

Motion Pattern Analysis for Biomedical Applications

Quantitative comparisons in 2D and 3D time-lapse microscopy

Robert Bensch

Freiburg im Breisgau, Germany

December 2017



Dissertation

zur Erlangung des Doktorgrades der Technischen Fakultät
der Albert-Ludwigs-Universität Freiburg im Breisgau

Dekan: Prof. Dr. Oliver Paul

Prüfungskommission: Prof. Dr.-Ing. Matthias Teschner (Vorsitz)
Prof. Dr. Rolf Backofen (Beisitz)
apl. Prof. Dr. Olaf Ronneberger (1. Gutachter)
Prof. Dr.-Ing. Thomas Brox (2. Gutachter)

Datum der Disputation: 15.12.2017

Unterstützt durch: The Excellence Initiative of the German Federal and State Governments
(BIOSS Centre for Biological Signalling Studies, EXC 294)

TO MY WIFE DANY &
MY CHILDREN LARISSA & JUSTUS



TO MY PARENTS &
MY BROTHER MAIK

Acknowledgements

At this point I want to thank all the people who accompanied my way and made it both possible and pleasant for me to work on, write and successfully complete my dissertation.

I want to start thanking Prof. Hans Burkhardt who taught me pattern recognition and image processing in his lectures and seminars and let me join his group of Pattern Recognition and Image Processing as a PhD student. He provided a great place to work at and always supported me.

I would like to sincerely thank Prof. Olaf Ronneberger whom I got to know as my seminar topic supervisor when he was a PhD student himself. He supervised my Master's thesis and let me continue my research in his Image Analysis Group. His support was far beyond average, I want to thank him for the time he invested and for all the interesting interdisciplinary biomedical projects. Thank you for all the great ideas and that there was always time for discussions. You were critical but always positive and constructive, your way of tackling and solving problems fascinated me. You were a great PhD supervisor in both professional and human terms.

I thank Prof. Thomas Brox for bringing new computer vision expertise to Freiburg and for always supporting me as a PhD student in his group. Thank you for generating such a nice working environment from both a professional and personal perspective. I thank you for stimulating and productive discussions on my work, new ideas and comments on my manuscripts. Many thanks for being the second supervisor and examiner of my PhD thesis.

For being part of the examination committee I further thank Prof. Rolf Backofen and Prof. Matthias Teschner.

Without all my colleagues the time at the group would not have been so nice as it was. I start thanking Dr. Thorsten Falk, Dr. Dominic Mai, Dr. Benjamin Drayer and Özgün Çiçek who kindly agreed to proofread my thesis. Many thanks to Thorsten whom I already know for long time and who became a friend. He was a great colleague not only in terms of his professional and technical expertise but also in human terms. He later became my boss when he luckily continued heading the Core Facility Image Analysis, after Olaf Ronneberger had been released to Google DeepMind. I thank you Thorsten for your friendship and for supporting me all the time. Dominic, thank you for the nice and funny time we had sharing an office together, thank you for your friendship, for introducing me to the juggling world and also for the great time we had when making music together and playing at the faculty summer party - many greetings to André, Babsi and Jenny as well. Benjamin, thank you for becoming a friend and being such a nice colleague with always positive attitude. Özgün, thank you for your company and nice kind during my time at the group and also thank you for proofreading, I wish you good luck and progress with your PhD thesis.

I want to thank Margret Keuper for the good time we had, from a scientific, personal and sportive point of view. Together with Alexandra Teynor, who was my office mate when I started my PhD, we formed a running group and also participated in running events. I thank Kun Liu for being such a kind and scientifically amazing colleague and office mate. I thank Henrik Skibbe and Matthias Schlachter, who were also writing their Master's thesis in the diploma room of the group where we got to know each other. I want to thank Phillip Fischer for bringing a "fresh breeze" into the group - he set up drinks and snakes for the group, organized several events and introduced highlights, such as the game "Beat the Profs" and also set up the very helpful server load web page.

I also appreciate all the other colleagues I got to know and try to list here, hopefully without anyone missing: Ahmed, Artemij, Anton, Eddy, Nikolaus, Gabriel, Nima, Maxim, Benjamin U., Huizhong, Christian, Mohammadreza, Peter, Alexey, Naveen, Frank, Maja, Qing, Phillip C., Lingyu, Wan, Nikos, Dominik R., Mario, Janis, Janina, Marco, and Lokesh, whom I did not really get to know, but whose office desk I took over when he left.

I greatly appreciate our secretaries Cynthia Findlay, Jasmin Anders and Lilian Siebert for their organizational and administrative work, and for having an open ear for all our questions and concerns. A great appreciation also for our system administrator Stefan Teister and his team Felix Ratzkowski,

Patrick Hornecker, including Thorsten. Thank you for providing such an excellent and robust technical system.

I am also very thankful that I had my friends Sergio Feo and Marco Muñiz around who did their PhD in the Software Engineering group, a floor below in the same building.

My thesis and work was not possible without the numerous interdisciplinary projects, project partners and co-authors. Many thanks to Jan Huisken, Nico Scherf, Wolfgang Driever, Sungmin Song, Mathias Simons, Deniz Saltukoglu, Klaus Palme, Alexander Dovzhenko, Robert F. Murphy, Karsten Voigt, Virginie Lecaudey, Sandra Ernst, Gerd Walz, Barbara Müller, Vadym Budnyk, Christian Fleck, Bettina Greese, Martin Hülskamp and finally Roland Nitschke for providing state-of-the-art microscopy techniques and courses in the Life Imaging Center.

I appreciate the funding from the Excellence Initiative of the German Federal and State Governments (BIOSS Centre for Biological Signalling Studies, EXC 294) without which my work and this dissertation would not have been possible in this way.

My final thanks are devoted my family, especially my parents and my brother who always supported me in so many ways and in bad times and always believed in me. Beyond that I am very grateful for my wife's love, her support and backing all the years. She gave me the freedom to pursue my work, she gave me love and our two lovely children who brightened up our lives.

Munich, July 2018

Robert Bensch

Abstract

In current biomedical research 3D time-lapse microscopy allows to investigate developmental processes, such as the growth of whole organisms, on the cellular level. Embryos of zebrafish (*Danio rerio*) or fruit fly (*Drosophila*), for example, can be recorded *in vivo* in 3D space. Processes such as cell division and cell migration can be observed, that lead to the formation and growth of tissue, and finally, to the development of the whole organism. Analyzing cell motion and the emerging motion patterns is crucial for the understanding of developmental processes and their underlying mechanisms. Typically, experiments in genomics, proteomics and metabolomics aim at comparing *wild-type* organisms to genetically manipulated organisms, usually referred to as *mutants*. The goal is to link observable changes in the manipulated organisms (changes in the *phenotype*) to the underlying genetic manipulation (manipulation of the *genotype*) to infer the function of certain genes. Generally, a quantitative and unbiased comparison is desired to discover differences and verify their statistical significance. Such experiments usually generate huge amounts of data, consisting of time sequences of 3D volumetric images. Since manual evaluation is neither feasible nor desired, techniques for automated image analysis have become indispensable.

This thesis presents approaches that enable motion pattern analysis in complex biomedical applications based on 2D and 3D time-lapse microscopy. They allow precise physical measurements for quantitative analysis and robust comparisons of motion patterns, which is essential for evaluating experiments in biomedical research and specifically in developmental biology.

Two methods that are based on a trajectory representation of motion provide the main contributions of this thesis. Trajectories yield a rich motion representation, including long-term motion information, that is independent of object appearance, which is very important in microscopy when comparing images that are affected by different imaging settings.

In chapter 2 we propose a general method to detect motion anomalies in 3D+time data. The setting of anomaly detection fits well to the usual biomedical tasks where wild-type patterns define a normal model and significant deviations, *i.e.* anomalies, are to be detected in mutants. We detect anomalies by placing spatiotemporally deformed instances of a prototype pattern to reconstruct a test pattern. In the test pattern, we regard poorly reconstructed patterns showing strong deviations from the elastically registered prototype patterns as anomalies. To define accepted variations a prototype model is learned from multiple training sequences. We propose a new method for elastic registration of 3D+time trajectory patterns, together with a new efficient and robust supertrajectory representation and a modified hashing approach to efficiently produce transformation hypotheses. The method performs well in detecting subtle anomalies on a new motion anomaly dataset of juggling patterns, and we demonstrated the applicability to biological motion patterns in zebrafish development.

The second trajectory-based method allows to detect specific motion patterns in 3D+time data using spatiotemporal geometrical models. In particular, we developed a model to detect cell intercalation which is an essential pattern in developmental biology. Cell intercalations occur when cells enter the space between adjacent cells and play an important role in tissue formation. The approach builds on single cell motion trajectories and specifies the motion pattern to be detected by spatiotemporal transition functions of a geometrical model. The method is robust to noisy and incomplete measurements and handles the variability within the class of 3D intercalations. We applied our method to biological data from zebrafish development and performed a quantitative comparison of cell intercalations and their motion statistics between wild-type and mutant embryos.

Two contour-based approaches form the second part of this thesis. Instead of trajectories, sequences of evolving contours are used to represent motion patterns.

We propose a new robust, effective, and surprisingly simple approach for the segmentation of cells in phase contrast microscopy. Phase contrast microscopy generates strong intensity gradients along interfaces of media with different physical densities. They allow to obtain clear boundary responses even for perfectly transparent samples. However, classical edge-based image segmentation fails due

to the complex intensity profile consisting of a bright-to-dark and a dark-to-bright transition at the boundary and other artifacts from phase contrast microscopy. Our algorithm exploits the properties of positive phase contrast microscopy where the true cell borders always appear as a dark-to-bright transition in outwards direction. The segmentation mask is effectively found by a fast min-cut approach. In contrast to classical min-cut our graph contains directed edges with asymmetric edge weights. This modification to classical min-cut allows to choose optimization parameters from a wider range without affecting segmentation performance and surpasses segmentation quality with symmetric edge weights. We outperformed the top ranked methods from the ISBI Cell Tracking Challenge (CTC) 2014 on the phase contrast dataset, and reached second place in the ISBI CTC 2015.

We were able to directly apply our approach for cell segmentation on phase contrast images to produce cell contour input data for our second contour-based method. It investigates migrating cells and their motion patterns. We developed a method to detect symmetry-breaking events, which enabled automatic browsing of large amounts of data for these cellular events of interest. To investigate motion patterns along the cell contour, we implemented a method to compute protrusion/retraction maps that visualize the contour velocities in a 2D map in polar contour coordinates over time. We applied the approach to the analysis of spontaneous and electric field-controlled front-rear polarization of human keratinocytes. Our approach enabled the quantification of several experimental conditions and led to the extraction of biologically relevant results.

In conclusion, this thesis contains several contributions to tackle the task of motion pattern analysis and quantitative comparison in biomedical applications from 3D+time data. The presented approaches build on trajectory- and contour-based representations and yielded robust methods that have been successfully applied in real biomedical applications and published in the field of computer vision, as well as in the field of biology and at the intersection of image analysis, imaging and biomedical applications.

It seems reasonable to focus future research into the direction of deep learning as it is currently the most exciting and promising field for advances. However, many open questions and challenges have to be addressed to solve the tasks considered in this thesis using deep learning. For the near future, using a mixture of conventional and deep learning methods seems to be the most promising approach.

Zusammenfassung

In der aktuellen biomedizinischen Forschung erlaubt es die 3D Zeitraffer-Mikroskopie Entwicklungsprozesse, wie zum Beispiel das Wachstum kompletter Organismen, auf zellulärer Ebene zu untersuchen. Embryos vom Zebrafish (*Danio rerio*) oder der Fruchtfliege (*Drosophila*) zum Beispiel können *in vivo* in 3D aufgenommen werden. Prozesse wie Zellteilung und Zellmigration können beobachtet werden, die zur Anordnung und zum Wachstum von Gewebe, und schließlich zur Entwicklung des gesamten Organismus führen. Die Analyse von Zellbewegungen und der entstehenden Bewegungsmuster ist entscheidend für das Verständnis von Entwicklungsprozessen und deren zugrundeliegenden Mechanismen. Typischerweise zielen Experimente in Genomik, Proteomik und Metabolomik auf den Vergleich von Wildtyp und genetisch manipulierten Organismen, auch Mutanten genannt, ab. Das Ziel ist es, beobachtbare Veränderungen in den manipulierten Organismen (Änderungen im Phänotyp) mit den zugrundeliegenden genetischen Manipulationen (Manipulation des Genotyps) zu verknüpfen, um die Funktion bestimmter Gene abzuleiten. Grundsätzlich ist ein quantitativer und unverfälschter Vergleich gewünscht, um Unterschiede zu entdecken und deren statistische Signifikanz nachzuweisen. Solche Experimente erzeugen gewöhnlich riesige Mengen an Daten, bestehend aus Zeitsequenzen von 3D volumetrischen Bildern. Da eine manuelle Auswertung weder machbar noch gewünscht ist, sind Techniken zur automatischen Bildanalyse unentbehrlich geworden.

Diese Arbeit stellt Ansätze vor, die die Analyse von Bewegungsmustern in komplexen biomedizinischen Anwendungen, basierend auf 2D und 3D Zeitraffer-Mikroskopie, ermöglichen. Sie erlauben präzise physikalische Messungen zur quantitativen Analyse und zum robusten Vergleich von Bewegungsmustern, was entscheidend für die Evaluierung von Experimenten in der biomedizinischen Forschung und besonders in der Entwicklungsbiologie ist.

Zwei Methoden, die auf einer Trajektorien-Repräsentation von Bewegung basieren, stellen die Hauptbeiträge dieser Arbeit dar. Trajektorien liefern eine mächtige Bewegungsrepräsentation, einschließlich Langzeit-Bewegungsinformationen, die unabhängig von der Objekterscheinung ist. Dies ist zum Beispiel in der Mikroskopie sehr wichtig, wenn Bilder verglichen werden sollen, die mit unterschiedlichen Einstellungen aufgenommen wurden.

In Kapitel 2 schlagen wir eine generelle Methode zur Detektion von Bewegungsanomalien in 3D+Zeit Daten vor. Das Szenario der Anomaliedetektion passt gut zu den üblichen biomedizinischen Fragestellungen, bei denen Wildtyp-Muster ein normales Modell definieren und signifikante Abweichungen, d.h. Anomalien, in Mutanten gefunden werden sollen. Wir detektieren Anomalien durch das Platzieren von raumzeitlich deformierten Instanzen eines Prototyp-Musters, um ein Testmuster zu rekonstruieren. Im Testmuster betrachten wir schlecht rekonstruierte Muster, die starke Abweichungen zu den elastisch registrierten Prototyp-Mustern aufweisen, als Anomalien. Um tolerierte Variationen zu definieren, wird ein Prototyp-Modell aus mehreren Trainingssequenzen gelernt. Wir schlagen eine neue Methode zur elastischen Registrierung von 3D+Zeit Trajektorienmustern vor, zusammen mit einer neuen effizienten und robusten Supertrajektorien-Repräsentation und einem modifizierten Hashing-Ansatz, um effizient Transformationshypothesen zu generieren. Die Methode erzielt gute Ergebnisse bei der Detektion von subtilen Anomalien in einem neuen Datensatz für Bewegungsanomalien von Jongliermustern. Außerdem haben wir die Anwendbarkeit auf biologische Bewegungsmuster anhand von Entwicklungsmustern vom Zebrafisch gezeigt.

