \otimes b - aR_{ij} \otimes b = [R_{ij}, a] \otimes b. \]

In particular, for \(c \in so(m)\), the action \([c \otimes \text{Id}, \cdot]\) on \(M_{m(k-1)}(\mathbb{R})\) preserves the type of the first factors from the splitting (12.5). Writing \(A\) as in (12.6) we have

\[ [R_{\alpha\beta} \otimes \text{Id}, A] = 0 + \sum_{i,j} [R_{\alpha\beta}, R_{ij}] \otimes A_{ij} + \sum_i [R_{\alpha\beta}, E_i] \otimes A_i. \]

Up to sign \([R_{\alpha\beta}, R_{ij}]\) is either zero or another \(R\)-matrix: the matrices \(R_{ai}\) and \(R_{ji}\) occur once each for every \(i \neq \alpha, \beta\). The condition \([R_{\alpha\beta} \otimes \text{Id}, A] = 0\) for all \(\alpha, \beta\) therefore implies that each \(A_{ij}\) is zero. The action of \([R_{\alpha\beta}, \cdot]\) on \(S_0^2V\) has complex eigenvalues \(\pm 2i, \pm i, 0\) and the intersection of the zero eigenspaces is a subspace of \(S_0^2V\) invariant under the action of \(SO(m)\), so is trivial. Therefore each \(A_i\) is also zero. Thus \(A = \text{Id} \otimes A_0\), which is exactly the Lie algebra of \(GL(k-1, \mathbb{R})\).

**Remark.** In the case \(m = 2\) the above proof changes at the point where we look at the action of \(R_{\alpha\beta}\) on the now one-dimensional \(so(2)\): this action is trivial and we conclude that \(g\) is the vector space sum of two copies of \(gl(k-1, \mathbb{R})\). This gives that \(G = GL(k-1, \mathbb{C})\) when we identify \(\mathbb{R}^2\) with \(\mathbb{C}\).

Up to now we have only looked at groups \(G\) commuting with the groups defining equivalence of (reduced) point sets. However, this is a stronger condition than necessary for a \(G\)-action to descend to shape space. The minimal condition is that \(G\) should send equivalence classes to equivalence classes. For reduced point sets this is the same as saying that \(G\) normalises the group of rotations and scalings, i.e., for each \(g \in G\), \(\alpha > 0\) and \(A \in SO(m)\) there is an \(\alpha' > 0\) and a \(A' \in SO(m)\) such that \(g \cdot (\alpha A) = \alpha' A' \cdot g\).

**Corollary 12.5.2** For \(m > 2\), the maximal connected group of linear transformations on reduced point sets which induces an action on shape space is \(SO(m) \times GL(k-1, \mathbb{R})\).

For \(m = 2\), the corresponding group is \(GL(k-1, \mathbb{C})\).

**Proof.** We have the decomposition (12.5) of \(M_m(\mathbb{R})\) under the action of \(SO(m)\) and for \(M_{k-1}(\mathbb{R})\) we have

\[ M_{k-1}(\mathbb{R}) = gl(k-1, \mathbb{R}) = \mathbb{R} + sl(k-1, \mathbb{R}), \]

where \(\mathbb{R}\) is spanned by the identity matrix and \(sl(k-1, \mathbb{R})\) consists of the trace-free matrices. Now

\[
M_{m(k-1)}(\mathbb{R}) \\
= M_m(\mathbb{R}) \otimes M_{k-1}(\mathbb{R}) \\
= (\mathbb{R} + so(m) + S_0^2V) \otimes (\mathbb{R} + sl(k-1, \mathbb{R})) \\
= \mathbb{R} + so(m) + sl(k-1, \mathbb{R}) \\
+ \left(S_0^2V + so(m) \otimes sl(k-1, \mathbb{R}) + S_0^2V \otimes sl(k-1, \mathbb{R})\right) \\
= so(m) + gl(k-1, \mathbb{R}) + m.
\]
The Lie algebra of the group of scalings and rotations is $\mathbb{R} + \mathfrak{so}(m)$. But $\mathbb{R}$ consists of multiples of the identity matrix and so commutes with all elements of $M_{m(k-1)}$. Thus the normaliser of $\mathbb{R} + \mathfrak{so}(m)$ consists of those $A$ such that $[A, \mathfrak{so}(m)] \subseteq \mathfrak{so}(m)$. Clearly this normaliser contains $gl(k-1, \mathbb{R}) + \mathfrak{so}(m)$. Now $[\mathfrak{so}(m), \mathfrak{m}] \subseteq \mathfrak{m}$, so the only elements of $\mathfrak{m}$ that can contribute to the normaliser are those that commute with $\mathfrak{so}(m)$; however we saw above, that for $m > 2$, there are no such elements.

In the case $m = 2$, the centraliser $GL(k-1, \mathbb{C})$ already contains the scalings and rotations, and the above argument shows that there are no other elements in the normaliser. □

This result shows that the centralisers found in Theorem 12.5.1 and Remark 12.5 are the interesting groups, since the extra elements we pick up in the normaliser lie in $SO(m)$ which acts trivially on shape space. Note that the diagonal matrices in $GL(k-1)$ also act trivially on shape space $\Sigma^k_m$ and so it is sufficient to only consider the subgroup $SL(k-1)$ consisting of matrices of determinant 1.

### 12.6 Shape Stabilisers

We have now established the action of a group $G$ on shape space $\Sigma^k_m$. To understand the action better and the associated geometry it is useful to determine which elements of $G$ preserve a given shape $[y] \in \Sigma^k_m$.

Let $H = \{ g \in G : g \cdot [y] = [y] \}$ be the stabiliser of $[y]$. This is a subgroup of $G$ and for two shapes $[y]$ and $[y']$ in the same orbit of $G$, the stabilisers are conjugate in $G$. It is therefore enough to calculate $H$ for one point $[y]$ of the open orbit $\Omega^k_m$ of $G$. As $G$ acts smoothly and transitively on $\Omega^k_m$, we will then have $\Omega^k_m = G/H$ as smooth manifolds [10, §IV.9].

Let us begin with the case $G = GL(k-1, \mathbb{R})$ and take $y = \sum_{i=1}^m e_i \otimes f_i$. The action of $SO(m)$ on $y$ is given by

$$A \cdot y = \sum_{i=1}^m A e_i \otimes f_i = \sum_{i,j=1}^m a_{ij} e_j \otimes f_i = \sum_{j=1}^m e_j \otimes A^T f_j,$$

where we regard $f_1, \ldots, f_m$ as elements of $\mathbb{R}^m \subseteq \mathbb{R}^{k-1}$. Writing an element $B \in GL(k-1, \mathbb{R})$ as

$$B = \begin{pmatrix} X & W^T \\ V & Y \end{pmatrix},$$

with $X \in M_m(\mathbb{R})$, $V, W \in M_{k-m-1,m}(\mathbb{R})$ and $Y \in M_{k-m-1}(\mathbb{R})$, we have

$$B \cdot y = \sum_{i=1}^m e_i \otimes \begin{pmatrix} Xf_i \\ Vf_i \end{pmatrix}.$$  

Thus $[B \cdot y] = [y]$ if and only if $V = 0$ and $X = \alpha A^T$ for some $\alpha > 0$ and $A \in SO(m)$. As $B$ is invertible, the condition $V = 0$ implies that $Y$ is invertible too. Therefore the isotropy group $H$ of $[\sum_{i=1}^m e_i \otimes f_i]$ is

$$\{ \begin{pmatrix} X & W^T \\ 0 & Y \end{pmatrix} : X \in SO(m) \times \mathbb{R}, W \in M_{k-m-1,m}(\mathbb{R}), Y \in GL(k-m-1, \mathbb{R}) \}. \quad (12.7)$$

It is now easy to check that $\dim G - \dim H = \dim \Sigma^k_m$ and thus confirm that $[\sum_{i=1}^m e_i \otimes f_i]$ lies in the open orbit.