Die zweite trajektorien-basierte Methode ermöglicht die Detektion von spezifischen Bewegungsmustern in 3D+Zeit Daten unter Verwendung von raumzeitlichen geometrischen Modellen. Speziell haben wir ein Modell zur Detektion von Zellinterkalationen entwickelt. Diese stellen ein wesentliches Muster in der Entwicklungsbiologie dar. Interkalationen treten auf, wenn Zellen in den Raum zwischen benachbarten Zellen eindringen und spielen eine wichtige Rolle bei der Entstehung von Gewebe. Der Ansatz basiert auf Bewegungstrajektorien von Einzelzellen und legt das zu detektierende Bewegungsmuster durch raumzeitliche Übergangsfunktionen eines geometrischen Modells fest. Die Methode ist robust gegenüber Rauschen und unvollständigen Messungen und kann mit der Variabilität innerhalb

der Klasse der 3D Interkalationen umgehen. Wir haben unsere Methode auf biologische Daten der Entwicklung des Zebrafisches angewandt und einen quantitativen Vergleich der Zellinterkalationen und deren Bewegungsstatistiken zwischen Wildtyp- und Mutant-Embryonen durchgeführt.

Zwei kontur-basierte Verfahren bilden den zweiten Teil dieser Arbeit. Anstatt von Trajektorien, werden Kontur-Sequenzen verwendet, um Bewegungsmuster zu repräsentieren.

Wir stellen ein neues, robustes, effektives und erstaunlich einfaches Verfahren zur Segmentierung von Zellen in der Phasenkontrast-Mikroskopie vor. Phasenkontrast-Mikroskopie erzeugt starke Intensitätsgradienten entlang der Schnittstelle von Medien mit unterschiedlicher physikalischer Dichte. Diese ermöglichen es, deutliche Antworten an Phasengrenzen zu erhalten, sogar für perfekt transparente Zellen. Klassische kanten-basierte Bildsegmentierung schlägt jedoch fehl, aufgrund des komplexen Intensitätsprofils, bestehend aus einem hell-zu-dunkel und einem dunkel-zu-hell Übergang an der Grenze und anderen Artefakten der Phasenkontrast-Mikroskopie. Unser Algorithmus nutzt die Eigenschaften der positiven Phasenkontrast-Mikroskopie aus, wobei die wahren Zellgrenzen immer als hell-zu-dunkel Übergang in Außenrichtung erscheinen. Die Segmentierungsmaske wird effektiv über einen schnellen Min-Cut Ansatz gefunden. Im Gegensatz zum klassischen Min-Cut, enthält unser Graph gerichtete Kanten mit asymmetrischen Kantengewichten. Diese Modifikation gegenüber dem klassischen Min-Cut ermöglicht es, Optimierungsparameter aus einem größeren Bereich zu wählen, ohne die Segmentierungsqualität zu beeinflussen und übertrifft die Segmentierungsqualität mit symmetrischen Kantengewichten. Wir haben auf dem Phasenkontrast-Datensatz die bestplatziertesten Methoden aus dem ISBI Cell Tracking Challenge (CTC) 2014 übertroffen und erreichten den zweiten Platz bei dem ISBI CTC 2015.

Es war uns möglich, unseren Ansatz zur Zellsegmentierung in Phasenkontrast-Bildern unmittelbar zur Berechnung von Zellkonturen für unseren zweiten kontur-basierten Ansatz zu verwenden. Dieser untersucht migrierende Zellen und deren Bewegungsmuster. Wir haben eine Methode zur Detektion von symmetrie-brechenden Zellereignissen entwickelt, die ein automatisches Durchsuchen von großen Datenmengen nach diesen zellulären Ereignissen ermöglicht. Um Bewegungsmuster entlang der Zellkontur zu untersuchen, haben wir eine Methode implementiert, um Protrusions/Retractions-Karten zu berechnen. Diese visualisieren die Konturgeschwindigkeiten in einer 2D-Karte in Polarkoordinaten über die Zeit. Diesen Ansatz haben wir zur Analyse von spontaner und durch ein elektrisches Feld gesteuerter Polarisation von menschlichen Keratinocytes angewendet. Unser Ansatz ermöglichte die Quantifikation von mehreren experimentellen Bedingungen und lieferte biologisch relevante Ergebnisse.

Zusammenfassend stellt diese Arbeit mehrere Verfahren vor, die zur Analyse von Bewegungsmustern und zum quantitativen Vergleich in biomedizinischen Anwendungen mit 3D+Zeit Daten beitragen. Die präsentierten Verfahren basieren auf trajektorien- und kontur-basierten Repräsentationen und liefern robuste Methoden, die erfolgreich auf reale biomedizinische Fragestellungen angewandt wurden und im Bereich Computer Vision, sowie im Bereich Biologie und an der Schnittstelle von Bildanalyse, Bildgebung und biomedizinischen Anwendungen veröffentlicht wurden.

Es erscheint sinnvoll, zukünftige Forschung in die Richtung von Deep Learning zu fokussieren, da dies das gegenwärtig aufregendste und vielversprechendste Feld für Fortschritte ist. Allerdings müssen zuerst viele offene Fragen und Herausforderungen behandelt werden, bevor die Aufgabenstellungen in dieser Arbeit mit Deep Learning Ansätzen gelöst werden können. Für die nahe Zukunft scheint es daher am aussichtsvollsten zu sein, eine Mischung aus konventionellen und Deep Learning Methoden einzusetzen.

Contents

1	Introduction	13
1.1	Motivation	13
1.2	Contributions	17
1.3	Challenges for Deep Learning	18
1.4	Structure	20
1.5	Publications	21
1.6	Collaborations	22
1.7	Notation	23
I	Trajectory-based Methods	25
2	Spatiotemporal Deformable Prototypes for Motion Anomaly Detection	27
2.1	Introduction	27
2.2	Related Work	29
2.3	Supertrajectory Representation	30
2.3.1	Hierarchical Clustering	31
2.3.2	Supertrajectories	31
2.4	Detection of Motion Patterns	32
2.4.1	Detection Hypotheses by Efficient Hashing	32
2.5	Elastic Registration of Motion Patterns	34
2.5.1	Trajectory Association Function	34
2.5.2	Rigid Pre-Alignment	34
2.5.3	Elastic Registration	35
2.5.4	Temporal discretization	36
2.5.5	Energy Optimization	37
2.6	Motion Anomaly Detection	37
2.6.1	Learning a Spatiotemporal Deformable Prototype Model	37
2.6.2	Reconstructing Motion Patterns by Prototype Placements	38
2.6.3	Computing Anomaly Scores	38
2.7	Experiments	39
2.7.1	Basic Validation Experiments of Elastic Registration	39
2.7.2	Anomaly Detection in Juggling Patterns	43
2.7.3	Anomaly Detection in Biological Motion Patterns	53
2.8	Conclusion	57
3	Spatiotemporal Geometrical Models for Motion Pattern Detection	59
3.1	Biological Motivation and Related Work	60
3.2	Cell-triple Intercalation Model	63
3.3	Detection of Cell Intercalations	64
3.3.1	Fitting the Intercalation Model	65
3.3.2	Detecting Cell Intercalations	68
3.4	Experiments	69

3.4.1	Datasets	69
3.4.2	Image Quantification Details	69
3.4.3	Statistical Analysis	72
3.4.4	Results	72
3.4.5	Biological Discussion	80
3.5	Conclusion and Outlook	81
3.6	Supplementary Material	81
II	Contour-based Methods	91
4	Asymmetric Graph Cut for Cell Segmentation in Phase Contrast Images	93
4.1	Introduction	93
4.1.1	Related work	95
4.2	Methods	95
4.2.1	Cell segmentation	95
4.2.2	Cell tracking	96
4.3	Experiments	99
4.3.1	Dataset	99
4.3.2	Implementation details	99
4.3.3	Evaluation	99
4.4	Conclusion and Outlook	104
5	Contour-based Motion Pattern Analysis of Migrating Cells	105
5.1	Introduction	105
5.1.1	Related Work	106
5.2	Detection of Symmetry-Breaking Events	107
5.2.1	Calculation of Eccentricity	107
5.2.2	Detection	107
5.3	Cell Contour Protrusion/Retraction Analysis	109
5.3.1	Contour Extraction and Normalization	109
5.3.2	Calculation of Contour Velocities	109
5.3.3	Protrusion/Retraction Measure	110
5.4	Experiments	110
5.4.1	Datasets	111
5.4.2	Implementation Details	111
5.4.3	Quality Control	112
5.4.4	Results	112
5.5	Conclusions	121
6	Conclusion	123
	Bibliography	127

1 Introduction

1.1 Motivation

The Biomedical Task and Setting

In current biomedical research developmental processes are investigated on the cellular level using 3D time-lapse microscopy. Living specimen are recorded *in vivo* in 3D space and processes such as cell division and cell migration, that lead to the formation of tissue, the development and growth of organs and finally the development of whole organisms, can be observed at the cellular level. The technological advances in 3D time-lapse microscopy, in particular light sheet microscopy (Keller et al., 2008, 2010; Tomer et al., 2012), are impressive. They revealed the early development of zebrafish (*Danio rerio*) and fruit fly (*Drosophila*) embryos at the cellular level. As an example, figure 1.1 depicts the complex cell dynamics in the early development of a zebrafish embryo. Studying developmental processes and cell migration can help researchers understand the underlying mechanisms and lead to the development of new therapies for diseases that occur when some of these mechanisms fail. Huge amounts of 3D+time data, *i.e.* time sequences of 3D volumetric images, are generated in such experiments. Since a manual evaluation is usually neither possible nor desired, techniques for automated image analysis have become indispensable.

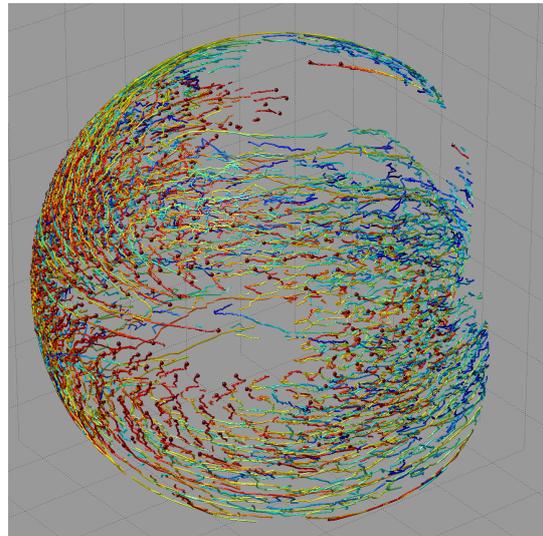


Figure 1.1: Complex biological motion pattern of cell dynamics in the early development of a zebrafish embryo. 3D rendering of cell motion trajectories with time color-coded (from *blue* to *red*).

To answer biomedical questions the observed phenomena have to be measured precisely in a robust and unbiased manner that yields a solid quantitative evaluation. Beyond the quantification of single samples, *i.e.* the development of single specimen, the comparison of multiple samples or groups of samples is required to draw biomedically relevant conclusions. Usually the *wild-type* specimen, which represent the original natural state, form the reference group to which groups of genetically manipulated specimen (*mutants*) are compared. The idea is to link observable changes in the manipulated specimen, *i.e.* changes in the *phenotype*, to the underlying genetic manipulation, *i.e.* the manipulation of the *genotype*. Using the link between genotype and phenotype, the final goal is to infer the role and function of specific genes. High throughput screening experiments present an efficient way to test for the effect of a large number of drugs. Screening allows to quickly detect drugs or modifications affecting the phenotype of different target groups. However, automatic localization and quantitative description of the differences requires high-content imaging and remains a very challenging task.

Challenges for Automated Image Analysis and Quantitative Comparisons

Properties of the Imaging Technique

Several issues make this task challenging, first of all, the properties of the imaging technique that is at the interface between biology, or medicine, and image analysis. In 3D microscopy the point spread function (PSF) characterizes the impulse response of the optical system and results in an anisotropic resolution which is reduced along the optical axis. Other optical effects such as light absorption, refraction and scatter occur together with different kinds of noise introduced at various stages of the image acquisition pipeline. When it comes to the comparison of different recordings additional issues arise. The use of different imaging settings for example introduces artificial variations that can bias the comparison. In certain scenarios however the use of different imaging settings is unavoidable, which becomes particularly critical in terms of quantitative comparisons if the used settings correlate with the different groups to be compared. This may introduce a systematic bias not related to any actual difference. Image analysis techniques are required to be robust not only to varying imaging settings but ideally operate robust across different microscopes, different imaging modalities, and in general be imaging device independent.

Amount of Training Data

Another limitation in biomedical applications, especially with 3D time-lapse microscopy, is that usually only few data samples are available, which is due to the time-consuming sample preparation and recording. This makes the problem difficult for learning-based approaches that rely on a sufficient amount of training data to capture variability. Moreover, creating ground truth annotations, that are required for supervised learning, can be very time-consuming, *e.g.* for 3D segmentation tasks. Consequently, the biomedical field, in contrast to the field of computer vision, cannot rely on large public image databases with annotations. Compared to standardized medical recording devices, in biology an ever increasing zoo of experimental microscopes emerges, leading to an amazing data variety in terms of content and quality. Additionally, often annotation quality of the few available public datasets is insufficient for supervised learning.

Distinguishing Relevant from Irrelevant Variations

A main challenge in pattern recognition is to distinguish *relevant* from *irrelevant* variations in the data. The same holds for quantitative comparisons and the determination of differences, which requires approaches that are invariant or robust to all irrelevant variations, such that only the relevant variations, *i.e.* discriminative features, remain that can be measured and compared. Relevance of variations is dependent on the specific task and the *prior knowledge* available or given by the expert. The kind of prior knowledge usually determines the design of the image analysis algorithm. In the biomedical setting, apart from irrelevant variations caused by the image acquisition process, common irrelevant variations include location, orientation or slight elastic deformation of the recorded specimen. *Normalization* is one general approach to achieve invariance. It comprises *detection* and *registration*, to compensate for spatial transformations, such as shift, rotation, scaling or elastic deformation. For time-lapse data, detection also incorporates the temporal dimension and, similar to elastic registration in space, dynamic time-warping allows to register time sequence information. After normalization the aligned datasets can be compared directly by some measure and the differences can be localized. *Localization* involves both, space and time, *e.g.* spatially w.r.t. the anatomy, and temporally w.r.t. the course of development of the organism. Most often differences must be localized relative to the anatomy, *e.g.* using anatomical coordinates or mapping to an atlas. Anatomical structures can be obtained by *segmentation* or fitting a model to the data. Localization alone however gives no clue about the nature of the differences. Instead a *model* is required that expresses the differences in meaningful parameters, that can be measured. If,

for example, the only difference is an increased speed of development then differences would be detected at many locations, but just a model that considers temporal transformations could explain the differences in a meaningful way.

Amount of Prior Knowledge - Three Scenarios

In the following, we give three examples with decreasing degree of prior knowledge, which makes the tasks increasingly difficult. The case that a hypothesis about the difference is a priori known or not is also included.

1. **Hypothesis and prior knowledge available:** In this example, the early development of zebrafish is investigated using 3D+time data showing the cell dynamics of wild-type and mutant embryos. A specific hypothesis is a priori available, namely that differences exist in the dynamic behavior of cells that perform cell intercalations, *i.e.* cells that enter the space between adjacent cells. Mainly two kinds of prior knowledge were used: 1) Since only the dynamic behavior of cells is of interest, a motion representation by cell trajectories was used. 2) Since only cell intercalations are in the focus, cell intercalation events were detected and analyzed. This scenario is described in detail in chapter 3.
2. **No hypothesis, but prior knowledge available:** This example also investigates the early zebrafish development using 3D+time data showing the cell dynamics of wild-type and morphant embryos. However, no specific hypothesis about differences is a priori available, apart from that only the dynamic behavior of cells is of interest. In this scenario, a motion representation by cell trajectories was used. The task was formulated more generally as an *anomaly detection* problem. A precise generative model of the commonly observed variations in the normal class, *i.e.* the wild-type group, is learned. Any deviation of a test sample that goes beyond the learned model will be detected as an anomaly, *i.e.* as a significant difference. The approach is more generally applicable and introduced in chapter 2.
3. **No hypothesis and no prior knowledge available:** This scenario states the most difficult case where neither a hypothesis about the difference nor any prior knowledge about the task is available. In this setting, only the raw data of the different groups is available. The possibility of addressing this universal task with a general learning approach is discussed in section 1.3 and in the outlook of this thesis in section 6.