For $m = 2$, we have the case $G = GL(k-1, \mathbb{C})$. Taking $y = w = (1, 0, \ldots, 0)^T \in \mathbb{C}^{k-1}$, we see that the stabiliser is

$$\{ \begin{pmatrix} z & u^T \\ 0 & B \end{pmatrix} : z \in \mathbb{C} \setminus \{0\}, u \in \mathbb{C}^{k-2}, B \in GL(k-2, \mathbb{C}) \}. \quad (12.8)$$  

116
12.7 One-Parameter Groups and Geodesics

We have described an open dense set $\Omega_{\mathcal{S}}$ of shape space as $G/H$ for a Lie group $G$ and a closed subgroup $H$, which is the stabiliser of a shape $[y]$. As $H$ is a closed subgroup of $G$, the Lie algebra $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$ and the tangent space at $[y]$ is given by the quotient vector space

$$T_{[y]}\Sigma^k_m = \mathfrak{g}/\mathfrak{h}.$$ 

An element $A$ of $\mathfrak{g}$ generates a one-parameter subgroup on $\Sigma^k_m$ by $[y'] \mapsto \exp(tA) \cdot [y']$. If $A$ lies in $\mathfrak{h}$ then the corresponding one-parameter subgroup fixes the shape $[y]$.

**Example.** One case studied in [35], is that of ‘length-ratio-preserving’ bending. Here one considers the shape space of triangles in the plane and look for a one-parameter group of $\text{SL}(2, \mathbb{C})$ that preserves the length ratio of two sides of the triangle, whilst allowing the angle between those sides to vary, Figure 12.6. Such an element has two fixed-points given by triangles where one of the two given sides has zero length. When centred, these triangles are $(-2x, x, x)$ and $(x, x, -2x)$ which are the shapes $[\sqrt{3}, 1]$ and $[0, 1]$ in $\text{CP}(1) = \Sigma^2_2$. Such a one-parameter group is thus generated by a complex $2 \times 2$ matrix $A$ with eigenvectors $(\sqrt{3}, 1)^T$ and $(0, 1)^T$. As $A$ lies in $\mathfrak{sl}(2, \mathbb{C})$, it has trace zero, and the two eigenvalues are equal and opposite. To obtain a periodic element, this eigenvalue must be pure imaginary. Thus up to scale $A$ is

$$\begin{pmatrix} \sqrt{3} & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} i & 0 \\ \frac{2i}{\sqrt{3}} & -i \end{pmatrix}$$

and the corresponding one-parameter group is

$$t \mapsto \begin{pmatrix} e^{it} & 0 \\ \frac{2i}{\sqrt{3}} \sin(t) & e^{-it} \end{pmatrix}.$$ 

If one changes the eigenvalues in this example to $\pm 1$ instead of $\pm i$, then one obtains a one-parameter group that alters the length ratio of the two sides of the triangle whilst preserving the angle between them.

It is important to emphasise that bending for all triangles is given by the same one-parameter group. The derivation above shows that this group comes from linear transformations, but that
the matrices involved are not unitary and so the transformation does not preserve the Procrustes metric of shape space. Bending of arbitrary configurations in two-dimensions at an arbitrary vertex may be similarly described [45].

Models for the quotient $\mathfrak{g} / \mathfrak{h}$ may be given by choosing a complement $\mathfrak{m}$ so that

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$$

as vector spaces. Different choices for $\mathfrak{m}$ pick out different families of one-parameter subgroups, i.e., different types of variability of the initial shape $[y]$. The map $A \mapsto \exp(A) \cdot [y]$ is smooth and its differential has kernel $\mathfrak{h}$. So for each choice of $\mathfrak{m}$ the restriction of $\exp$ to $\mathfrak{m}$ is a diffeomorphism from a neighbourhood of $0 \in \mathfrak{m}$ to a neighbourhood of $[y]$. If $[y'] = g \cdot [y]$ then we get a complement at $[y']$ as $g \mathfrak{m} g^{-1}$. For a given $[y']$, this depends on the choice of $g \in G$ unless $h \mathfrak{m} h^{-1} = \mathfrak{m}$ for all $h \in H$. When $H$ is connected, this latter condition is equivalent to $[\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}$, in which case the decomposition is said to be reductive. Unfortunately, the actions of $GL(k-1, \mathbb{R})$ and $GL(k-1, \mathbb{C})$ on shape space do not have reductive complements.

For $m = 2$, there are particular choices of $\mathfrak{m}$ that have special interest. We first discuss these at the base point $[w] = [(1, 0, \ldots, 0)']$.

### 12.7.1 The Procrustes Complement

Recall that the Procrustes metric comes from regarding preshape space as a round sphere. The elements of $GL(k-1, \mathbb{C})$ that preserve preshape space are precisely those complex matrices that preserve the standard inner product on $\mathbb{C}^{k-1}$. This is the unitary group $U(k-1)$, which has Lie algebra

$$u(k-1) = \left\{ A \in M_{k-1}(\mathbb{C}) : A^T = -A \right\}.$$  

We define the Procrustes complement to be

$$\mathfrak{m}_p = \left\{ P(v) = \begin{pmatrix} 0 & -v^T \\ v & 0 \end{pmatrix} : v \in \mathbb{C}^{k-2} \right\}.$$ 

The elements of $\mathfrak{m}_p$ lie in $u(k-1)$ and so their exponentials preserve the Procrustes metric.

In fact more is true. The stabiliser $H_1$ of $[w]$ under the action of $U(k-1)$ is simply the intersection of (12.8) with $U(k-1)$. We have

$$u(k-1) = \mathfrak{h}_1 + \mathfrak{m}_p,$$

with $[\mathfrak{h}_1, \mathfrak{m}_p] \subset \mathfrak{m}_p$ and $[\mathfrak{m}_p, \mathfrak{m}_p] \subset \mathfrak{h}_1$. This means that $U(k-1)/H_1$ is a symmetric space and has the following important consequence.

**Theorem 12.7.1** The paths $\exp(tA) \cdot [w]$ for $A \in \mathfrak{m}_p$ are exactly the geodesics emanating from $[w]$. □

See Chapter XI, Theorem 3.2 in [32] for a proof of this general result. However, the natural metric on the symmetric space is complete, and so there are geodesics joining any two given points and we have that $A \mapsto \exp(A) \cdot [w]$ is a surjection $\mathfrak{m}_p \to \Sigma_k$.

Explicitly, for $v$ a unit vector in $\mathbb{C}^{k-2}$, we have

$$\exp(tP(v)) \cdot [w] = [\cos t \, w + \sin t \, \tilde{v}],$$

(12.9)

where $\tilde{v}^T = (0 \, v^T)$. Note that the Procrustes distance from $[w]$ to $\exp(tA) \cdot [w]$ is $t$, so this map has a certain distance preserving property, which can be viewed as the basis for Kent’s tangent plane analysis (cf. [17]). The Procrustes complement at an arbitrary point $[y]$ in shape space is given by $g \mathfrak{m}_p g^{-1}$, where $g \in U(k-1)$ is such that $g \cdot [w] = [y]$. Note that $g \mathfrak{m}_p g^{-1} \subset u(k-1)$, so linear averaging techniques on Procrustes generators can not yield anything other than isometries of shape space. In particular, non-linear bending can never be obtained this way.
12.7.2 The Killing Complement

The most obvious choice of $\mathfrak{m}$ from the description of $H$ in (12.8) is

$$\mathfrak{m}_k = \left\{ \begin{bmatrix} K(v) = \begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix} : v \in \mathbb{C}^{k-2} \right\}.$$  

The matrices $A \in \mathfrak{m}_k$ have $A^2 = 0$, so are nilpotent. In this case

$$\exp(tA) = \text{Id} + tA$$

defines a linear deformation of the initial shape, however on shape space this becomes

$$\exp(tK(v)) \cdot [w] = [w + t\hat{v}] = [\cos \theta \, w + \sin \theta \, \hat{v}],$$

where $\tan \theta = t$. This is a reparameterisation of the part of the Procrustes path (12.9) for $t \in (-\pi/2, \pi/2)$. The map $A \rightarrow \exp(A) \cdot [w]$ from $\mathfrak{m}_k$ to $\Sigma_k^2$ is not surjective: the image consists of points $[y]$ so that $w^T y \neq 0$, which is $\Sigma_k^2 \setminus \Sigma_k^{k-1}$.