Focus on Motion Patterns

This thesis focuses on *motion patterns* in particular, as motion, apart from the appearance, is an important means to capture temporal dynamics and evolution. It allows recognizing actions or detecting special events, for example, in the context of human motion or patterns in a traffic scene. Motion in the biomedical context can be considered from the macroscopic level, such as *e.g.* the motion of a mouse down to the microscopic level, where *e.g.* the motion of single cells is observed. Here, we consider *biomedical motion patterns* at the microscopic level, where cell migration is an essential mechanism that drives the development and growth of organisms. Using motion as the primary feature yields a clear advantage: The representation is inherently invariant, or at least robust, to appearance changes, due to different microscopy settings, different imaging devices or even different imaging modalities. Thus, the motion representation introduces prior knowledge into the task that helps to obtain robust solutions.

There are two principle ways to represent motion as illustrated in figure 1.2, at the top. Both originate from fluid dynamics in physics and differ in the point of view of the observer. In the *Lagrangian* representation the observer is located on the particles that flow through space driven by the flow field. Particle trajectories $\mathbf{X}(\mathbf{x}_0, t)$ are functions of a reference location \mathbf{x}_0 at reference time t_0 and time t .

In the *Eulerian* representation the observer is located at a fixed position \mathbf{x} , where the observed flow $\mathbf{u}(\mathbf{x}, t)$ is a function of location \mathbf{x} and time t . While the Lagrangian case yields a dynamic representation using trajectories the Eulerian case yields a stationary representation of the flow observed at fixed locations.

This thesis focuses on the Lagrangian representation of motion based on trajectories and is further divided into two parts. **Part I** deals with *trajectory-based* representations, as illustrated in figure 1.2 at the bottom left. In the example, the motion of three adjacent cells is illustrated using trajectories, where the central cell enters the space between the adjacent cells and performs a cell intercalation. This is an example of a *sparse* trajectory representation, where single objects, in this case cells, are represented by single trajectories. We apply this representation for motion patterns of cell nuclei in zebrafish, both for the detection of cell intercalations in chapter 3, and for anomaly detection in section 2.7.3. If in contrast to individual object tracking the motion pattern of the whole scene is of interest, a *dense* trajectory representation can be used. In this case each point on a regular grid on the image can be tracked. We initialize our representation by dense point trajectories for anomaly detection in chapter 2. Tracking can be a very challenging task and is thus not reliably possible in all scenarios. However, once trajectories are extracted they serve as a very robust, efficient and rich representation able to encode even long term motion information. Tracking algorithms themselves are not subject of this thesis, we rather use available techniques to obtain trajectories. Dense point trajectories were computed by large displacement optical flow (Sundaram et al., 2010), while trajectories of cell nuclei in zebrafish were provided by the biologists directly.

Part II is concerned with *contour-based* representations, as illustrated in figure 1.2 at the bottom in the middle. The example shows the evolving contour of a cell that changes its shape while starting to move. Six contour points are marked and connected to the corresponding points in the next frame. The contour-based representation is based on a closed contour that segments the object of interest, e.g. the cell, and the association of contour points across time. The motion pattern represented is the evolution of a contour. It is a special case of the trajectory-based representation with trajectories sampled only along the contour, plus the spatial connectivity and ordering a contour implies.

Grid-based representations, as illustrated in figure 1.2 at the bottom right, correspond to the Eulerian representation of motion. The example depicts a cell dividing into two compartments, cell contours are only included for better visualization. The vectors of the flow field are indicated for each grid cell: There is no flow outside the cell, randomly oriented flow inside the cell, and reduced flow at the evolving inner cell boundary. During my work I experimented with the grid-based representation using dense optical flow and was able to gain promising preliminary results in an application to intra-cellular motion patterns of growing and dividing plant cells (*Protoplasts*). Grid-based methods are however not part of this thesis.

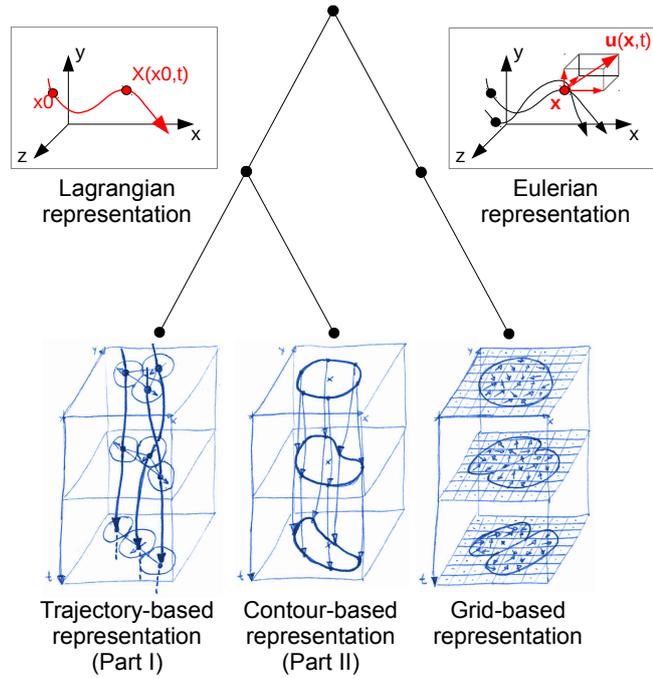


Figure 1.2: Representations of motion patterns. At the top the two principle representations of motion are shown: *Lagrangian* and *Eulerian* representation. At the bottom the representations of motion patterns considered in this thesis are illustrated. Grid-based methods are not part of this thesis.

Goal

In conclusion, the goal of this thesis is to develop approaches that allow robust quantification and comparison of motion patterns in the context of biomedical applications and specifically 2D and 3D time-lapse microscopy. Robustness in terms of microscopy requires device independent approaches. The restricted number of samples, especially in 3D time-lapse microscopy, requires approaches that can cope with a small number of training samples. Quantitative comparison requires precise measurements, possibly within an anatomical reference coordinate system. Importantly, the comparison should, on the one hand, yield only the relevant differences and, on the other hand, represent differences in a meaningful and interpretable way, such as physical units or w.r.t. the parameters of a specific model.

1.2 Contributions

This thesis presents methods that enable motion pattern analysis in complex biomedical applications from 3D+time data, specifically 2D and 3D time-lapse microscopy. The developed methods enable precise measurements for quantitative analysis and robust comparisons of motion patterns, which is essential for evaluating experiments in biomedical research and specifically developmental biology. The approaches have been successfully applied to real world problems within several collaborations and have been published in the area of computer vision (Bensch et al., 2017, 2015), as well as in the field of biology (Saltukoglu et al., 2015; Bensch et al., 2013) and life sciences (Ulman et al., 2017) and at the intersection of image analysis, imaging and biomedical applications (Bensch and Ronneberger, 2015). The thesis divides into two parts, trajectory-based and contour-based methods.

Trajectory-based Methods

Spatiotemporal Deformable Prototypes for Motion Anomaly Detection

We developed a general method that enables motion anomaly detection in 3D+time data (Bensch et al., 2017, 2015). It is based on a new efficient and robust representation by “supertrajectories”. An efficient hashing approach and a new method for elastic registration of 3D+time trajectory patterns allows placing spatiotemporally deformed instances of a prototype pattern to reconstruct a whole test pattern. A prototype model is learned from multiple sequences to define accepted variations. Anomalies are detected in the test pattern as poorly reconstructed patterns, which show strong deviations from the elastically registered prototype patterns. Unlike most existing methods, our method enables processing of 3D+time data, as opposed to 2D+time data, features anomaly detection in the context of a prototype model and explicitly models spatiotemporal deformations. The method performs well in detecting subtle anomalies on a new motion anomaly dataset of juggling patterns and outperforms an anomaly detection approach based on chaotic invariants (Wu et al., 2010). The strength of our method is that it can cope with few training samples while still generalizing well to patterns in different 3D position, orientation and with spatiotemporal deformations. Motion anomaly detection is highly relevant to the comparison of biological motion patterns. We demonstrated the applicability of our approach on the quantitative comparison of motion patterns from the early development of zebrafish embryos (Bensch et al., 2017, 2015). As a result, we obtained a detailed spatiotemporal analysis of differences between wild-type and morphant embryos (genetically modified).

Spatiotemporal Geometrical Models for Motion Pattern Detection

We developed a method to detect specific motion patterns in 3D+time data using spatiotemporal geometric models. In particular, we developed a model to detect an essential pattern in developmental biology referred to as *cell intercalation* (Bensch et al., 2013). Cell intercalations occur when cells enter the space between adjacent cells and play an important role in tissue formation. The method builds on

motion trajectories of single cells and models the motion pattern to be detected using a spatiotemporal geometrical model. The presented method enables the robust detection of cell intercalations in real 3D+time biological data. It is robust to noisy and interrupted measurements and handles the variability within the class of 3D intercalations. We successfully applied our method to obtain quantitative comparisons of cell intercalations and their motion statistics between wild-type and mutant embryos in the early development of zebrafish (Bensch et al., 2013). While the biological results and contributions are reported in Bensch et al. (2013), this thesis in addition elaborates on the methods for cell intercalation detection.

Contour-based Methods

Asymmetric Graph Cut for Cell Segmentation in Phase Contrast Images

We developed a new robust, effective, and surprisingly simple approach for the segmentation of cells in phase contrast microscopy (Bensch and Ronneberger, 2015). Phase contrast microscopy generates strong intensity gradients along interfaces of media with different physical densities. They allow to obtain clear boundary responses even for perfectly transparent samples. However, classical edge-based image segmentation fails due to the complex intensity profile consisting of a bright-to-dark and a dark-to-bright transition at the boundary and other artifacts from phase contrast microscopy. The key feature of our algorithm is that it strongly favors dark-to-bright transitions at the boundary of (arbitrarily shaped) segmentation masks. This exploits the appearance of true cell borders in positive phase contrast microscopy which form a dark-to-bright transition in outwards direction. The segmentation mask can be effectively found by a fast min-cut approach. The small but essential difference to standard min-cut based approaches is that our graph contains directed edges with asymmetric edge weights, which has never been applied to phase contrast microscopy, to the best of our knowledge. Our evaluation on phase contrast datasets from the ISBI Cell Tracking Challenge (CTC) (ISB) shows that asymmetric edge weights yield better results while being less sensitive to the selected graph cut parameters. Our method outperforms the top ranked methods from the ISBI CTC 2014 on the phase contrast dataset and reached second place, winning against the previously 1st ranked method, in the ISBI CTC 2015 (Bensch and Ronneberger, 2015; Ulman et al., 2017; ISB). The accurate segmentation of cells is an important basis for further analysis and quantification in biomedical applications.

Contour-based Motion Pattern Analysis of Migrating Cells

We successfully applied our approach for cell segmentation in phase contrast images to analyze cell migration and cell contour motion patterns. We developed a method that, based on the extracted cell contours from 2D time-lapse data, detects so-called *symmetry-breaking* events. This enabled automatic browsing of large amounts of data to focus on these cellular events of interest. To investigate motion patterns along the cell contour, we implemented a method to compute so-called protrusion/retraction maps. These visualize the (signed) contour velocities in a 2D map with polar contour coordinates on the y-axis and time on the x-axis. Our approach enabled quantitative comparisons of several experimental conditions and we contributed to the biological results published in Saltukoglu et al. (2015) and contained in Saltukoglu (2015).

1.3 Challenges for Deep Learning

In the light of today's success of deep learning (LeCun et al., 2015; Rusk, 2016), the question may arise whether the presented approaches in this thesis are still up-to-date. To answer this question, we first observe that current deep learning approaches, both in computer vision and in biomedical image analysis, solve rather "basic" tasks, such as classification, detection, segmentation and registration.

Tracking, specifically multi-target tracking, is a more difficult problem as it combines several tasks. Solving the problem *end-to-end* is a desired objective in deep learning. It means that only the raw input data and a ground truth for the final output data must be given and the deep learning approach will learn “everything in between” as a kind of black-box.

However, to date no deep learning approach has tackled the complex tasks presented in this thesis in an end-to-end fashion. The tasks at hand require a quantitative comparison of groups of 4D datasets (3D+time volumetric data) based on precise physical measurements. Moreover, beyond a localization of the differences, the comparison should yield a meaningful description of the differences and often the assessment of statistical significance is required. Several reasons make the tasks at hand difficult for deep learning approaches:

- The tasks at hand usually consist of several subtasks. Compared to conventional approaches that solve the problem step-by-step, deep learning as an end-to-end “black-box” approach has the disadvantage that it is hard to understand the learned internal function and to access intermediate results, which hampers engineering and development of complex approaches.
- In our setting training data is scarce, which is a problem for deep learning as it relies on large amounts of training data. This problem is even more severe with 3D+time data, as the number of possible variations is significantly increased.
- Usually a lot of prior knowledge about the task is available and required for the analysis, *e.g.*: Where is the focus of the analysis? What to measure and compare? Is the relation to anatomical structures important? Which spatial or temporal transformations are irrelevant, or which might be relevant? Without incorporating any prior knowledge it might be very hard to achieve the desired result with deep learning, or even to obtain a reasonable training network. While explicitly embedding prior knowledge through network design can be difficult, prior knowledge can be induced implicitly in an elegant way using data augmentation by *e.g.* sampling from and applying irrelevant transformations. However, including very specific prior knowledge, *e.g.* what to measure and compare exactly, might be very difficult and moreover deviates from the idea of end-to-end learning.
- Achieving robustness to irrelevant variations, especially fulfilling the demand for device-independent approaches (imaging settings, imaging modalities, microscopes) is hard because of the restricted amount of training data. Additionally, even if a sufficient amount of training data containing *e.g.* variations of the imaging setting is available, the behavior and result for a new dataset with so far unseen imaging settings is unpredictable. Recent domain adaptation approaches may help to solve this issue. [Kamnitsas et al. \(2016\)](#) showed promising results for transferring knowledge between different CT scan modes using adversarial networks. In our approaches in part I, we directly build upon a motion representation by trajectories, which basically neglects the appearance and is thus inherently robust to changes of the appearance. Building upon a motion representation is an example for including prior knowledge, without which the task for an end-to-end deep learning approach would be much more difficult from the beginning.
- Localization of the differences in the input data space (sequence of images) is possible with deep learning, however obtaining a meaningful description of differences and especially the assessment of statistical significance is difficult with deep learning (see section “Pitfalls” in [Angermueller et al. \(2016\)](#)).
- The huge amount of data that has to be processed for comparing groups of 4D volumetric datasets might pose a practical challenge for applying deep learning in terms of memory consumption and training time.

In conclusion, applying deep learning in an end-to-end learning fashion for quantitative comparisons in the complex biomedical tasks at hand poses several challenges to be addressed. For complex tasks it might even be hard to achieve a solution using an end-to-end deep learning approach. Therefore, the approaches in this thesis are still up-to-date and present state-of-the-art. One exception is the cell segmentation approach presented in chapter 4. This rather “basic” task has been outperformed by the U-net (Ronneberger et al., 2015), which is a convolutional neural network (CNN) for biomedical image segmentation. We want to note that the U-net relies on a sufficiently large amount of training data and data augmentation techniques, while our approach (Bensch and Ronneberger, 2015) requires only very few training data.

The conclusion gives an outlook on future research and discusses perspectives of deep learning for the tasks considered in this work.