The complement $\mathfrak{m}_k$ consists of those $A$ that are Killing orthogonal to all elements $B$ in $\mathfrak{h}$, meaning $\text{Tr}(A^T B) = 0$ for all such $B$. In order to get a well-defined complement at arbitrary points $[y]$ of $\Sigma_k^2$, we consider $g \mathfrak{m}_k \, g^{-1}$ with $g \in U(k-1)$. This complement is again given by the condition that $\text{Tr}(A^T B) = 0$ for all $B$ in the Lie algebra of the stabiliser of $[y]$.

Explicitly, suppose $[y_1]$ and $[y_2]$ are two shapes, which we choose to be represented by reduced point sets $y_1$ and $y_2$ of unit norm. Let $\Lambda = \langle y_2, y_1 \rangle$ and put $w = \frac{1}{2} y_2 - y_1$. Then $A = w \otimes y^T$ is the unique element in the Killing complement at $[y_1]$ with $\exp(A) \cdot [y_1] = [y_2]$.

Note that $\mathfrak{m}_k$ contains the element $f_2 \otimes f_1^T$. If $v_1$ and $v_2$ are mutually orthogonal unit vectors in $\mathbb{C}^{k-1}$, then we can find a unitary matrix $g$ whose first column is $v_1$ and whose second column is $v_2$. We then have that $g \mathfrak{m}_k \, g^{-1}$ contains $v_2 \otimes v_1^T$ and can deduce that the subspaces $g \mathfrak{m}_k \, g^{-1}$ generate $\mathfrak{sl}(k-1, \mathbb{C})$ linearly. In other words, linear averaging techniques on Killing generators do have a chance of giving non-linear bending.

12.7.3 General Complements

The most general complement invariant under $U(k-1)$ is determined by a single complex number $\sigma$ and is given by

$$\mathfrak{m}_\sigma = \left\{ \begin{bmatrix} 0 & -\sigma v^T \\ v & 0 \end{bmatrix} : v \in \mathbb{C}^{k-2} \right\}.$$  

This follows from the fact that $\mathfrak{m}_\sigma$ should be a linear subspace of dimension $k-2$, complementary to the Lie algebra $\mathfrak{h}_1$ and invariant under the action of $H_1$. Thus $\mathfrak{m}_\sigma$ defines a linear map $\mathbb{C}^{k-2} \rightarrow M_{k-1}(\mathbb{C})$ and that linear map is equivariant under the action of $H_1 = U(1) \times U(k-2)$. Decomposing $M_{k-1}(\mathbb{C})$ under the action of $H_1$ and using Schur’s Lemma one finds that $\mathfrak{m}_\sigma$ has the claimed form.

12.8 Existence of One-Parameter Groups

As the open orbit $\Omega^k_m$ of $GL(k-1, \mathbb{R})$ is not compact, it is not obvious that one can join two given shapes in $\Omega^k_m$ by a one-parameter group generated by $GL(k-1, \mathbb{R})$. Indeed there are two potential problems.

Firstly, $\Omega^k_m$ is only connected if $k > m + 1$. For $k = m + 1$, the orbit $\Omega^k_m$ has two connected components, call these $\Omega^m_{m+1}(\pm)$.  

119
Secondly, consider the exponential map \( \exp: M_{k-1}(\mathbb{R}) \rightarrow GL_4(k-1,\mathbb{R}) \), where \( GL_4 \) denotes the elements in \( GL \) with positive determinant. This map is not surjective for \( k \geq 3 \), cf. pages 24 and 30 in [14]. Indeed, it is instructive to look at the case \( k = 3 \) in some detail. Any element \( A \in M_2(\mathbb{R}) \) may be written as \( A = \mu \text{Id} + A_0 \), where \( \text{Tr} A_0 = 0 \). Now either (i) \( A_0 \) is diagonalisable over \( \mathbb{R} \) with eigenvalues \( \pm \nu \), or (ii) \( A_0 \) is diagonalisable over \( \mathbb{C} \) with pure imaginary eigenvalues \( \pm i\nu \) or (iii) \( A_0 \) is not diagonalisable, in which case \( A_0 \) has rank one and has zero as its only eigenvalue. Let \( B = \exp(A) = \exp(\mu) \exp(A_0) \). Then in case (i), \( B \) is diagonalisable over \( \mathbb{R} \) with eigenvalues \( e^{\mu \pm i\nu} > 0 \). For case (ii), \( B \) is diagonalisable over \( \mathbb{C} \), with eigenvalues \( e^{\mu \pm i\nu} \). Note, that here \( B \) can have real eigenvalues, but only for \( \nu = n\pi \), in which case both eigenvalues are equal. For case (iii), \( B \) is not diagonalisable, but has a single eigenvalue \( e^\mu > 0 \). Thus the image of \( \exp \) is all of \( GL_4(2,\mathbb{R}) \) except for those matrices with distinct negative eigenvalues and those non-diagonalisable matrices whose single eigenvalue is negative.

It turns out that on \( \Omega_{m}^k \) only the first of these problems is inescapable.

**Theorem 12.8.1** Let \( \Omega_{m}^k \subset \Sigma_{m}^k \) be the open orbit of \( GL(k-1,\mathbb{R}) \).

If \( k > m+1 \), then any two shapes in \( \Omega_{m}^k \) may be transformed into each other by some one-parameter subgroup of \( GL(k-1,\mathbb{R}) \).

For \( k = m+1 \), the open orbit \( \Omega_{m}^{m+1} \) decomposes into connected components as \( \Omega_{m}^{m+1} = \Omega_{m}^{m+1}(+) \cup \Omega_{m}^{m+1}(-) \) and two shapes are related by a one-parameter group if and only if they lie in the same connected component.

**Proof.** Let \([y]\) and \([y']\) be two shapes in \( \Omega_{m}^k \). Using an element of \( GL(k-1,\mathbb{R}) \) we may assume that \( y \) is the standard element \( \sum_{i=1}^{m} e_i \otimes f_i \). Write \( y' = \sum_{i=1}^{m} e_i \otimes v_i \) and let \( V \) be the \((k-1) \times m\) matrix with columns \( v_i \). Write \( V = \begin{pmatrix} X & Y \end{pmatrix} \), with \( X \in M_m(\mathbb{R}) \).

The matrix \( X \) has an oriented singular-value decomposition \( X = CDB^T \), meaning that \( B, C \in SO(m) \) and that \( D = \text{diag} (\lambda_1, \ldots, \lambda_m) \) with
\[
\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{m-1} \geq |\lambda_m| \geq 0.
\]
(This may be easily obtained by modifying the usual singular-value decomposition.) Now \( \tilde{V} = VBC^T = \begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix} \) is obtained from \( V \) by a rotation in \( SO(m) \) so represents the same shape as \( y' \).

Moreover, we have \( \tilde{X} = CDC^T \), so \( \tilde{X} \) is diagonalisable.

If \( k = m+1 \), then \( \tilde{Y} \) is not present and \( \tilde{X} \) is in the image of the exponential map only if \( \lambda_m > 0 \). This is exactly the condition that \([y]\) and \([y']\) lie in the same connected component of \( \Omega_{m}^k \).