1.4 Structure

This thesis is organized into two parts, trajectory-based methods (chapter 2 and chapter 3) and contour-based methods (chapter 4 and chapter 5). Chapter 2 presents our approach to motion anomaly detection using spatiotemporal deformable prototypes. It first introduces the supertrajectory representation and then focuses on efficient detection and the new method for elastic registration of motion trajectory patterns. Based on this, our approach to motion anomaly detection is described and extensive experiments are presented. Chapter 3 introduces our approach to motion pattern detection using spatiotemporal geometric models. The cell-triple intercalation model and our approach to detect cell intercalations is described, followed by the experiments. Contour-based methods first present our approach for cell segmentation in phase contrast images using asymmetric graph cut in chapter 4. First, the properties of images from phase contrast microscopy are introduced to motivate the use of asymmetric edge weights. The description of cell segmentation using min-cut with asymmetric costs and cell tracking is followed by the experiments, including results on the ISBI Cell Tracking Challenge (ISB; Ulman et al., 2017). Chapter 5 describes our method for contour-based motion pattern analysis of migrating cells. Cell contours are based on the cell segmentation approach presented in chapter 4. We first introduce our method to detect symmetry-breaking events, then describe how cell contour protrusion/retraction analysis is performed, followed by the experiments. In chapter 6 we summarize the results of the thesis, draw our conclusions and provide an outlook on future research.

1.5 Publications

Main parts of this thesis have been published, either in the form of journal articles or as conference papers.

Journal Articles

- R. Bensch, N. Scherf, J. Huisken, T. Brox, and O. Ronneberger. Spatiotemporal deformable prototypes for motion anomaly detection. *International Journal of Computer Vision*, 122(3):502–523, May 2017
- R. Bensch, S. Song, O. Ronneberger, and W. Driever. Non-directional radial intercalation dominates deep cell behavior during zebrafish epiboly. *Biology Open*, 2(8):845–854, 2013
- V. Ulman, M. Maška, K. E. G. Magnusson, O. Ronneberger, C. Haubold, N. Harder, P. Matula, P. Matula, D. Svoboda, M. Radojevic, I. Smal, K. Rohr, J. Jaldén, H. M. Blau, O. Dzyubachyk, B. Lelieveldt, P. Xiao, Y. Li, S.-Y. Cho, A. Dufour, J. C. Olivo-Marin, C. C. Reyes-Aldasoro, J. A. Solis-Lemus, R. Bensch, T. Brox, J. Stegmaier, R. Mikut, S. Wolf, F. A. Hamprecht, T. Esteves, P. Quelhas, Ö. Demirel, L. Malmström, F. Jug, P. Tomancák, E. Meijering, A. Muñoz-Barrutia, M. Kozubek, and C. O. de Solor. An objective comparison of cell-tracking algorithms. *Nature Methods*, 14:1141–1152, 2017
- D. Saltukoglu, J. Grünwald, N. Strohmeyer, R. Bensch, M. Ulbrich, O. Ronneberger, and M. Simons. Spontaneous and electric field-controlled front-rear polarization of human keratinocytes. *Molecular Biology of the Cell*, 26(24):4373–4386, December 2015

Conference Papers (peer-reviewed)

- R. Bensch, T. Brox, and O. Ronneberger. Spatiotemporal deformable prototypes for motion anomaly detection. In *British Machine Vision Conference (BMVC)*, pages 189.1–189.12, September 2015. (oral presentation)
- R. Bensch and O. Ronneberger. Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1220–1223, April 2015. (oral presentation)

This thesis does not report on the following publications that have been published during my time as a research assistant under the supervision of Prof. Olaf Ronneberger, Prof. Hans Burkhardt and Prof. Thomas Brox.

Journal Articles

- A. M. Middleton, C. D. Bosco, P. Chlap, R. Bensch, H. Harz, F. Ren, S. Bergmann, S. Wend, W. Weber, K. Hayashi, M. D. Zurbriggen, R. Uhl, O. Ronneberger, K. Palme, C. Fleck, and A. Dovzhenko. Data-driven modeling of intracellular auxin fluxes indicates a dominant role of the er in controlling nuclear auxin uptake. *Cell Reports*, 22(11):3044–3057, March 2018
- T. Yasunaga, S. Hoff, C. Schell, M. Helmstädter, O. Kretz, S. Kuechlin, T. Yakulov, C. Engel, B. Müller, R. Bensch, O. Ronneberger, T. Huber, S. Lienkamp, and G. Walz. The polarity protein intuned links nphp4 to daam1 to control the subapical actin network in multiciliated cells. *The Journal of Cell Biology*, 211(5):963–973, 2015
- B. Greese, K. Wester, R. Bensch, O. Ronneberger, J. Timmer, M. Hülskamp, and C. Fleck. Influence of cell-to-cell variability on spatial pattern formation. *IET Systems Biology*, 6(4):143–153, 2012

Conference Papers (peer-reviewed)

- M. Keuper, R. Bensch, K. Voigt, A. Dovzhenko, K. Palme, H. Burkhardt, and O. Ronneberger. Semi-supervised learning of edge filters for volumetric image segmentation. In *Pattern Recognition (Proc. DAGM)*, LNCS, pages 462–471. Springer, 2010
- R. Bensch, O. Ronneberger, B. Greese, C. Fleck, K. Wester, M. Hülkamp, and H. Burkhardt. Image analysis of arabidopsis trichome patterning in 4d confocal datasets. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 742–745, June 2009

1.6 Collaborations

Parts of the work for this thesis was done in collaboration with others. To clarify the attribution of the work described in this thesis, collaborations beyond the supervision of my Ph.D. advisor Prof. Olaf Ronneberger, the supervision of Prof. Thomas Brox and the discussions with colleagues is detailed below.

- The experiments on anomaly detection in juggling patterns in section 2.7.2 were enabled by Julia Koch, Andreas Krämer, Tobias Paxian, Dominic Mai and Olaf Ronneberger who contributed their juggling expertise and agreed to perform diverse juggling patterns in front of our Kinect camera.
- The experiments on anomaly detection in biological motion patterns in section 2.7.3 originate from the collaboration with Nico Scherf (Institute for Medical Informatics and Biometry, TU Dresden) and Jan Huiskens (Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG), Dresden) who provided the biological task together with the trajectory data from Schmid et al. (2013) and any further information we requested.
- The topic described in chapter 3 was initiated by a collaboration with Sungmin Song and Prof. Wolfgang Driever (Developmental Biology, Institute Biology I, Faculty of Biology, University of Freiburg) who provided the biological task, datasets and experiments and the biological expertise. While the author performed all computational analysis, the preparation of the figures and the writing of the manuscript with reference to Bensch et al. (2013) was performed by all the coauthors, including the author of this thesis, Sungmin Song, Prof. Olaf Ronneberger and Prof. Wolfgang Driever. Especially this concerns the experiments in section 3.4, and specifically the results and the biological discussion in section 3.4.4 and section 3.4.5, that require biological expertise.
- The topic described in chapter 5 originates from a collaboration with Deniz Saltukoglu (Center for Systems Biology, University of Freiburg and Renal Division, University Hospital Freiburg) and Matias Simons (Center for Systems Biology, University of Freiburg; Renal Division, University Hospital Freiburg and Imagine Institute, Paris Descartes University-Sorbonne Paris Cité) who provided the biological task, datasets and experiments and the biological expertise. The biological results in section 5.4.4 and a compressed version of the methods about segmentation, detection of symmetry-breaking events and protrusion/retraction analysis are published in Saltukoglu et al. (2015) and are contained in the PhD thesis by Deniz Saltukoglu (Saltukoglu, 2015).

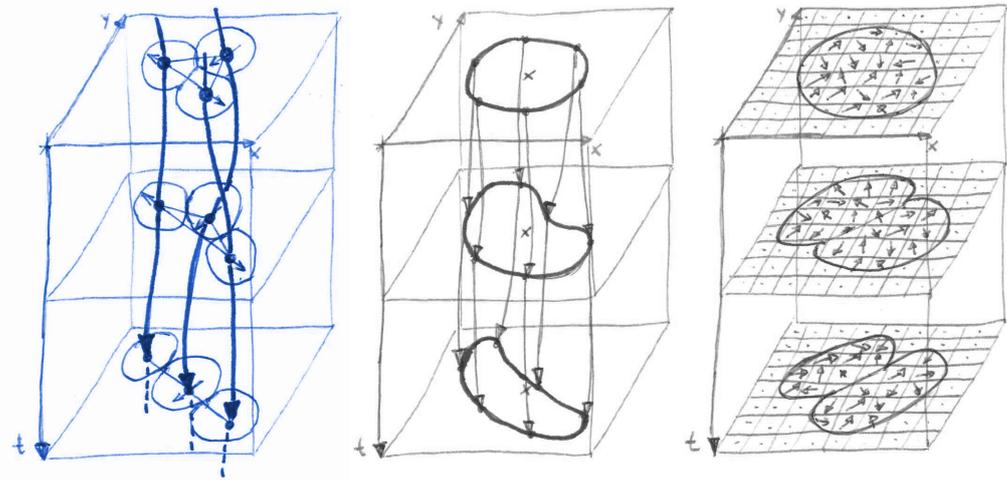
1.7 Notation

The following notations are used throughout this thesis. In most cases, fields and matrices are represented by capital letters, while single values or vectors are represented by lower case letters. Further notations in the thesis are introduced directly on their first appearance.

\mathbb{N}	Set of natural numbers (positive integers)
\mathbb{R}	Set of real numbers
\mathbb{C}	Set of complex numbers
$x \in \mathbb{R}, \mathbb{C}$	Scalar real or complex value
$\mathbf{x} \in \mathbb{R}^n, \mathbb{C}^n$	N-dimensional vector of real or complex values
$\mathbf{X} \in \mathbb{R}^{n \times m}, \mathbb{C}^{n \times m}$	Matrix of real or complex values
$\mathbf{x}^T, \mathbf{X}^T$	Vector-, Matrix transpose
$\ \cdot\ $	L_2 norm of a vector
$ \cdot $	Absolute value
$\langle \mathbf{x}, \mathbf{y} \rangle$	Inner product of \mathbf{x} and \mathbf{y}
$\mathbf{x} \times \mathbf{y}$	Cross product of \mathbf{x} and \mathbf{y}
X	N-dimensional scalar field (<i>e.g.</i> , a gray-valued image)
\mathbf{X}	N-dimensional vector field (<i>e.g.</i> , a gradient field)
$X(\mathbf{x})$	Scalar value (<i>e.g.</i> , an intensity value) at position \mathbf{x} of scalar field X
$\mathbf{X}(\mathbf{x})$	Vectorial value (<i>e.g.</i> , a gradient) at position \mathbf{x} of vector field \mathbf{X}
$\nabla X = \left(\frac{\partial X}{\partial x_1}, \dots, \frac{\partial X}{\partial x_n} \right)^T$	Gradient operator on scalar field $X : \mathbb{R}^n \rightarrow \mathbb{R}^n$
$\nabla X(\mathbf{x})$	Gradient of scalar field X at position \mathbf{x}
$f : X \rightarrow Y : x \mapsto y$	Function that maps an element in domain X , $x \in X$, to an element in codomain Y , $y \in Y$
$f : \mathbb{R}^n \rightarrow \mathbb{R}^m : \mathbf{x} \mapsto \mathbf{y}$	Function that maps a point in n-dimensional real coordinate space, $\mathbf{x} \in \mathbb{R}^n$, to a point in m-dimensional real coordinate space, $\mathbf{y} \in \mathbb{R}^m$
$f[a, b](x, y)$	Function with variables x and y , and parameters a and b
\hat{f}, \hat{a}	Estimate of function f , estimate of parameter a
Ω	Domain of a function
\exp	Exponential function
\mathbf{R}	Rotation matrix
$\Omega \subset \mathbb{N} \times \mathbb{R}$	Motion pattern domain, with domain of trajectories $i \in \mathbb{N}$ and time $t \in \mathbb{R}$
$\mathbf{x}(i, t) : \Omega \rightarrow \mathbb{R}^3$	Motion pattern position function. $\mathbf{x}(i, t)$ yields the position of trajectory i at time t
$w(i, t) : \Omega \rightarrow \{0, 1\}$	Motion pattern validity function. $w(i, t) = 1$ means that trajectory i exists at time t and the position is valid
$\mathbf{u}(i, t) : \Omega \rightarrow \mathbb{R}^3$	Motion pattern spatial deformation function. $\mathbf{u}(i, t)$ yields the spatial deformation vector of trajectory i at time t
$\tau(i, t) : \Omega \rightarrow \mathbb{R}$	Motion pattern temporal warping. $\tau(i, t)$ yields the temporal warping of trajectory i at time t
$\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt}$	Time derivative, short-hand notation
$\dot{\mathbf{x}}(i, t)$	Motion pattern velocity vector of trajectory i at time t

Part I

Trajectory-based Methods



2 Spatiotemporal Deformable Prototypes for Motion Anomaly Detection

In this chapter, we present an approach for detecting motion anomalies. We formulate the problem as a reconstruction task, in which a test sample is explained using a prototype pattern. To achieve this, the prototype pattern is detected and elastically registered against a test sample to detect anomalies in the test sample. The prototype model is learned from multiple sequences to define accepted variations. We rely on dense point trajectories as the underlying motion representation. From these, “supertrajectories” are computed based on hierarchical clustering that serve as an efficient and robust representation of motion patterns. An efficient hashing approach provides transformation hypotheses that are refined by a spatiotemporal elastic registration. We propose a new method for elastic registration of 3D+time trajectory patterns that induces spatial elasticity from trajectory affinities. Finally, anomalies are detected in the test sample as poorly reconstructed patterns, which show strong deviations from the elastically registered prototype patterns.

This task is highly relevant to the comparison of biological motion patterns, in which typically the wild-type pattern (normal behavioral pattern) is compared to a pattern resulting from genetically modified specimen (*e.g.* mutants). Usually the task is to quantify whether the classes are significantly different, and if this is the case, moreover to localize the deviations. The method is evaluated on a new motion anomaly dataset of juggling patterns and performs well in detecting subtle anomalies. Moreover, we demonstrate the applicability to biological motion patterns.

The material presented here has been published in the International Journal of Computer Vision (Bensch et al., 2017) and presented at the British Machine Vision Conference (Bensch et al., 2015) in a previous version.