Now take \( k > m+1 \) and consider the \((k-1) \times (k-1)\) matrix
\[
F(\alpha, \beta) = \begin{pmatrix} \tilde{X} & \alpha \\ \tilde{Y} & \beta \end{pmatrix}
\]
with \( \alpha \in M_{m,k-m-1}(\mathbb{R}) \) and \( \beta \in M_{k-m-1}(\mathbb{R}) \). If we can write \( F(\alpha, \beta) = \exp(A) \), then we will have that \( \exp(A) \cdot [y] = [y'] \), as required. Note that as \( \exp(KAK^{-1}) = K \exp(A)K^{-1} \), it is enough to find \( \alpha, \beta \) so that \( F(\alpha, \beta) \) is similar to a matrix in the image of the exponential map.

First suppose that \( \lambda_{m-1} > \lambda_m \) and that the last column of \( \tilde{Y}C = YB \) is non-zero. Then we can find an invertible matrix \( G \in M_{k-m-1}(\mathbb{R}) \) such that \( G^{-1}YB \) has last column \((1,0,\ldots,0)^T \).

Now
\[
\begin{pmatrix} C^T & 0 \\ 0 & G^{-1} \end{pmatrix} F(\alpha, \beta) \begin{pmatrix} C & 0 \\ 0 & G \end{pmatrix} = \begin{pmatrix} D & C^T \alpha G \\ G^{-1}YB & G^{-1}G \end{pmatrix}.
\]

(12.10)
Choose \( \alpha \) such that \( C^T \alpha G \) has only one non-zero entry \( a \) in its lower-left corner. Choose \( \beta = G \text{diag}(b, \lambda, \ldots, \lambda)G^{-1} \). Then (12.10) is nearly lower-triangular, with a central \( 2 \times 2 \) block

\[
\begin{pmatrix}
\lambda_m & a \\
1 & b
\end{pmatrix}.
\]

This block has eigenvalues \( \mu_\pm = \frac{1}{2}(\lambda_m + b \pm \sqrt{(\lambda_m - b)^2 + 4a}) \). We may thus choose \( a \) and \( b \) so that these two eigenvalues are distinct positive real numbers different from each of \( \lambda_1, \ldots, \lambda_{m-1} \). Now choosing \( \lambda > 0 \) to be distinct from \( \mu_\pm \) and \( \lambda_1, \ldots, \lambda_{m-1} \), we obtain an \( F(\alpha, \beta) \) which is diagonalisable with positive real eigenvalues. This \( F(\alpha, \beta) \) is therefore in the image of the exponential map.

If the last column of \( YB \) is zero, then \( \lambda_m \) is non-zero. We may then take \( G = \text{Id}, a = 0, b = \lambda_m \) and \( \lambda > 0 \) distinct from \( \lambda_1, \ldots, \lambda_m \). The discussion of the exponential map on \( 2 \times 2 \) matrices above shows that (12.10), and hence \( F(\alpha, \beta) \), lie in the image of the exponential map.

Finally, suppose \( \lambda_r = \cdots = \lambda_m \) for some \( r < m \). If \( \lambda_r > 0 \), then \( F(0, \lambda_r, \text{Id}) \) is similar to the direct sum of a diagonalisable matrix with positive eigenvalues \( \lambda_1, \ldots, \lambda_{r-1} \) and a lower triangular matrix with the single eigenvalue \( \lambda_r > 0 \). Thus \( F(0, \lambda_r, \text{Id}) \) is in the image of the exponential map.

If \( \lambda_r = \cdots = \lambda_m = 0 \) for some \( r < m \), then in (12.10) consider \( G = \text{Id} \) and \( \beta = 0 \). Write the right-hand side as

\[
\begin{pmatrix}
D' & 0 & \alpha' \\
0 & 0_{m-r+1} & \alpha_r \\
Y' & Y_r & 0
\end{pmatrix},
\]

with \( D' = \text{diag}(\lambda_1, \ldots, \lambda_{r-1}) \) and \( 0_{m-r} \) the zero matrix in \( M_{m-r}(\mathbb{R}) \). Choose \( \alpha \) so that \( \alpha' = 0 \) and \( \alpha_r = -Y_r^T \). Since \( V \) has rank \( m \), the matrix \( Y_r \) has singular value decomposition \( Y_r = C_rD_rB_r^T \), with \( C_r \) and \( B_r \) orthogonal matrices and \( D_r \) an \( k - m - 1 \) by \( m - r - 1 \) diagonal matrix with diagonal entries \( \mu_r, \ldots, \mu_m > 0 \). We now have that

\[
\begin{pmatrix}
0 & -Y_r^T \\
Y_r & 0
\end{pmatrix} = \begin{pmatrix}
B_r & 0 \\
0 & C_r
\end{pmatrix} \begin{pmatrix}
0 & -D_r^T \\
D_r & 0
\end{pmatrix} \begin{pmatrix}
B_r & 0 \\
0 & C_r
\end{pmatrix}^T.
\]

But the middle matrix on the right is a direct sum of \( 2 \times 2 \) blocks \( \begin{pmatrix} 0 & -\mu_j \\ \mu_j & 0 \end{pmatrix} \) and is thus in the image of the exponential map. As the block matrix (12.12) has pure imaginary eigenvalues \( \pm i\mu_j \), it has no eigenvalues in common with \( D' \). So by the Primary Decomposition Theorem, the matrix (12.11) is similar to a matrix of the same form with \( Y' = 0 \). Thus (12.11) is in the image of the exponential map and we have shown that for any \( [y'] \in \Omega_m^k \) there is an \( A \in GL(k - 1, \mathbb{R}) \) such that \( \exp(A) \cdot [y] = [y'] \), when \( k > m + 1 \).

The above proof describes an algorithm for constructing a one-parameter group joining two shapes in \( \Omega_m^k \). In Figure 12.7, we demonstrate such a transformation. The shapes are in \( \Sigma_{100}^k \), given by taking 100 sample points on two smooth curves: the first \( [y] \) a spiral lying on the surface of a torus, the last \( [y'] \) a helical curve lying on the surface of a sphere. The pictures show \( \exp(tA) \cdot [y] \) for various values of \( t \), where \( A \in M_{99}(\mathbb{R}) \) has \( \exp(A) \cdot [y] = [y'] \). An application where such transformations may well be of use, is the study of folding of protein chains.

We remark that for \( m > 2 \) there is no analogue of the invariant complements discussed in §12.7. Indeed the subgroup of \( GL(k - 1, \mathbb{R}) \) preserving the Procrustes metric is the orthogonal group \( O(k - 1) \), which being compact can not act transitively on the non-compact open orbit \( \Omega_m^k \). In fact, the action of \( O(k - 1) \) on \( \Sigma_m^k \) is of cohomogeneity \( m - 1 \), i.e., the dimension of a generic \( O(k - 1) \)-orbit is \( m - 1 \) less than the dimension of shape space. It is thus interesting to develop techniques that find preferred one-parameter subgroups joining given shapes.
Figure 12.7: From a spiral on a torus to a helix on a sphere via a one-parameter group found using 100 sample points on each curve.
12.9 Conclusions and Future Applications

We have analysed the expressive power of the simple class of linear point set transformations for the modelling of shape variability. Of the presented results we highlight, for \( k \) points in \( m \)-dimensional Euclidean space \((k > m)\):

- The subgroup \( G = GL(k - 1, \mathbb{R}) \) acts on Kendall’s shape space, and transitively on an open and dense orbit. Thus any generic shape may be transformed into any generic shape. For \( m = 2 \), the same is true of \( G = GL(k - 1, \mathbb{C}) \).
- Moreover for \( k > m + 1 \), any generic shape may be continuously transformed into any other generic shape by some one-parameter subgroup of \( GL(k - 1, \mathbb{R}) \). Continuous shape variabilities may thus be modelled by elements of \( \mathfrak{gl}(k - 1, \mathbb{R}) \) via the exponential map, and are capable of connecting practically all of Kendall’s shape space.