2.1 Introduction

An anomaly is generally a deviation from what is regarded as normal. Since there are no examples from which distinct features of the anomaly could be learned, anomaly detection cannot be modeled

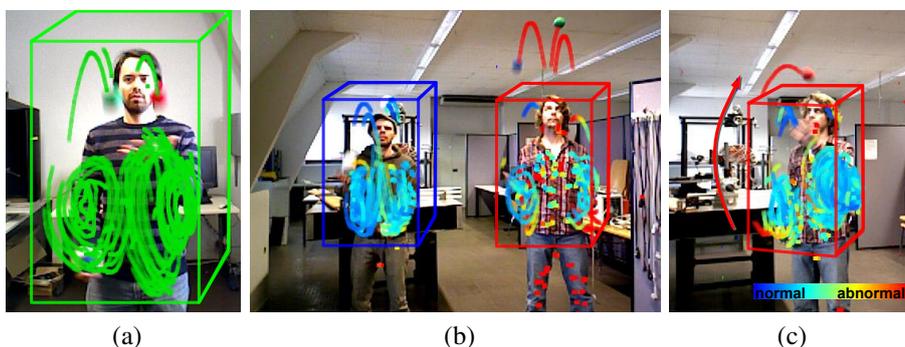


Figure 2.1: Motion anomaly detection in juggling patterns. **(a)** Motion pattern prototype: Standard 3-ball cascade pattern. **(b-c)** Detection of local anomalies in context of prototype detections (*bounding boxes*) of different jugglers in different 3D poses. The anomaly score is plotted for supertrajectories from low/normal (*blue*) to high/abnormal (*red*).

as a discriminative classification task. We must rather learn a precise generative model of normal patterns, for which we have examples, and detect anomalies as cases that are not sufficiently explained by this model. In this work, we consider motion-based anomaly detection from video and present a new setting of anomaly detection. Compared to existing approaches in the literature, we formulate anomaly detection as the task of detecting subtle anomalies in the context of a well-defined, reproducible motion pattern, termed motion pattern prototype. We want to detect instances of this prototype and localize anomalies in its context. Anomaly detection is performed by reconstructing an unseen motion pattern by prototype placements. Subpatterns that remain poorly reconstructed are detected as abnormal. For reconstruction, we start with a robust detection followed by a spatiotemporal elastic registration of a deformable prototype. Our method copes with 3D+time data, while existing methods deal with 2D+time data. Specifically, in our setting *3D+time* data denotes motion trajectories in 3D space. For detection, we allow for temporal shift and 3D spatial translation and rotation. 3D scaling is not modelled explicitly and therefore will be recognized anomalous in principle. However, elastic registration is able to compensate for scaling to some extent. A statistical prototype model is learned from training data. It defines the accepted spatiotemporal deformations and deviations. Figure 2.1 illustrates our setting of anomaly detection with an example of juggling patterns. In this example, the prototype defines a standard 3-ball juggling pattern. It is robustly detected under various transformations. Deviations from the standard pattern are localized as anomalies. We stress that, in this work, we are interested in anomalies in the motion pattern rather than the object appearance. For example, we do not want to detect an anomaly, if a “normal” motion pattern is performed but the person is wearing a different shirt.

The presented material extends a preliminary version that has been published in [Bensch et al. \(2015\)](#). Mainly, the experimental evaluation has been extended considerably. Experiments on biological motion patterns are added, together with extended evaluation of anomaly localization, elastic registration and supertrajectory representation that give additional insights. Moreover, the method description has been extended and made more comprehensible, in particular the detection and elastic registration of motion patterns. Furthermore, the section on related work has been extended and mathematical formalisations have been simplified, among many other minor improvements.

An overview of our method is given in figure 2.2. In the next section, we first summarize the related work. Then the main parts of our approach, namely the supertrajectory representation, the detection and elastic registration of motion patterns and the anomaly detection are explained. The experimental section provides basic validation experiments of elastic registration and evaluates the performance of our approach on a new motion anomaly dataset. Moreover, we demonstrate the relevance and general applicability of our method in experiments on biological motion patterns. While in the motion anomaly dataset 3D+time data originates from Kinect 2D+depth data, the biological motion patterns stem from 3D volumetric microscopy recordings.

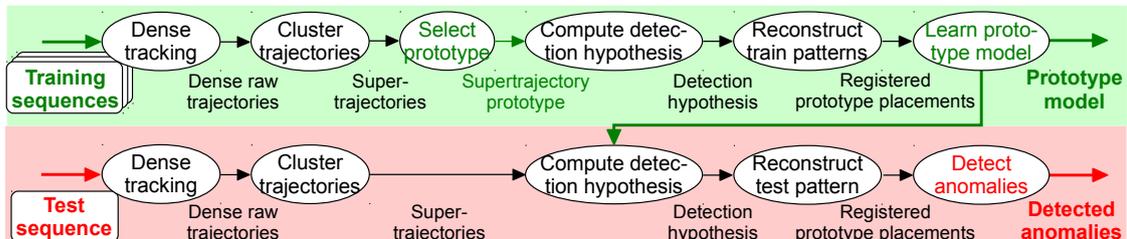


Figure 2.2: Overview of our approach.

2.2 Related Work

In general, “*anomaly detection* refers to the problem of finding patterns in data that do not conform to expected behavior” (Chandola et al., 2009). Accordingly wide is the variety of applications, ranging from intrusion detection for cyber-security and fraud detection for credit cards to textual anomaly detection, industrial damage, or medical anomaly detection and image processing (Chandola et al., 2009).

This work focuses on video-based anomaly detection (Popoola and Wang, 2012; Saligrama et al., 2010). In this field, predominantly video surveillance applications are found in the literature, such as surveillance scenarios for crowded scenes of people or traffic scenes (Mahadevan et al., 2010; Saligrama and Chen, 2012; Kratz and Nishino, 2009; Piciarelli et al., 2008; Hu et al., 2006).

In these scenarios usually a fixed scene and camera setting is assumed and absolute position information is the predominant feature. Commonly, a fixed spatial (and temporal) grid representation is used and local statistics of grid cells are learned. Anomalies are detected for cells in which the observed statistics deviate from the learned model. Mahadevan et al. (2010) model the normal crowd behavior by mixtures of dynamic textures (MDTs) that jointly model appearance and dynamics of the scene. Various methods use optical flow to represent motion dynamics (Saligrama and Chen, 2012; Adam et al., 2008; Kim and Grauman, 2009). In Saligrama and Chen (2012) local statistical aggregates are built upon local feature descriptors based on optical flow computed at each spatio-temporal location. Adam et al. (2008) maintain probabilities of optical flow in local regions, using histograms. Kim and Grauman (2009) extract local optical flow and enforce consistency across locations through Markov Random Field (MRF) models. Benzeth et al. (2009) uses background subtraction to extract motion labels, these local features are then modeled using a 3D MRF. Kratz and Nishino (2009) extract spatio-temporal gradients to fit a Gaussian model and use a Hidden Markov Model (HMM) to detect abnormal events.

Trajectory-based methods constitute another category of approaches, that usually uses object-trajectories of tracked persons, or cars in traffic scenes for example (Piciarelli et al., 2008; Sillito and Fisher, 2008; Hu et al., 2006; Li et al., 2013; Nait-Charif and McKenna, 2004; Dee and Hogg, 2004), but also temporally shorter and dense trajectory representations, such as representative trajectories for crowd flow obtained from clustering particle trajectories in Wu et al. (2010). In Piciarelli et al. (2008) single-class support vector machine (SVM) clustering is used to identify anomalous trajectories in traffic scenes. Sillito and Fisher (2008) presents a method that uses cubic spline curves to parametrise trajectories and an incremental one-class learning approach using Gaussian mixture models. Hu et al. (2006) uses a hierarchical clustering of trajectories depending on spatial and temporal information and a chain of Gaussian distributions to represent motion patterns. Li et al. (2013) introduces trajectory sparse reconstruction analysis (SRA) that constructs a normal dictionary set which is used to reconstruct test trajectories. A rather different approach is presented in Dee and Hogg (2004), that uses inexplicability scores to measure the extent to which a trajectory can be regarded as goal-directed.

Trajectory-based methods share the drawback that they have to rely on a robust tracking result, which becomes more challenging when long trajectories need to be extracted or difficult scenarios appear such as crowded scenes. However, once a trajectory representation is available, a clear advantage over local feature representations is the strong semantic information contained in long trajectories representing long term motion behavior of objects or particles.

The majority of methods discussed so far, especially methods in the field of surveillance, rely on fixed scene representations, and for instance learn models on fixed spatio-temporal grids. However, in several scenarios absolute position or the exact location of events is not relevant for the characterization of anomalies. Consider cases where anomalies occur invariant to their absolute position, or rather relative to a certain spatiotemporal context. Moreover, spatio-temporal transformations, such as temporal and spatial shifts, spatial rotations or even deformations may act on motion patterns that are however not relevant for characterizing anomalies.

Several approaches in the literature exist that are able to deal with the scenario described above to some extent. While the approach described in [Dee and Hogg \(2004\)](#) presents a rather abstract way of measuring anomalies that only considers how goal-directed a movement is, another type of approaches relies on bag of invariant features. For instance, in [Wu et al. \(2010\)](#) chaotic invariants are computed that are invariant to position and magnitude, and a global probabilistic scene model of normality is learned using Gaussian mixture models (GMMs). In [Mehran et al. \(2009\)](#) a social force model induces interaction forces that serve as input for a bag of words approach. With these approaches localization of anomalies is only possible indirectly by projecting anomalous features back to their spatiotemporal domain.

Apart from basic bag of feature (BOF) approaches that usually do not consider relationships and contextual information, which however is crucial for anomaly detection, another group of methods exist. These methods are termed dictionary-based, or more general reconstruction-based or analysis-by-synthesis approaches ([Saligrama et al., 2010](#)). Reconstruction-based approaches consider anomaly detection as a reconstruction task, in which anomalies remain as poorly reconstructed entities. Antic and Ommer use the term “video parsing” for jointly explaining the foreground from normal training samples ([Antic and Ommer, 2011](#)). In [Cong et al. \(2011\)](#) a sparse reconstruction cost is used to perform sparse dictionary selection given an over-complete spatiotemporal basis. In [Cong et al. \(2013\)](#) the problem is formulated as matching against spatiotemporal segments in the training data. [Boiman and Irani \(2007b,a\)](#) and similarly [Roshtkhari and Levine \(2013\)](#) consider the problem as spatio-temporal composition from a database of patch ensembles or a codebook. Reconstruction based approaches provide a more direct explanation of the test data by concrete instances from the training data.

The presented approach is most related to reconstruction-based and analysis-by-synthesis approaches. However, several aspects distinguish the proposed method from previous work. First of all, we want to detect anomalies in context of a specific motion pattern, termed motion pattern prototype. Moreover, our method involves detection of all instances of this prototype pattern robust to globally rigid and locally elastic spatio-temporal transformations. This allows anomaly detection specific to a certain context robust to strong variations of the normal motion pattern, both globally and locally. Detection and elastic registration of a single prototype pattern can already explain large parts of observed variations and contributes to a strong generalization capability. Another component is a statistical approach that models allowed deformations and deviations in context of the prototype model. This increases the capability of the method to distinguish normal and abnormal motion on a finer level. Furthermore, a trajectory-based representation is used together with detection and elastic registration that induces spatial elasticity from trajectory affinities, which to the best of our knowledge has not been applied to anomaly detection before. The trajectory-based representation provides an expressive as well as efficient representation. Moreover, the presented approach deals with 3D+time data, while existing methods deal with 2D+time data.

2.3 Supertrajectory Representation

We suggest to represent motion patterns by “supertrajectories” describing the motion of local groups of similarly moving points, see figure 2.3b. We initialize our representation by dense point trajectories ([Sundaram et al., 2010](#)). These basic motion trajectories from tracked points constitute the lowest level of our representation and are denoted as raw trajectories, see figure 2.3a. Trajectories are allowed to start and end at arbitrary points in time. We define motion patterns by two functions,

$$\begin{aligned} \mathbf{x}_{\text{raw}} &: \Omega_{\text{raw}} \rightarrow \mathbb{R}^3, \\ w_{\text{raw}} &: \Omega_{\text{raw}} \rightarrow \{0, 1\}, \end{aligned}$$

where $\Omega_{\text{raw}} \subset \mathbb{N} \times \mathbb{R}$ denotes the domain of raw trajectories $i_{\text{raw}} \in \{1, \dots, N_{\text{raw}}\} \subset \mathbb{N}$ and time $t \in \mathbb{R}$. The position is denoted by $\mathbf{x}_{\text{raw}}(i_{\text{raw}}, t)$ and the validity by $w_{\text{raw}}(i_{\text{raw}}, t)$. $w_{\text{raw}}(i_{\text{raw}}, t) = 1$

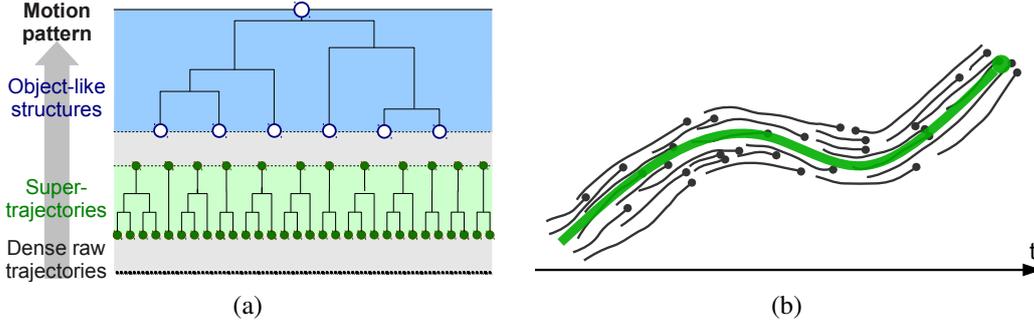


Figure 2.3: Supertrajectories are inferred from a hierarchical representation of motion patterns. **(a)** Hierarchical representation of motion patterns by agglomerative clustering of motion trajectories. **(b)** A supertrajectory (*green*) provides a smooth, robust and efficient representation of a bundle of raw trajectories (*black*).

means that trajectory i_{raw} exists at time point t and the corresponding position is valid.

2.3.1 Hierarchical Clustering

Motion patterns often exhibit a natural hierarchical composition of subpatterns. We propose a clustering step to represent bundles of similarly moving points at different hierarchical levels by “supertrajectories”, see figure 2.3. We build a hierarchical representation by agglomerative clustering (Everitt et al., 2011), that iteratively groups trajectories in bottom-up manner. Low levels represent local trajectory bundles, while higher levels represent object-like structures, as shown in figure 2.3a. Hierarchical clustering is defined by a distance metric between elements and a linkage criterion that defines distances between sets of elements. We define pairwise distances between trajectories (raw trajectories or supertrajectories) by their maximum Euclidean distance in the overlapping time window:

$$d(i, j) = \max_{t \in \mathbb{R}} (w(i, t) \cdot w(j, t) \cdot \|\mathbf{x}(i, t) - \mathbf{x}(j, t)\|). \quad (2.1)$$

The distance between temporally non-overlapping trajectories, *i.e.* $w(i, t) \cdot w(j, t) = 0, \forall t \in \mathbb{R}$, is defined as $d(i, j) = +\infty$. For bottom-up grouping we apply centroid linkage (Everitt et al., 2011), that defines distances between sets of elements by their centroids. Here, the centroid of a set of raw trajectories is computed as the supertrajectory (described in the next section). Pairs of temporally non-overlapping trajectories are not evaluated during linkage. The resulting hierarchical cluster tree can be cut at arbitrary levels to obtain 1 to N_{raw} clusters. By splitting the hierarchy at a certain level, we obtain “supertrajectories” (the clusters at the split level).

In practice, the choice of the level depends on the task and the required degree of detailedness. In our experiments on juggling patterns in section 2.7.2, we chose the level such that the motion of a juggling ball is represented by a single supertrajectory. This is achieved by setting the split level at distance $d = 8$ cm, which is the diameter of a juggling ball. Experimental section 2.7.2 gives more details and figure 2.11 illustrates supertrajectories at different levels including the level used in our experiments.

2.3.2 Supertrajectories

As illustrated in figure 2.3b, supertrajectories provide a smooth, robust and efficient representation of the dominant motion of a bundle of raw trajectories. We denote the set of raw trajectories that form one supertrajectory by $\mathcal{X}_i \subset \{1, \dots, N_{\text{raw}}\}$, where $i \in \{1, \dots, N_{\text{super}}\}$. A supertrajectory is computed

by averaging the positions of all grouped raw trajectories at each time point

$$\mathbf{x}(i, t) = \begin{cases} \frac{\sum_{i_{\text{raw}} \in \mathcal{X}_i} w_{\text{raw}}(i_{\text{raw}}, t) \cdot \mathbf{x}_{\text{raw}}(i_{\text{raw}}, t)}{\sum_{i_{\text{raw}} \in \mathcal{X}_i} w_{\text{raw}}(i_{\text{raw}}, t)} & \text{if } w(i, t) = 1 \\ \mathbf{0} & \text{else,} \end{cases}$$

where $w(i, t) = \max_{i_{\text{raw}} \in \mathcal{X}_i} w_{\text{raw}}(i_{\text{raw}}, t)$. Analogous to raw trajectories, the position is denoted by a function \mathbf{x} and the validity is denoted by a function w , on the domain $\Omega_{\text{super}} \subset \mathbb{N} \times \mathbb{R}$. Because the underlying raw trajectories may start and end asynchronously and frequently, high frequencies can be introduced along the supertrajectory. To suppress high frequencies and reduce noise the supertrajectory \mathbf{x} is finally low-pass filtered in temporal direction (average filter).

2.4 Detection of Motion Patterns

We aim for reconstructing an entire test pattern by prototype placements. Based on that, we detect anomalies in the context of the prototype. Prototype detection shall be invariant to temporal shift and 3D spatial translation and rotation, as well as spatiotemporal deformations. In this section, we describe how a motion pattern prototype, represented by supertrajectories, can be efficiently detected to yield initial transformation parameters and robust correspondences. The following section deals with the elastic registration to the underlying test pattern. The prototype pattern is denoted by the supertrajectories \mathbf{x}^a . The test sequence is represented by the supertrajectories \mathbf{x}^b .

2.4.1 Detection Hypotheses by Efficient Hashing

To efficiently detect a prototype pattern in a new test sequence we modified the hashing approach by Winkelbach et al. (2006) to deal with our spatiotemporal setting. Random point pairs of \mathbf{x}^a and \mathbf{x}^b are chosen and stored in a hash table using rotationally invariant features as table indices. Hash collisions provide transformation hypotheses with a time complexity of $O(n)$ for the first hypothesis, that converges to $O(1)$ for further hypotheses.

The original algorithm by Winkelbach et al. (2006) works on *oriented points*, a combination of surface points and their surface normals. To adapt it to our setting, we use points on supertrajectories and their velocity vectors instead: $(\mathbf{x}(i, t), \dot{\mathbf{x}}(i, t))$ for trajectory i at time t , where $\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt}$. An *oriented point pair* in our setting is termed *constellation* and consists of two oriented points and two time points:

$$\mathbf{q} = (\mathbf{x}(i, t_1); \dot{\mathbf{x}}(i, t_1); \mathbf{x}(j, t_2); \dot{\mathbf{x}}(j, t_2); t_1; t_2). \quad (2.2)$$

Figure 2.4a depicts a two point constellation on a pair of supertrajectories from the prototype pattern, that potentially matches a constellation on a pair of supertrajectories in the test pattern.

To find similar constellations in the test pattern, we compute rotation and translation invariant features of constellations. Using

$$\begin{aligned} \mathbf{d} &= \mathbf{x}(j, t_2) - \mathbf{x}(i, t_1), & \mathbf{e}_2 &= \frac{\mathbf{d}}{\|\mathbf{d}\|}, \\ \mathbf{n}_1 &= \frac{\dot{\mathbf{x}}(i, t_1)}{\|\dot{\mathbf{x}}(i, t_1)\|}, & \mathbf{n}_2 &= \frac{\dot{\mathbf{x}}(j, t_2)}{\|\dot{\mathbf{x}}(j, t_2)\|} \end{aligned} \quad (2.3)$$

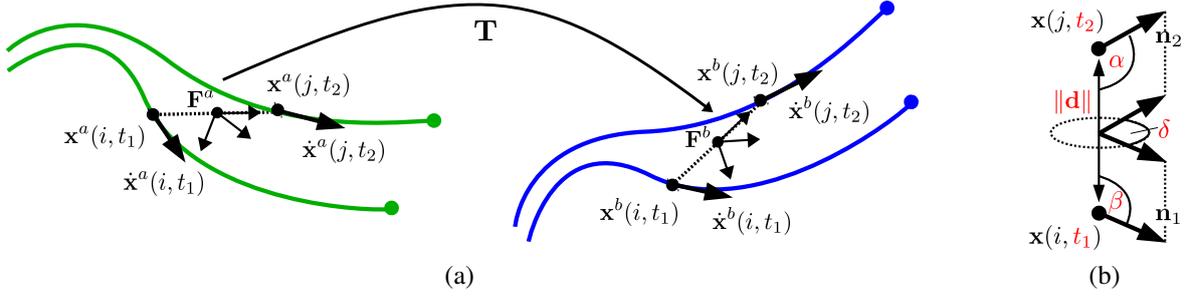


Figure 2.4: Transformation hypotheses for the detection of motion patterns using a hashing approach by Winkelbach et al. (2006) transferred to motion patterns. **(a)** Relative transformation T between oriented point pairs (constellations) on pairs of supertrajectories in the prototype pattern (green) and test pattern (blue). Oriented points, i.e. points on supertrajectories ($x(i, t)$) and their velocity vectors ($\dot{x}(i, t)$), the connecting line (dashed) and the local coordinate system in the center are indicated. **(b)** Relations between oriented points $x(i, t_1)$ and $x(j, t_2)$, and invariant features (highlighted in red).

we construct the rotation and translation invariant 6-component feature vector

$$\mathbf{f} = \begin{pmatrix} \|\mathbf{d}\| \\ \cos \alpha \\ \cos \beta \\ \gamma \\ t_1 \\ t_2 - t_1 \end{pmatrix} = \begin{pmatrix} \|\mathbf{d}\| \\ \langle \mathbf{e}_2, \mathbf{n}_1 \rangle \\ \langle \mathbf{e}_2, \mathbf{n}_2 \rangle \\ \text{atan2}(\mathbf{n}_1 \cdot (\mathbf{e}_2 \times \mathbf{n}_1), (\mathbf{n}_1 \times \mathbf{e}_2) \cdot (\mathbf{e}_2 \times \mathbf{n}_2)) \\ t_1 \\ t_2 - t_1 \end{pmatrix}. \quad (2.4)$$

The features including the involved angles (α, β, γ) between velocity directions $\mathbf{n}_1, \mathbf{n}_2$ and connecting vector \mathbf{d} are illustrated in figure 2.4b.

Additionally, we use the constellation to define a local coordinate system with the unit vectors and origin

$$\begin{aligned} \mathbf{e}_1 &= \frac{\mathbf{e}_2 \times (\mathbf{n}_1 + \mathbf{n}_2)}{\|\mathbf{e}_2 \times (\mathbf{n}_1 + \mathbf{n}_2)\|}, & \mathbf{e}_3 &= \mathbf{e}_1 \times \mathbf{e}_2, \\ \mathbf{t} &= (\mathbf{x}(j, t_2) + \mathbf{x}(i, t_1)) / 2, \end{aligned} \quad (2.5)$$

and the corresponding Frame $\mathbf{F} \in \mathbb{R}^{4 \times 4}$

$$\mathbf{F} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (2.6)$$

For the construction of a transformation hypothesis $T_{\text{hyp}} \in \mathbb{R}^{4 \times 4}$ from a constellation \mathbf{q}^a in the prototype pattern and a similar constellation \mathbf{q}^b in the test pattern the two extracted frames $\mathbf{F}^a, \mathbf{F}^b$ only need to be multiplied, i.e.

$$T_{\text{hyp}} = \mathbf{F}^b \cdot (\mathbf{F}^a)^{-1} = \begin{pmatrix} \mathbf{R} & \mathbf{b} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2.7)$$

where \mathbf{R} denotes the 3×3 rotation matrix and \mathbf{b} the translation vector that parametrize the spatial rigid transformation that maps the constellations.

The rest of the algorithm including hypotheses generation and verification was adapted accordingly. The output of the algorithm is a number of rigid transformation hypotheses. We parameterize them as

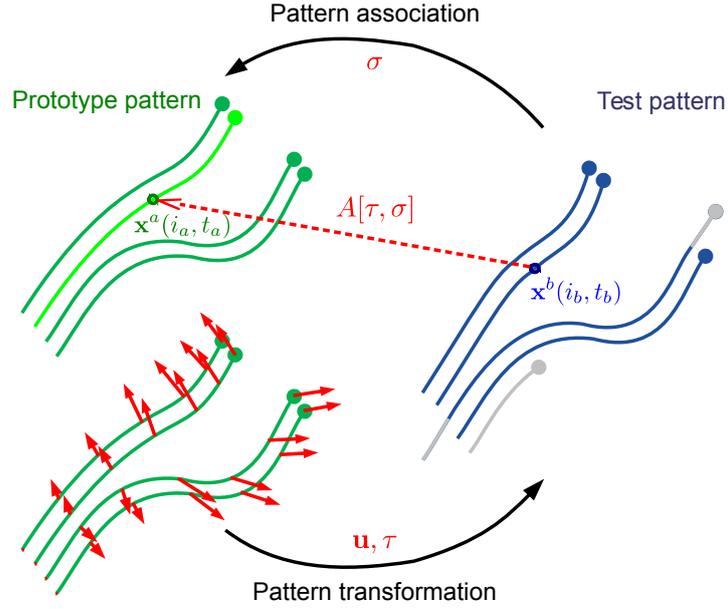


Figure 2.5: Pattern association and transformation during elastic registration. The test pattern (*blue*) is associated to the prototype pattern (*green*), which is transformed to fit the test pattern. Exemplary an assignment (dashed red line) of test pattern point $\mathbf{x}^b(i_b, t_b)$ to a prototype trajectory (*light green*) and prototype point is depicted. Unassociated test pattern (*gray*). A spatio-temporally smooth deformation is illustrated (*red arrows*), where spatial elasticity is induced from trajectory affinities, *i.e.* similar neighboring trajectories are strongly coupled.

a temporal shift t_{shift} , and a spatial rigid transformation $\mathbf{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ with $\mathbf{T}(\mathbf{x}(i, t)) = \mathbf{R}\mathbf{x}(i, t) + \mathbf{b}$, such that

$$\mathbf{x}'(i, t) = \mathbf{T}(\mathbf{x}(i, t - t_{\text{shift}})), \quad \forall i \in \{1, \dots, N_{\text{super}}\} \quad (2.8)$$

and $w'(i, t) = w(i, t - t_{\text{shift}})$ respectively.

2.5 Elastic Registration of Motion Patterns

2.5.1 Trajectory Association Function

To perform registration, both the correspondences and the transformation between prototype and test pattern have to be estimated. We define trajectory correspondences by association function

$$\sigma(i_b, t_b) = i_a,$$

that assigns test trajectory i_b of pattern \mathbf{x}^b at time t_b to a prototype trajectory i_a of pattern \mathbf{x}^a . The assignment is undefined $\sigma(i_b, t_b) := 0$, if no temporally warped prototype trajectory i_a is valid at time t_b . Trajectory association is allowed to change in time, which is important, for instance, when a temporally long trajectory corresponds to multiple temporally consecutive trajectories. Figure 2.5 illustrates the association of a point on a test trajectory to a prototype trajectory.

2.5.2 Rigid Pre-Alignment

The transformation of each detection hypothesis is refined using a scheme similar to the iterative closest point algorithm (Besl and McKay, 1992). The energy to be minimized is the sum of squared distances

of all points of test pattern \mathbf{x}^b to the associated points in pattern \mathbf{x}^a :

$$E_{\text{data}}(\mathbf{T}, t_{\text{shift}}, \sigma) = \sum_{\substack{(i_b, t_b) \in \Omega_b \\ w(i_b, t_b) = 1 \\ \sigma(i_b, t_b) \neq 0}} \Psi \left(\|\mathbf{T}(\mathbf{x}^a(A[t_{\text{shift}}, \sigma](i_b, t_b))) - \mathbf{x}^b(i_b, t_b)\|^2 \right) + \sum_{\substack{(i_b, t_b) \in \Omega_b \\ w(i_b, t_b) = 1 \\ \sigma(i_b, t_b) = 0}} d_{\text{undef}}^2. \quad (2.9)$$

For easier notation, we define association function

$$A[t_{\text{shift}}, \sigma](i_b, t_b) = (i_a, t_a), \quad (2.10)$$

that assigns test trajectory i_b at time t_b to prototype trajectory i_a at time t_a . The function depends on the temporal shift t_{shift} and the trajectory association function σ . To be robust against outliers, we apply a truncated squared norm denoted by function $\Psi(d^2) = d^2$ for $d \leq d_{\text{max}}$ and $\Psi(d^2) = d_{\text{max}}^2$, otherwise. The maximum distance parameter $d_{\text{max}} > 0$ is defined during training and reduces the influence of outliers beyond that distance. In addition, unassociated points, *i.e.* $\sigma(i_b, t_b) = 0$, are penalized with $d_{\text{undef}} \geq 0$. The energy is minimized with respect to the rigid transformation parameters \mathbf{T} and t_{shift} , and association function σ .

We note that for simplicity we use sum notations in equations 2.9, 2.11 and equations 2.5.3-2.5.3 for the time domain as well, even it is continuous and would require integral notation. In practice however, time domain is discretized as well, which is detailed below in section 2.5.4.

We give an overview of pattern association and transformation in figure 2.5 to accompany the following section.

2.5.3 Elastic Registration

After rigid pre-alignment we perform a spatiotemporal elastic registration. The elastic transformation is parametrized by a spatial deformation function \mathbf{u} and a temporal warping τ , where

$$\mathbf{u}(i, t) : \Omega_a \rightarrow \mathbb{R}^3 \quad \text{and} \quad \tau(i, t) : \Omega_a \rightarrow \mathbb{R}.$$

The spatiotemporally transformed motion pattern reads

$$\begin{aligned} \mathbf{x}'(i, t) &= \mathbf{x}(i, t - \tau(i, t)) + \mathbf{u}(i, t - \tau(i, t)) \quad \text{and} \\ w'(i, t) &= w(i, t - \tau(i, t)) \quad \text{accordingly.} \end{aligned}$$

Data term The data term is defined analogous to equation 2.9, where the rigid transformations $(\mathbf{T}, t_{\text{shift}})$ get replaced by deformation and temporal warping (\mathbf{u}, τ) . The data term is the sum of squared distances of all points of test pattern \mathbf{x}^b to the associated points in pattern \mathbf{x}^a , subject to spatial deformation and temporal warping

$$E_{\text{data}}(\mathbf{u}, \tau, \sigma) = \sum_{\substack{(i_b, t_b) \in \Omega_b \\ w(i_b, t_b) = 1 \\ \sigma(i_b, t_b) \neq 0}} \Psi \left(\|\mathbf{x}^a(A[\tau, \sigma](i_b, t_b)) + \mathbf{u}(A[\tau, \sigma](i_b, t_b)) - \mathbf{x}^b(i_b, t_b)\|^2 \right) + \sum_{\substack{(i_b, t_b) \in \Omega_b \\ w(i_b, t_b) = 1 \\ \sigma(i_b, t_b) = 0}} d_{\text{undef}}^2. \quad (2.11)$$

Analogous to equation 2.10, we define an association function $A[\tau, \sigma](i_b, t_b) = (i_a, t_a)$, which instead depends on the temporal warping τ .

Smoothness constraints Furthermore, we formulate the following smoothness assumptions: Firstly, the elastic transformation should be smooth both spatially (across neighboring trajectories) and temporally (along trajectories). Secondly, the assignment function should be temporally smooth as well. To this end, we formulate the total energy

$$E(\mathbf{u}, \tau, \sigma) = E_{\text{data}}(\mathbf{u}, \tau, \sigma) + \alpha_{\text{spatial}} E_{\text{spatial}}(\mathbf{u}, \tau) + \alpha_{\text{temp}} E_{\text{temp}}(\mathbf{u}, \tau) + \alpha_{\text{assign}} E_{\text{assign}}(\sigma), \quad (2.12)$$

with weighting parameters $\alpha_{\text{spatial}} \geq 0$, $\alpha_{\text{temp}} \geq 0$ and $\alpha_{\text{assign}} \geq 0$.