One application of these results is to analysis of data collected via video sequences. A typical sample will consist of shapes \([y_0], [y_1], \ldots, [y_n]\) measured at \( t = 0, 1, \ldots, n \). An approximate one-parameter group of linear transformations may be found by an optimization process for an approximate solution \( A \in GL(k - 1) \) to the equations

\[
A \cdot [y_i] = [y_{i+1}], \quad i = 0, \ldots, n - 1.
\]

Note that this does not require determination of a mean shape. Writing \( A = \exp(B) \), with \( B \) suitably minimised, gives a one-parameter group \( \exp(tB) \), taking \([y_0]\) to \([y_i]\) at time \( t = i \).

Preliminary work on video sequences of moving sperm cells shows this approach is very promising, even if one replaces the above system of equations by the corresponding linear equations on pre-shapes

\[
A \cdot y_i = y_{i+1}
\]

and then takes the least squares solution for \( A \). One can then use usual linear statistical analysis on the matrices \( B \), to compare movement characteristics of different sperm cells from different video sequences. This will be pursued in future work. Similar techniques will also be tried on the Vilman rat dataset of [8, p. 68].

Future work will also consider probabilistic models of non-linear shape ensembles expressed by probability distributions on \( \mathfrak{gl}(k - 1, \mathbb{R}) \). We believe, that the exact model structure should be chosen depending on the application, and with a possible choice of inference from examples in consideration. But we note that the vector space structure of \( \mathfrak{gl}(k - 1, \mathbb{R}) \) allows for the use of the well known Gaussian model with mean 0 and acting on a “mean shape”.
Part V

References
Bibliography


Part VI

Appendices
Appendix A

Manually Defined True Cells

In this appendix we present figures displaying a circular approximation to the equator-contour of each of the true cells manually defined by Søren Ziebe. These figures were generated from the center of mass and mean diameter of the manually defined outline.
Sphere 1 in 200100018_00_02_10.jpg, z = −5.0 um

Sphere 5 in 200100018_00_02_13.jpg, z = −20.0 um

Sphere 1 in 200100019_00_02_03.jpg, z = 30.0 um

Sphere 2 in 200100019_00_02_08.jpg, z = 5.0 um

Sphere 3 in 200100019_00_02_08.jpg, z = 5.0 um

Sphere 4 in 200100019_00_02_13.jpg, z = −20.0 um

Sphere 1 in 200100022_00_01_15.jpg, z = −30.0 um

Sphere 2 in 200100022_00_01_09.jpg, z = 0.0 um

Sphere 3 in 200100022_00_01_09.jpg, z = 0.0 um

Sphere 4 in 200100022_00_01_09.jpg, z = 0.0 um

Sphere 1 in 200100023_00_01_08.jpg, z = 5.0 um

Sphere 2 in 200100023_00_01_07.jpg, z = 10.0 um

Sphere 3 in 200100023_00_01_07.jpg, z = 10.0 um

Sphere 4 in 200100023_00_01_07.jpg, z = 10.0 um

Sphere 1 in 200100024_00_02_08.jpg, z = 5.0 um

Sphere 2 in 200100024_00_02_08.jpg, z = 5.0 um

Sphere 3 in 200100024_00_02_08.jpg, z = 5.0 um

Sphere 4 in 200100024_00_02_08.jpg, z = 5.0 um
There were no true cell annotations for this sequence.
Appendix B

Evaluation I - Automated Cell Detections

In this appendix we present figures displaying the equator-contour of each of the detected cells (accepted in green and rejected in red) in the image which is the equator plane of the cell.
None accepted
None rejected
Appendix C

1000 Strongest Detections

In this appendix we present figures displaying the equator-contour of each of the 1000 strongest raw cell detections. The last figure is a plot of the absolute value of the filtered correlation values. These figures were investigated interactively using a zoom facility, in order to check if the threshold heuristic was the cause of false negative detections. Here we present only the figures for the example sequence 2 (200100013), in order to save paper and ink.
61 suggested cells in 200100013_00_01_01.jpg, z = 40.0 um

177 suggested cells in 200100013_00_01_02.jpg, z = 35.0 um

90 suggested cells in 200100013_00_01_03.jpg, z = 30.0 um

91 suggested cells in 200100013_00_01_04.jpg, z = 25.0 um

117 suggested cells in 200100013_00_01_05.jpg, z = 20.0 um

91 suggested cells in 200100013_00_01_06.jpg, z = 15.0 um

106 suggested cells in 200100013_00_01_07.jpg, z = 10.0 um

89 suggested cells in 200100013_00_01_08.jpg, z = 5.0 um

69 suggested cells in 200100013_00_01_09.jpg, z = 0.0 um

44 suggested cells in 200100013_00_01_10.jpg, z = −5.0 um

30 suggested cells in 200100013_00_01_11.jpg, z = −10.0 um

22 suggested cells in 200100013_00_01_12.jpg, z = −15.0 um

11 suggested cells in 200100013_00_01_13.jpg, z = −20.0 um

5 suggested cells in 200100013_00_01_14.jpg, z = −25.0 um

11 suggested cells in 200100013_00_01_15.jpg, z = −30.0 um

7 suggested cells in 200100013_00_01_16.jpg, z = −35.0 um

5 suggested cells in 200100013_00_01_17.jpg, z = −40.0 um

Cell evidence for the strongest extrema in 200100013/200100013_00_01.d2.s.mat
Appendix D

Results of Visual Inspection of Automated Cell Detections

In this appendix, we present our notes from the visual inspection of the automatic detections of cells in each of the 21 image sequences.

First we present the correspondence between the automatic detections (the accepted and rejected cells from appendix B, numbered starting with the accepted cells) and the true cells (presented in appendix A). False detections show up in this table as a minus sign ("%") for the (non)-corresponding cell number.

For each the true detections we have calculated the displacement in z-position in micro meters. In this table “D” is for the detected cell, “T” is for the true cell. dz is the difference in z-position in micrometers.

For each of the false detections we comment on possible type and cause.
1. False Positive Types:
   1. Multiple detection by out of focus cell.
   2. Wrong sign of extrema.
   3. Coincidence between two (or more) structures/cells.
   4. Fragment.
   5. Wrong size, partly triggered by coincident structure
   6. Wrong z-position.

2. False Negative Types:
   1. Not suggested.
   2. Suggested but overlapping with true positive.
   3. Suggested but overlapping with false negative.
   4. Not detected as 1 of the 1000 strongest.

1. False Positive Causes:
   1. Wrong cell image model.
   2. Bad overlap criteria
   3. False negative didn’t rule out this false positive.
   4. Wrong model match measure?
   5. Wrong model size (too small/large).
   6. Too low threshold.
   7. True detections not strong enough.

2. False Negative Causes:
   1. (Not suggested) Too high threshold for suggested.
   2. (Not suggested) Cell image lacks (white/dark) edges, (due to influence from other cells?)
   3. (Not suggested) Model image lacks out of focus dark shadow halo.
   4. (Not suggested) model image is not elliptical.
   5. (Suggested but overlapping) Non flexible sphere model. (Strong elliptic rules out too much)
   6. (Suggested but overlapping) Wrong z-positions from Greedy algorithm.
   7. (Suggested but overlapping) Wrong number of cells from Greedy algorithm.
   8. (Suggested but overlapping) Lack of integrated match measure for many cells.
   9. (Suggested but overlapping) Too strict overlapping criteria.
   10. Wrong image model
   11. (Suggested but overlapping) Model is not elliptic. (week because elliptic)
12

D  1  2  3  4  5
T  3  3  1  4  %
dz 15  0  35  15

D5T%: Type: 1.1  Cause: 1.1, (1.2?)