Spatial smoothness. The elastic coupling within the prototype pattern is described by a smoothness energy on the spatiotemporal deformation functions for each pair of prototype trajectories

$$E_{\text{spatial}}(\mathbf{u}, \tau) = \sum_{\substack{i,j,t \\ (i,t) \in \Omega_a \wedge (j,t) \in \Omega_a \\ w(i,t)=1 \wedge w(j,t)=1}} C(i, j) \cdot \left(\|\mathbf{u}(i, t) - \mathbf{u}(j, t)\|^2 + \beta_{\text{temp}} (\tau(i, t) - \tau(j, t))^2 \right).$$

Spatial smoothness across trajectories is defined, such that the elastic coupling between trajectories

$$C(i, j) = \exp(-d(i, j)^2 / 2\sigma_{\text{spatial}}^2) \quad (2.13)$$

depends on the pairwise distances $d(i, j)$ (equation 2.1), where σ_{spatial} is a spatial scaling parameter (we used $\sigma_{\text{spatial}} = 10$ cm in our experiments in section 2.7.2). The weighting between spatial deformation and temporal warping is determined by $\beta_{\text{temp}} \geq 0$. This results in strong coupling between similar trajectories (similar moving points) and weak coupling between dissimilar trajectories, that therefore can be transformed rather independently. Figure 2.5 illustrates such a spatio-temporally smooth deformation on the domain of prototype trajectories.

Temporal smoothness. A temporal smooth transformation along trajectories is enforced by the energy term

$$E_{\text{temp}}(\mathbf{u}, \tau) = \sum_{\substack{(i,t) \in \Omega_a \\ w(i,t)=1}} \|\dot{\mathbf{u}}(i, t)\|^2 + \beta_{\text{temp}} (\dot{\tau}(i, t))^2, \quad (2.14)$$

where $\dot{\mathbf{u}} := \frac{d\mathbf{u}}{dt}$ and $\dot{\tau} := \frac{d\tau}{dt}$ respectively.

Assignment smoothness. Moreover, a temporally smooth assignment is preferred by the smoothness term

$$E_{\text{assign}}(\sigma) = \sum_{\substack{i,t \\ (i,t) \in \Omega_b \wedge (i,t-\Delta t) \in \Omega_b \\ w(i,t)=1 \wedge w(i,t-\Delta t)=1 \\ \sigma(i,t) \neq 0 \wedge \sigma(i,t-\Delta t) \neq 0}} d^2(\sigma(i, t), \sigma(i, t - \Delta t)),$$

that penalizes temporal assignment changes from one to another trajectory by their pairwise distance $d(i, j)$. In this way, assignment changes between dissimilar trajectories get strong penalization. Let Δt_{\min} be the shortest interval for assignment changes of σ , then $\Delta t \in (0, \Delta t_{\min}] \subset \mathbb{R}$ must be chosen.

2.5.4 Temporal discretization

In practice motion patterns, specifically trajectories, are not given time continuous. Instead, trajectories are given discretized by samples along time dimension. In this respect, sum notation for time domain becomes exact (see equations 2.9, 2.11 and equations 2.5.3-2.5.3). To compute intermediate positions on sampled trajectories we use linear interpolation between the two adjacent sampled positions.

2.5.5 Energy Optimization

We found an approximate solution for minimizing the total energy in equation 2.12, for both the rigid pre-alignment and the elastic registration. We apply alternating optimization of the transformation and the assignment σ .

Given a fixed elastic transformation for example, the assignment function can be obtained by minimizing the energy

$$E'(\mathbf{u}, \tau, \sigma) = E_{\text{data}}(\mathbf{u}, \tau, \sigma) + \alpha_{\text{assign}} E_{\text{assign}}(\sigma) \quad (2.15)$$

with respect to σ . This can be done by exact inference and computed efficiently by dynamic programming. The rigid transformation \mathbf{T} and the temporal shift t_{shift} are found by a Procrustes algorithm (Umeyama, 1991) using all point correspondences with distance $d < d_{\text{max}}$.

The elastic transformation is obtained by minimizing the energy

$$E''(\mathbf{u}, \tau, \sigma) = E_{\text{data}}(\mathbf{u}, \tau, \sigma) + \alpha_{\text{spatial}} E_{\text{spatial}}(\mathbf{u}, \tau) + \alpha_{\text{temp}} E_{\text{temp}}(\mathbf{u}, \tau) \quad (2.16)$$

with respect to (\mathbf{u}, τ) . We solve this optimization problem by L-BFGS (Byrd et al., 1995), which is a popular optimization algorithm particularly well suited for continuous optimization problems with a large number of variables when first- or second-order information is available.

Alternating optimization is repeated until the estimated transformation converges. Both parts of the optimization can be solved globally optimal: 1) Optimization of the transformation only involves quadratic energy terms and thus yields a convex energy that can be solved globally optimal. 2) Optimization of the assignment function can even be done by exact inference, since the problem only involves linear chains of depending variables along separate trajectories. No pairwise trajectory terms are involved here, and thus no cyclic dependencies are present.

2.6 Motion Anomaly Detection

2.6.1 Learning a Spatiotemporal Deformable Prototype Model

First a concrete motion pattern is selected as the spatiotemporal prototype which represents a clean and segmented instance of the “normal” motion pattern of interest, see figure 2.1a and movie 2.17. The “normal” variations of spatiotemporal deformation and remaining deviations observed in training sequences are learned and together with the prototype build the prototype model. The complete training pipeline is shown in figure 2.2 (top row).

In our experiments, we manually selected a specific spatio-temporal section of the training data and used it as the prototype pattern. In case of juggling patterns the selected prototype and its spatio-temporal extents are depicted in figure 2.11. However, only a single prototype pattern needs to be selected manually. Learning the prototype model from observed variations in a set of entire training sequences, as depicted in figure 2.15, is automatic.

For learning the prototype model, the prototype pattern is detected in training sequences and D detections are selected. For each detection d , the prototype pattern is rigidly pre-aligned and elastically registered. This results in rigid transformation parameters $((\mathbf{R}_d, \mathbf{b}_d), t_{\text{shift},d})$ and elastic transformation parameters (\mathbf{u}_d, τ_d) . Elastic spatial transformation parameters are transformed to the prototype coordinate system by $\mathbf{u}'_d = \mathbf{R}_d^{-1} \mathbf{u}_d$. We build a statistical model that captures:

Global spatiotemporal deformations and data fitting costs after registration For both, we define bounds for validating prototype registrations. Global deformation parameters are computed by standard PCA on the concatenation of deformation parameters $\mathbf{u}'_d(i, t)$ and $\tau_d(i, t)$, where $(i, t) \in \Omega_a \wedge w(i, t) = 1, \forall d \in \{1, \dots, D\}$. Let \mathbf{y}_d be the representation of deformations for detection d in PCA space, with elements $y_d(k)$ and $k \in \{1, \dots, K\}$. We define a bounding hyper-cuboid in PCA

space based on the lower and upper bounds $y_{\min}(k) = \min_d(y_d(k))$ and $y_{\max}(k) = \max_d(y_d(k))$, for each k . Prototype data fitting costs c_d are computed analogous to the data term in equation 2.11, but averaging over valid points on the prototype domain Ω_a and applying association functions mapping from Ω_a to Ω_b instead. We define an upper bound based on the maximum $c_{\max} = \max_d(c_d)$.

Residual distances remaining after elastic registration For each prototype pattern point $\mathbf{x}^a(i, t)$ the obtained residual distances with associated training pattern points are locally aggregated and learned using a Gaussian residual model.

2.6.2 Reconstructing Motion Patterns by Prototype Placements

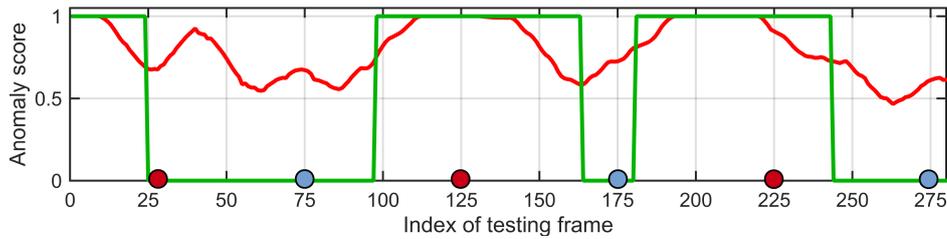
For reconstructing a whole test pattern by prototype placements we apply a greedy search algorithm. It iteratively finds best placements of prototype patterns into the test pattern. In other words, it finds rigid transformations of the prototype pattern that are refined by elastic registration to fit parts of the test pattern, with the goal of explaining or covering the test pattern as much as possible. The algorithm is detailed in the following. Candidate placements are obtained from prototype detections. A priority list defines the order of placements. It is sorted by the score $s' = r \cdot s$, where $s = d_{\max}^2 - c$, with data fitting cost c , described in section 2.6.1, and maximum value d_{\max}^2 . The ratio of the candidate pattern that overlaps with so far unreconstructed test pattern is denoted by r . Thus, a high score is achieved only with good data fitting combined with good overlapping with so far unreconstructed test pattern. In each step, the k-best candidates are elastically registered. Registered candidates are accepted only if all deformation and data fitting parameters are within the learned bounds. Rejected candidates are removed from the priority list. The scores of accepted candidates are updated (temporarily) and the best accepted candidate is selected for reconstruction and removed from the priority list. The unreconstructed test pattern is updated accordingly. The algorithm stops, if the test pattern is reconstructed completely, or if no candidates remain that can reconstruct parts of the unreconstructed test pattern.

2.6.3 Computing Anomaly Scores

Pointwise Anomaly Score We compute an anomaly score for each test pattern point expressing how much it deviates from the prototype model. The pointwise anomaly score is computed by taking the minimum residual distance to all registered prototype patterns and applying the locally learned residual model from the associated prototype pattern point.

Framewise Anomaly Score We map the anomaly scores of all trajectory points within one video frame to a framewise anomaly score by computing the maximum. The maximum measure is better suited for detecting local fine-grained anomalies, compared to the average measure, which is sufficient for detecting global anomalies. In presence of strong noise, that results in outliers, using a more robust maximum measure, such as an upper percentile is more appropriate. We use the 98%tile for very noisy biological motion patterns, see section 2.7.3.

An anomaly profile is computed that yields the temporal progression of motion anomalies. A sample profile is plotted in figure 2.6a. Overall it shows high anomaly scores in anomalous sections and decreasing and lower scores in normal sections, indicated by the ground truth profile. The video snapshots in figure 2.6b accompany the description at the marked time points on the time axis in the anomaly profile. Decreasing anomaly scores across the edge of anomalous sections are due to continuous transitions between normal and abnormal motion, and the fact that the plotted anomaly profile is smoothed (average filter with size of 15 frames). Also, a peak of rather high anomaly score is present in the first normal section roughly at frame 40. The reason for this is that in this example we use the maximum measure to compute framewise anomaly scores. This measure is maximally sensitive, meaning that even a single anomalous trajectory in the pattern yields an increased anomaly score.



(a) Anomaly profile (*red*). Ground truth profile (*green*). Markers on the x-axis indicate the position of the frames shown in (b-g). The *color* indicates whether the frame shows a normal (*blue*) or abnormal motion (*red*).

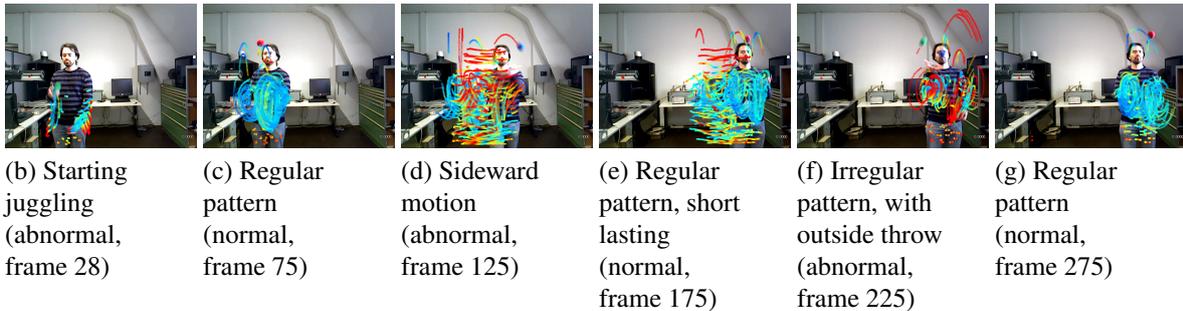


Figure 2.6: Anomaly profile for a juggling sequence containing several motion anomalies. **(a)** Anomaly profile. **(b-g)** Video frames corresponding to the frame markers in the anomaly profile are shown. Supertrajectories are plotted with a delay of 45 frames (≈ 1.5 sec). The color corresponds to the obtained anomaly scores and ranges from *blue* (normal motion) to *red* (abnormal motion).

Global Anomaly Score In case a global anomaly assessment is desired, a single score is computed from all pointwise anomaly scores. For quantifying biological motion patterns we compute global anomaly scores using the 98%tile, see section 2.7.3.

2.7 Experiments

Our anomaly detection approach is basically different from most existing approaches. Accordingly, we found existing benchmarks to be inappropriate for demonstrating our method. Popular datasets such as the UMN dataset¹, the UCSD dataset (Mahadevan et al., 2010) or the Subway dataset (Adam et al., 2008) present surveillance scenarios with fixed scene and camera, where absolute position is relevant and fixed spatiotemporal grid representations are sufficient. Apart from dealing with 2D data only, particularly the task of detecting anomalies in context of a specific motion pattern of interest is not present. To demonstrate our method we recorded a new motion anomaly dataset from persons juggling balls using a Kinect camera (see figure 2.10).

In this section, we present experiments on juggling patterns, with evaluation of anomaly detection and localization, as well as insights into supertrajectory representation and elastic registration. An application on real biological motion patterns concludes the experiments. Beforehand, basic validation experiments demonstrate the effect of the smoothness terms on elastic registration.

2.7.1 Basic Validation Experiments of Elastic Registration

We first show basic experiments on simple 2D+time patterns to demonstrate the effect of temporal, spatial and assignment smoothness terms. Figures 2.7, 2.8 and 2.9 first show the initial configuration

¹Unusual crowd activity dataset made available by the University of Minnesota at: <http://mha.cs.umn.edu/Movies/Crowd-Activity-All.avi>

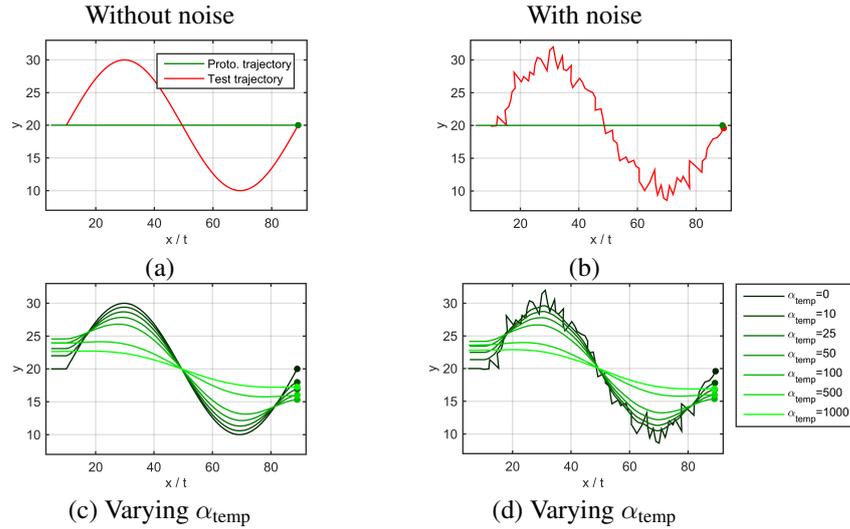


Figure 2.7: Effect of temporal smoothness on elastic registration. **(a-b)** Minimal sample pattern with one prototype (*green*) and one test trajectory (*red*), without **(a)** and with noise **(b)** applied to the test trajectory. Note that in this minimal case of single trajectory patterns both spatial and assignment smoothness terms have no influence. **(c-d)** Registration results for increasing temporal smoothness parameter α_{temp} (plots from *dark* to *light green*), for both cases (a-b).

of prototype pattern and test pattern, followed by elastic registration results for varying smoothness parameters.