13

D  1  2  3  4
T  4  3  2  1
dz 20  25  25  20

18

D  1  2  3  %
T  5  3  1  2  4
dz 20  30  15

T2D%: Type: 2.1  Cause: 2.1, 2.2, 2.3
T4D%: Type: 2.1  Cause: 2.1, (2.2), 2.3, 2.4

19

D  1  2  3  4  %
T  4  2  3  %  1
dz 25  30  25

D4T%: Type: 1.2, 1.3  Causes: 1.1 (lack of sign), 1.3
T1D%: Type: 2.2  Causes: 2.5, 2.6

22

D  1  2  3  4
T  3  4  1  2
dz 30  35  15  25

23

D  1  2
T  1  2
dz 25  25

Note: D2.R<<T2.R  Causes: Greed, badz, with overlap

24

D  1  %  %  %  %
D1T%: Type: 1.3  Cause: 1.3, 1.4
T(1-4)D% Type: 2.3  Cause: 2.7, 2.8

D 1 2 3 4
T 1 3 2  % 4
dz 15 15 10

D4T%: Type: 1.1,(1.3)  Cause: 1.1, 1.4, 1.5
T4D% Type: 2.2  Cause: 2.5, 2.9

D 1 2 3 4 5 6 7  %  %
T 2 4 1 6  %  %  % 3 5
dz 10 5 0 10

D5T%: Type: 1.4  Cause: 1.5
D6T%: Type: 1.3  Cause: 1.5
D7T%: Type: 1.2  Cause: 1.4,1.5,coicidence,1.6
T3D% Type: 2.1  Cause: 2.1
T4D% Type: 2.4  Cause: 2.10

Is this embryo typical?
How important is it?

D 1 2 3 4
T 3 1 4 2
dz 35 20 20 10

D 1 2 3 4 5 6  %
T 4 3 2 4  %  %  1
dz 10 15 20 0

D5T%: Type: 1.2,1.3  Cause: 1.1 (sign)
D6T%: Type: 1.1  Cause: 1.1 (focusing missing in model)
T1D% Type: 2.1  Cause: 2.1,2.2,2.4

D 1  %  %  %
<table>
<thead>
<tr>
<th>D</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>%</td>
</tr>
<tr>
<td>dz</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

D1T%: Type: 1.3  Cause: 1.3, 1.1 (elliptical)
T1D%: Type: 2.3  Cause: 2.5, 2.10
T2D%: Type: 2.3  Cause: 2.5
T3D%: Type: 2.3  Cause: 2.5

32

<table>
<thead>
<tr>
<th>D</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>%</td>
</tr>
<tr>
<td>dz</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D5T%: Type: 1.2,1.3  Cause: 1.1 (sign)

33

<table>
<thead>
<tr>
<th>D</th>
<th>1</th>
<th>2</th>
<th>%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>2</td>
<td>%</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>dz</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D2T%: Type: 1.2,1.3  Cause: 1.1 (sign)
T1D%: Type: 2.3  Cause: 2.7, 2.8, 2.11
T3D%: Type: 2.3  Cause: 2.7, 2.8, 2.11

Maybe D1 should have been classified as false positive:
D1T%: Type: 1.3  Cause: 1.3
T2D%: Type: 2.1  Cause: 2.4

34

<table>
<thead>
<tr>
<th>D</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>%</th>
<th>%</th>
<th>%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

D1T%: Type: 1.3  Cause: 1.3, 1.4
D2T%: Type: 1.4  Cause: 1.5, 1.6
D3T%: Type: 1.4  Cause: 1.5, 1.6
D4T%: Type: 1.3  Cause: 1.3, 1.6
T1D%: Type: 2.3  Cause: 2.7, 2.8
T2D%: Type: 2.3, 2.1  Cause: 2.7, 2.8, 2.4, 2.2
T3D%: Type: 2.3  Cause: 2.7, 2.8,
T4D%: Type: 2.3, 2.1  Cause: 2.7, 2.8, 2.4, 2.2

37

<table>
<thead>
<tr>
<th>D</th>
<th>1</th>
<th>2</th>
<th>%</th>
<th>%</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>%</td>
<td>%</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

D1T%: Type: 1.3  Cause: 1.3, 1.4
D2T%: Type: 1.2, 1.5  Cause: 1.4, 1.5, 1.6
T1D%: Type: 2.3, 2.1  Cause: 2.7, 2.8, 2.11, 2.1, 2.4, 2.3
T2D%: Type: 2.3  Cause: 2.7, 2.8, 2.11
T3D%: Type: 2.3  Cause: 2.7, 2.8, 2.11
T4D%: Type: 2.3  Cause: 2.7, 2.8, 2.11

38

D 1 % % % %
T % 1 2 3 4

D1T%: Type: 1.3  Cause: 1.3, 1.4
T1D%: Type: 2.3  Cause: 2.7, 2.8, 2.11
T2D%: Type: 2.3  Cause: 2.7, 2.8, 2.11
T3D%: Type: 2.3  Cause: 2.7, 2.8
T4D%: Type: 2.3  Cause: 2.7, 2.8

42

eg2-file missing.

It seems like the sequence doesn’t really span enough z-values!?

D 1 2 3 4 5 6 % % %
T 1 % % % % 2 3 4

D2T%: Type: 1.5  Cause: 1.3, 1.4
D3-5T%: Type: 1.2  Cause: 1.1, 1.4, 1.6
D6T%: Type: 1.1  Cause: 1.6, 1.3
T2D%: Type: 2.1  Cause: 2.1, 2.2, 2.4
T3D%: Type: 2.3  Cause: 2.7, 2.8
T4D%: Type: 2.3  Cause: 2.7, 2.8

43

D 1 % % % %
T % 1 2 3 4

D1T%: Type: 1.3  Cause: 1.3, 1.4
T1D%: Type: 2.3  Cause: 2.7, 2.8
T2D%: Type: 2.3  Cause: 2.7, 2.8, 2.11
T3D%: Type: 2.3  Cause: 2.7, 2.8
T4D%: Type: 2.3  Cause: 2.7, 2.8, 2.11

45

D 1 % % % %
T % 1 2 3 4

D1T%: Type: 1.3  Cause: 1.3, 1.4
Are the true cells 2 and 4 typical and important?

D 1 % % % %
T % 1 2 3 4

D1T%: Type: 1.3 Cause: 1.3, 1.4
T1D%: Type: 2.3, 2.1 Cause: 2.7, 2.8, 2.11, 2.4
T2D%: Type: 2.4 Cause: 2.4, 2.2
T3D%: Type: 2.3 Cause: 2.7, 2.8
T4D%: Type: 2.4 Cause: 2.4, 2.2
Appendix E

Evaluation II - First Improvements 1

In this appendix we present figures displaying the equator-contour of each of the detected cells (accepted in green and rejected in red) in the image which is the equator plane of the cell. These figures were generated from the Select written matlab files “run_20_1_40_3_4/(*.mat)”. 

12

13
Sphere 1 in 200100029_00_02_03.jpg, z = 30.0 um

Sphere 2 in 200100029_00_02_06.jpg, z = 15.0 um

Sphere 3 in 200100029_00_02_14.jpg, z = −25.0 um

Sphere 4 in 200100029_00_02_08.jpg, z = 5.0 um

Sphere 1 in 200100030_00_02_09.jpg, z = 0.0 um

Sphere 2 in 200100030_00_02_13.jpg, z = −20.0 um

Sphere 3 in 200100030_00_02_13.jpg, z = −20.0 um

Sphere 4 in 200100030_00_02_06.jpg, z = 15.0 um

Sphere 1 in 200100030_00_02_12.jpg, z = −15.0 um

Sphere 2 in 200100030_00_02_05.jpg, z = 20.0 um

Sphere 1 in 200100031_00_02_10.jpg, z = −5.0 um

Sphere 2 in 200100031_00_02_02.jpg, z = 35.0 um

Sphere 3 in 200100031_00_02_01.jpg, z = 40.0 um
Sphere 1 in 200100032_00_02_01.jpg, z = 40.0 um

Sphere 2 in 200100032_00_02_01.jpg, z = 40.0 um

Sphere 3 in 200100032_00_02_02.jpg, z = 35.0 um

Sphere 4 in 200100032_00_02_10.jpg, z = −5.0 um

Sphere 1 in 200100033_00_02_07.jpg, z = 10.0 um

Sphere 2 in 200100033_00_02_06.jpg, z = 15.0 um

Sphere 1 in 200100033_00_02_17.jpg, z = −40.0 um

Sphere 1 in 200100034_00_02_06.jpg, z = 15.0 um

Sphere 2 in 200100034_00_02_04.jpg, z = 25.0 um

Sphere 3 in 200100034_00_02_10.jpg, z = −5.0 um

Sphere 1 in 200100034_00_02_02.jpg, z = 35.0 um

Sphere 1 in 200100037_00_02_10.jpg, z = −5.0 um

Sphere 2 in 200100037_00_02_05.jpg, z = 20.0 um

Sphere 3 in 200100037_00_02_01.jpg, z = 40.0 um

Sphere 4 in 200100037_00_02_11.jpg, z = −10.0 um
Sphere 1 in 200100045_00_02_09.jpg, z = 0.0 um
Sphere 2 in 200100045_00_02_01.jpg, z = 40.0 um
Sphere 3 in 200100045_00_02_02.jpg, z = 35.0 um
Sphere 4 in 200100045_00_02_01.jpg, z = 40.0 um

Sphere 1 in 200100047_00_02_10.jpg, z = -5.0 um
Sphere 2 in 200100047_00_02_05.jpg, z = 20.0 um
Sphere 3 in 200100047_00_02_08.jpg, z = 5.0 um
Sphere 1 in 200100047_00_02_01.jpg, z = 40.0 um
Appendix F

Evaluation II - First Improvements 2

In this appendix we present figures displaying the equator-contour of each of the detected cells (accepted in green and rejected in red) in the image which is the equator plane of the cell. These figures were generated from the Select written matlab files “run_10_1_40_5_8/*.mat”.

12

13
Sphere 1 in 200100042_00_02_02.jpg, \( z = 35.0 \) \( \mu \text{m} \)

Sphere 2 in 200100042_00_02_06.jpg, \( z = 15.0 \) \( \mu \text{m} \)

Sphere 3 in 200100042_00_02_12.jpg, \( z = -15.0 \) \( \mu \text{m} \)

Sphere 4 in 200100042_00_02_15.jpg, \( z = 99.0 \) \( \mu \text{m} \)

Sphere 1 in 200100043_00_02_05.jpg, \( z = 20.0 \) \( \mu \text{m} \)

Sphere 2 in 200100043_00_02_01.jpg, \( z = 40.0 \) \( \mu \text{m} \)

Sphere 3 in 200100043_00_02_07.jpg, \( z = -40.0 \) \( \mu \text{m} \)

Sphere 4 in 200100043_00_02_01.jpg, \( z = 40.0 \) \( \mu \text{m} \)

Sphere 1 in 200100043_00_02_08.jpg, \( z = 5.0 \) \( \mu \text{m} \)

Sphere 2 in 200100043_00_02_01.jpg, \( z = 40.0 \) \( \mu \text{m} \)
Sphere 1 in 200100045_00_02_07.jpg, z = 10.0 μm

Sphere 2 in 200100045_00_02_01.jpg, z = 40.0 μm

Sphere 3 in 200100045_00_02_02.jpg, z = 35.0 μm

Sphere 4 in 200100045_00_02_01.jpg, z = 40.0 μm

Sphere 1 in 200100047_00_02_09.jpg, z = 0.0 μm

Sphere 2 in 200100047_00_02_04.jpg, z = 25.0 μm

Sphere 3 in 200100047_00_02_03.jpg, z = 30.0 μm

173
Appendix G

Evaluation II - First Improvements

In this appendix we present figures displaying the equator-contour of each of the detected cells (accepted in green and rejected in red) in the image which is the equator plane of the cell. These figures were generated from the Select written matlab files “run_20_1_40_5_8/*.mat”.

12

13

18
Sphere 1 in 200100026_00_02_11.jpg, z = −10.0 μm

Sphere 2 in 200100026_00_02_05.jpg, z = 20.0 μm

Sphere 3 in 200100026_00_02_08.jpg, z = 5.0 μm

Sphere 4 in 200100026_00_02_02.jpg, z = 35.0 μm

Sphere 5 in 200100026_00_02_06.jpg, z = 15.0 μm

Sphere 1 in 200100029_00_02_03.jpg, z = 30.0 μm

Sphere 2 in 200100029_00_02_06.jpg, z = 15.0 μm

Sphere 3 in 200100029_00_02_11.jpg, z = −10.0 μm

Sphere 4 in 200100029_00_02_08.jpg, z = 5.0 μm

Sphere 1 in 200100030_00_02_12.jpg, z = −15.0 μm

Sphere 2 in 200100030_00_02_09.jpg, z = 0.0 μm

Sphere 3 in 200100030_00_02_13.jpg, z = −20.0 μm

Sphere 4 in 200100030_00_02_07.jpg, z = 10.0 μm

Sphere 1 in 200100030_00_02_06.jpg, z = 15.0 μm

Sphere 2 in 200100030_00_02_13.jpg, z = −20.0 μm
Sphere 1 in 200100031_00_02_10.jpg, z = −5.0 μm

Sphere 2 in 200100031_00_02_04.jpg, z = 25.0 μm

Sphere 3 in 200100031_00_02_01.jpg, z = 40.0 μm

Sphere 1 in 200100032_00_02_01.jpg, z = 40.0 μm

Sphere 2 in 200100032_00_02_01.jpg, z = 40.0 μm

Sphere 3 in 200100032_00_02_02.jpg, z = 35.0 μm

Sphere 4 in 200100032_00_02_09.jpg, z = 0.0 μm

Sphere 1 in 200100033_00_02_07.jpg, z = 10.0 μm

Sphere 2 in 200100033_00_02_06.jpg, z = 15.0 μm

Sphere 1 in 200100033_00_02_16.jpg, z = −35.0 μm

Sphere 2 in 200100033_00_02_02.jpg, z = 35.0 μm

Sphere 3 in 200100033_00_02_03.jpg, z = 30.0 μm

Sphere 1 in 200100034_00_02_06.jpg, z = 15.0 μm

Sphere 2 in 200100034_00_02_04.jpg, z = 25.0 μm

Sphere 3 in 200100034_00_02_08.jpg, z = 5.0 μm

Sphere 4 in 200100034_00_02_03.jpg, z = 30.0 μm

Sphere 1 in 200100034_00_02_02.jpg, z = 35.0 μm
Appendix H

Evaluation III - Combinatoric Search

In this appendix we present figures displaying the equator-contour of each of the detected cells in the image which is the equatorplane of the cell. They have not been marked as accepted or rejected.
1.4: 200100032_00_02, z=30.0 μm
2.4: 200100032_00_02, z=25.0 μm
3.4: 200100032_00_02, z=20.0 μm
4.4: 200100032_00_02, z=15.0 μm

1.3: 200100033_00_02, z=5.0 μm
2.3: 200100033_00_02, z=20.0 μm
3.3: 200100033_00_02, z=30.0 μm
4.3: 200100033_00_02, z=−10.0 μm

1.4: 200100034_00_02, z=15.0 μm
2.4: 200100034_00_02, z=35.0 μm
3.4: 200100034_00_02, z=25.0 μm
4.4: 200100034_00_02, z=5.0 μm

1.4: 200100037_00_02, z=5.0 μm
2.4: 200100037_00_02, z=20.0 μm
3.4: 200100037_00_02, z=30.0 μm
4.4: 200100037_00_02, z=−5.0 μm

1.4: 200100038_00_02, z=20.0 μm
2.4: 200100038_00_02, z=30.0 μm
3.4: 200100038_00_02, z=20.0 μm
4.4: 200100038_00_02, z=−5.0 μm
Appendix I

Embryo Optics - Observed Image Focus Sequences

In this appendix we present observed image focus sequences analyzed in chapter 6 on “Embryo Optics” of 4 oocytes using Hoffman Modulation Contrast, Paraxial Illuminated Bright Field, and Oblique Ring Bright Field.

Oocyte 1 HMC sequence
Oocyte 1 Axial Bright Field

200  190  180  170  160  150  140
  130  120  110  100  90  80  70
  60  50  40  30  20  10  0
 -10  -20  -30  -40  -50  -60  -70
 -80  -90  -100  -110  -120  -130  -140
 -150  -160  -170  -180  -190  -200  -250
 -300  -350  -400  -450  -500  -550  -600
 -650  -700  -750  -800  -850  -900  -950
 -1000  -1050  -1100  -1150  -1200  -1250  -1300
 -1350  -1400  -1500  -1600  -1700  -1800
Oocyte 2 HMC sequence
Oocyte 2 Axial Bright Field
### Oocyte 2 Oblique Ring Bright Field

<table>
<thead>
<tr>
<th>100</th>
<th>90</th>
<th>80</th>
<th>70</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
</tr>
<tr>
<td>50</td>
<td>40</td>
<td>30</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
</tr>
<tr>
<td>0</td>
<td>-10</td>
<td>-20</td>
<td>-30</td>
<td>-40</td>
</tr>
<tr>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>-50</td>
<td>-60</td>
<td>-70</td>
<td>-80</td>
<td>-90</td>
</tr>
<tr>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
</tr>
<tr>
<td>-100</td>
<td>-150</td>
<td>-200</td>
<td>-250</td>
<td>-300</td>
</tr>
<tr>
<td><img src="image21.png" alt="Image" /></td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
<td><img src="image25.png" alt="Image" /></td>
</tr>
<tr>
<td>-350</td>
<td>-400</td>
<td>-450</td>
<td>-500</td>
<td>-550</td>
</tr>
<tr>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
<td><img src="image28.png" alt="Image" /></td>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
</tr>
<tr>
<td>-600</td>
<td>-650</td>
<td>-700</td>
<td>-750</td>
<td>-800</td>
</tr>
<tr>
<td><img src="image31.png" alt="Image" /></td>
<td><img src="image32.png" alt="Image" /></td>
<td><img src="image33.png" alt="Image" /></td>
<td><img src="image34.png" alt="Image" /></td>
<td><img src="image35.png" alt="Image" /></td>
</tr>
<tr>
<td>-850</td>
<td>-900</td>
<td>-950</td>
<td>-1000</td>
<td></td>
</tr>
</tbody>
</table>
Oocyte 3 HMC sequence

<table>
<thead>
<tr>
<th>100</th>
<th>95</th>
<th>90</th>
<th>85</th>
<th>80</th>
<th>75</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
</tr>
<tr>
<td>65</td>
<td>60</td>
<td>55</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
</tr>
<tr>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
</tr>
<tr>
<td>30</td>
<td>25</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
<td><img src="image21.png" alt="Image" /></td>
</tr>
<tr>
<td>-5</td>
<td>-10</td>
<td>-15</td>
<td>-20</td>
<td>-25</td>
<td>-30</td>
<td>-35</td>
</tr>
<tr>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
<td><img src="image25.png" alt="Image" /></td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
<td><img src="image28.png" alt="Image" /></td>
</tr>
<tr>
<td>-40</td>
<td>-45</td>
<td>-50</td>
<td>-55</td>
<td>-60</td>
<td>-65</td>
<td>-70</td>
</tr>
<tr>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
<td><img src="image31.png" alt="Image" /></td>
<td><img src="image32.png" alt="Image" /></td>
<td><img src="image33.png" alt="Image" /></td>
<td><img src="image34.png" alt="Image" /></td>
<td><img src="image35.png" alt="Image" /></td>
</tr>
<tr>
<td>-75</td>
<td>-80</td>
<td>-85</td>
<td>-90</td>
<td>-95</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image36.png" alt="Image" /></td>
<td><img src="image37.png" alt="Image" /></td>
<td><img src="image38.png" alt="Image" /></td>
<td><img src="image39.png" alt="Image" /></td>
<td><img src="image40.png" alt="Image" /></td>
<td><img src="image41.png" alt="Image" /></td>
<td><img src="image42.png" alt="Image" /></td>
</tr>
<tr>
<td>-100</td>
<td>-150</td>
<td>-200</td>
<td>-250</td>
<td>-300</td>
<td>-350</td>
<td>-400</td>
</tr>
<tr>
<td><img src="image43.png" alt="Image" /></td>
<td><img src="image44.png" alt="Image" /></td>
<td><img src="image45.png" alt="Image" /></td>
<td><img src="image46.png" alt="Image" /></td>
<td><img src="image47.png" alt="Image" /></td>
<td><img src="image48.png" alt="Image" /></td>
<td><img src="image49.png" alt="Image" /></td>
</tr>
<tr>
<td>-450</td>
<td>-500</td>
<td>-550</td>
<td>-600</td>
<td>-650</td>
<td>-700</td>
<td>-750</td>
</tr>
<tr>
<td><img src="image50.png" alt="Image" /></td>
<td><img src="image51.png" alt="Image" /></td>
<td><img src="image52.png" alt="Image" /></td>
<td><img src="image53.png" alt="Image" /></td>
<td><img src="image54.png" alt="Image" /></td>
<td><img src="image55.png" alt="Image" /></td>
<td><img src="image56.png" alt="Image" /></td>
</tr>
<tr>
<td>-800</td>
<td>-850</td>
<td>-900</td>
<td>-950</td>
<td>-1000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Oocyte 3 Axial Bright Field

<table>
<thead>
<tr>
<th>100</th>
<th>95</th>
<th>90</th>
<th>85</th>
<th>80</th>
<th>75</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>65</td>
<td>60</td>
<td>55</td>
<td>50</td>
<td>45</td>
<td>40</td>
<td>35</td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>30</td>
<td>25</td>
<td>20</td>
<td>15</td>
<td>10</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>-5</td>
<td>-10</td>
<td>-15</td>
<td>-20</td>
<td>-25</td>
<td>-30</td>
<td>-35</td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>-40</td>
<td>-45</td>
<td>-50</td>
<td>-55</td>
<td>-60</td>
<td>-65</td>
<td>-70</td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>-75</td>
<td>-80</td>
<td>-85</td>
<td>-90</td>
<td>-95</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-100</td>
<td>-150</td>
<td>-200</td>
<td>-250</td>
<td>-300</td>
<td>-350</td>
<td>-400</td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>-450</td>
<td>-500</td>
<td>-550</td>
<td>-600</td>
<td>-650</td>
<td>-700</td>
<td>-750</td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
</tr>
<tr>
<td>-800</td>
<td>-850</td>
<td>-900</td>
<td>-950</td>
<td>-1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td><img src="image" alt="" /></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Oocyte 3 Oblique Ring Bright Field

Oocyte 4 HMC sequence
Oocyte 4 Axial Bright Field
Oocyte 4 Oblique Ring Bright Field

100  95  90  85  80  75  70

65  60  55  50  45  40  35

30  25  20  15  10  5   0

-5 -10 -15 -20 -25 -30 -35

-40 -45 -50 -55 -60 -65 -70

-75 -80 -85 -90 -95 -100 -150

-200 -250 -300 -350 -400 -450 -500

-550 -600 -650 -700 -750 -800 -850

-900 -950 -1000