Temporal smoothness As illustrated in figure 2.7 applying no temporal smoothness ($\alpha_{\text{temp}} = 0$) results in an elastic deformation that tightly fits the prototype pattern to the test pattern, and overfits in case of noise (figure 2.7c-d). Applying slight temporal smoothness instead, also yields smooth results in case of noise (see figure 2.7d). Further increasing temporal smoothness, increases the temporal stiffness of the resulting deformation and converges towards a (trajectory-wise) rigid transformation in the limit, see the progression of registration results for increasing α_{temp} in figure 2.7c-d. Note that in this minimal example of single trajectory patterns spatial and assignment smoothness terms have no influence.

Spatial and assignment smoothness Figure 2.8 illustrates the effect of spatial and assignment smoothness using four extreme cases. A prototype pattern consisting of two parallel running trajectories is registered to a single sinusoidal test trajectory. Compared to the previous minimal single trajectory example in figure 2.7, a second prototype trajectory adds spatial constraints, and both spatial and assignment smoothness terms influence the result. In case smooth assignment of test to prototype trajectories is disabled ($\alpha_{\text{assign}} = 0$) the assignment changes immediately in favor of fulfilling the data term instead, see left columns in figure 2.8b-c. In contrast, when assignment smoothness is applied (in this example $\alpha_{\text{assign}} = 1000$) the assignment stays constant for the whole time range the test trajectory exists, see right columns in figure 2.8b-c. On the other hand, adding spatial smoothness leads to an elastic coupling of neighboring trajectories in the prototype pattern. As the example in figure 2.8b shows, the two neighboring trajectory are smoothly deformed together when spatial smoothness is applied ($\alpha_{\text{spatial}} = 1$, bottom row) compared to the other extreme where spatial smoothness is disabled ($\alpha_{\text{spatial}} = 0$, top row).

Spatio-temporal smoothness In addition to the previous examples figure 2.9 demonstrates the effect of varying spatial and temporal smoothness parameters using more complex patterns of circular

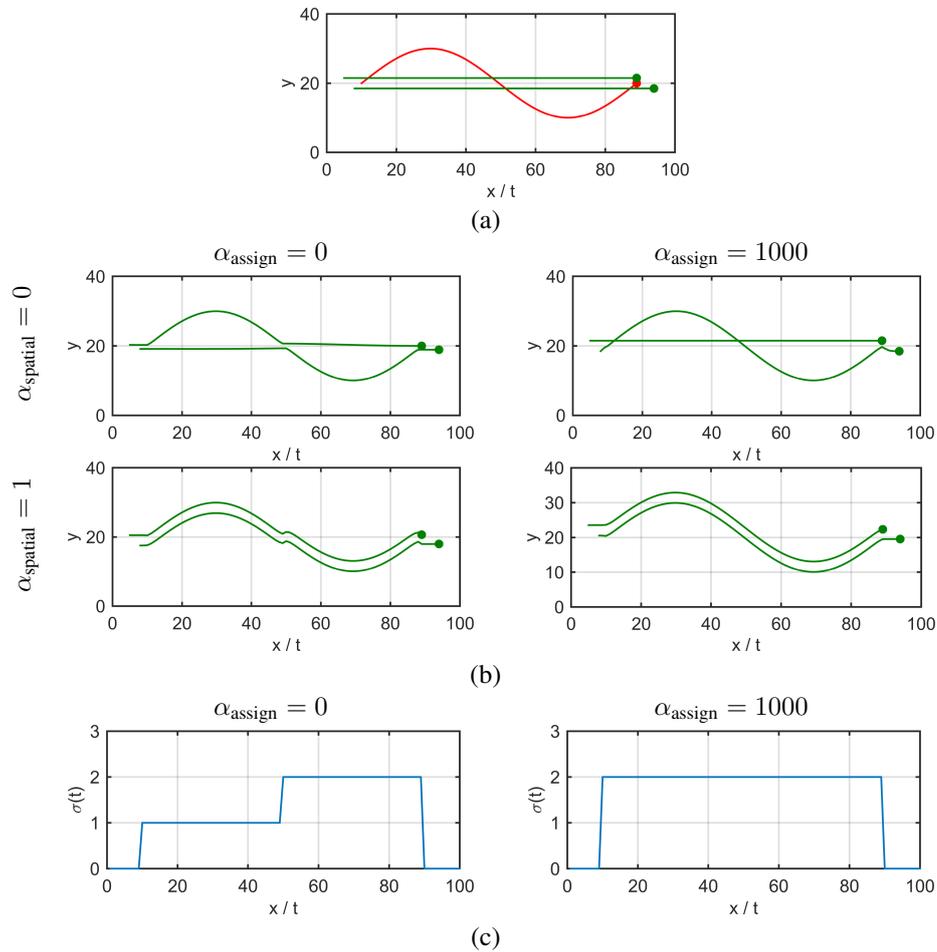


Figure 2.8: Effect of spatial and assignment smoothness on elastic registration. **(a)** Minimal sample pattern with two prototype (*green*) and one test trajectory (*red*). Note that in this minimal case of two prototype trajectories, only a single degree of spatial elasticity is present. **(b)** Registration results in four extreme cases of spatial and assignment smoothness parameter combinations. In all cases $\alpha_{\text{temp}} = 0.5$ is chosen. **(c)** Trajectory association functions σ shows the assignment of test trajectory to prototype trajectories for both cases $\alpha_{\text{assign}} = 0$ and $\alpha_{\text{assign}} = 1000$ in (b).

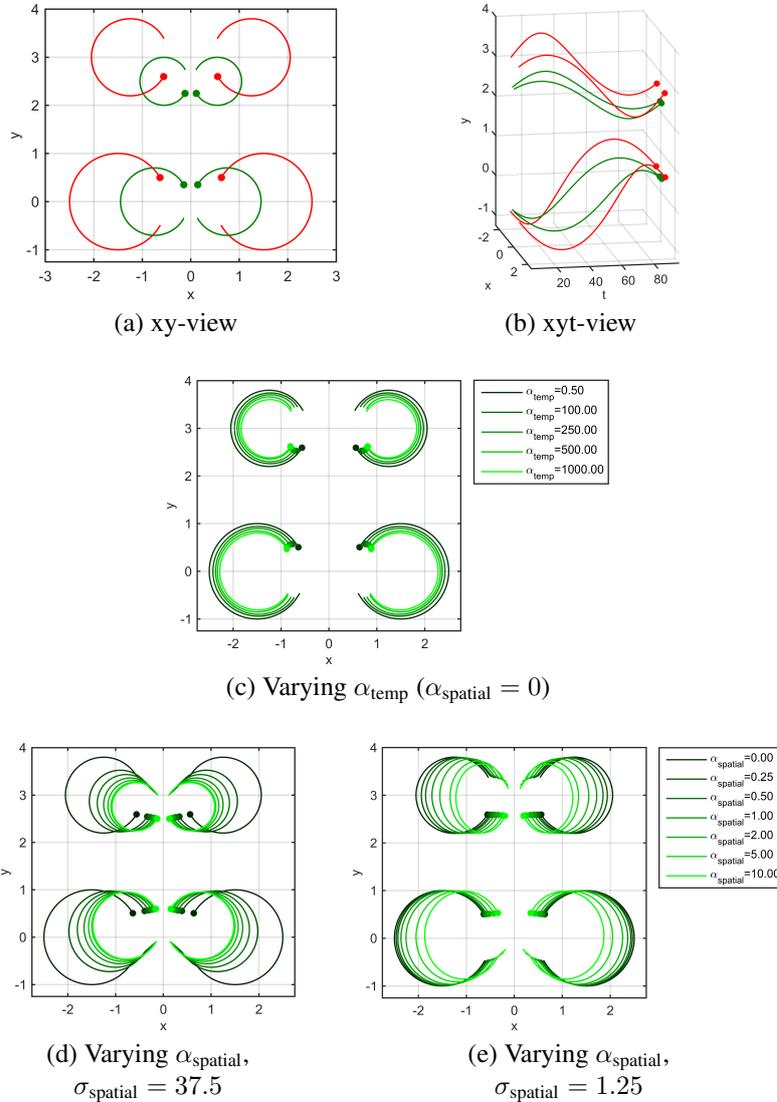


Figure 2.9: Effect of spatio-temporal smoothness on elastic registration. **(a-b)** Sample pattern with four prototype (*green*) and four test trajectories (*red*). Two pairs of contrary circular motions are present (*top* and *bottom*). **(a)** xy-view **(b)** xyt-view **(c)** Registration results for increasing temporal smoothness parameter α_{temp} (plots from *dark* to *light green*), $\alpha_{spatial} = 0$. **(d-e)** Registration results for increasing spatial smoothness parameter $\alpha_{spatial}$ (plots from *dark* to *light green*), $\alpha_{temp} = 0.5$. **(d)** Large $\sigma_{spatial}$. **(e)** Small $\sigma_{spatial}$.

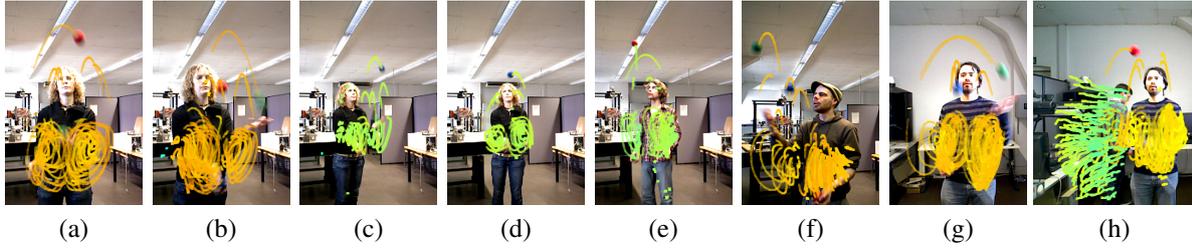


Figure 2.10: Juggling pattern test datasets of different persons performing the standard 3-ball cascade pattern including various anomalies. (a-d) One person from four different viewpoints. (e-g) Three further persons performing the same juggling pattern. (h) Background motion. Supertrajectories are shown over a range of 45 frames (≈ 1.5 sec), the *color* corresponds to the depth obtained from Kinect camera.

motions, see figure 2.9a-b. Also, the influence of the elastic coupling parameter σ_{spatial} is shown (figure 2.9d-e). Figure 2.9c demonstrates two properties: 1) Prototype trajectories are allowed to deform independently from each other towards the test pattern when spatial smoothness is disabled. 2) Increasing temporal smoothness, increases temporal stiffness towards a (trajectory-wise) rigid transformation (light green plot), while low temporal smoothness results in high temporal elasticity and tight fit to the test pattern (dark green plot). Figures 2.9d-e show that with increasing spatial smoothness the resulting deformation becomes increasingly spatially stiff. While the elastic coupling is primarily induced by pairwise trajectory distances, it is scalable by parameter σ_{spatial} (see equation 2.13). In this example, using a large value results in strong coupling between all trajectories and with increasing spatial smoothness yields results towards a globally rigid transformation, see figure 2.9d. In contrast, when using smaller values, coupling between trajectories is reduced and with increasing spatial smoothness rather yields locally rigid transformations, see figure 2.9e where only the coupling within the two pairs of circular motion remains strong.

2.7.2 Anomaly Detection in Juggling Patterns

Dataset

To demonstrate our method we recorded a new motion anomaly dataset using a Kinect camera from persons juggling balls. From the obtained RGB-D data we extract 3D+time motion patterns as follows. Motion trajectories are generated by large displacement optical flow tracking (Sundaram et al., 2010), which was adapted to use depth in addition to RGB data. The tracking algorithm generates 2D trajectories, which we transform to 3D by adding the depth coordinates. Background motion is removed using a threshold on the maximum velocity.

Multiple persons performing a standard 3-ball cascade juggling pattern were recorded from different viewpoints, see figure 2.10. For testing, we recorded similar sequences, but asked the persons to include "anomalies" randomly. Given training sequences and a selected prototype pattern, the task is to detect anomalies in test sequences, in context of the learned prototype model.

For training, three sequences (200 frames each) were used, that contain juggling patterns from three different persons, recorded in frontal view with 1.5 m distance to the camera. We chose a rather small training set with a single viewpoint to demonstrate generalization capabilities of our method. For testing, we used 29 sequences (150-480 frames each; in total more than 11.000 frames). They contain juggling patterns from five different persons including anomalies, recorded in different viewpoint and positioning settings: frontal and angular side view, close and far distance to the camera (1.5 m and 2.5 m) and different positions of the person. Challenging sequences are included that contain background motion (see figure 2.10h), persons moving while juggling and two persons juggling side-by-side (see figure 2.1b). Anomalies included are manifold and comprise anomalous single ball throws

Table 2.1: Juggling patterns test set overview

Sequence	Person	Frames	Distance to juggler	View angle
1	1	141	1.5m	0°
2	1	460	1.5m	±20°
3	1	220	2.5m	0°
4	1	460	2.5m	±20°
5	2	250	1.5m	0°
6	2	440	2.5m	0°
7	3	410	1.5m	0°
8	3	430	1.5m	±20°
9	3	420	2.5m	0°
10	3	420	2.5m	±20°
11	4	430	1.5m	0°
12	4	360	1.5m	±20°
13	4	440	2.5m	0°
14	4	440	2.5m	±20°
15	5	460	1.5m	0°
16	5	420	1.5m	±20°
17	5	440	2.5m	0°
18	5	410	2.5m	±20°
19	3+5	370	2.5m	0°
20	3+5	480	2.5m	0°
21	1	280	1.5m	0° ... ± 20°
22	1	410	1.5m...2.5m	0° ... ± 20°
23	1	350	1.5m...2.5m	0° ... ± 20°
24	1	320	1.5m	0°
25	1	250	1.5m	0°
26	1	390	1.5m	0°
27	2	360	1.5m...2.5m	0° ... ± 20°
28	2	400	1.5m	0° ... ± 20°
29	2	450	1.5m	±20°

In total: #persons: 5, #sequences: 29, #frames: 11.111

(high or outside throws, see figure 2.1b-c), completely different juggling patterns or juggling mistakes resulting in ball drop for example. Complementary to figure 2.10, see movie 2.17 to get a better impression of the recorded training and test sequences. An overview of the test set is given in table 2.1.

We generated a framewise anomaly ground truth. Additionally, we provide a segmentation of juggling relevant motion patterns (hands and arms of persons, and juggling balls) for both, the train and test set. The segmentation is given as a pointwise labelling of supertrajectories from our representation.

Evaluation

We quantitatively evaluate anomaly detection, *i.e.* classifying each video frame as normal or abnormal, see figure 2.14 and table 2.2, and show qualitative results for anomaly localization in figure 2.15 and figure 2.16. In addition, qualitative results give insights into the supertrajectory representation (see figure 2.11) and the elastic registration (see figure 2.12 and figure 2.13).

Anomaly detection ROC curves are generated by thresholding the anomaly profile (figure 2.6a) at different levels. We compare against an existing anomaly detection method, namely chaotic invariants (CI) for anomaly detection in crowded scenes (Wu et al., 2010)². Among the methods with code available, CI is the most related one as it provides invariance to position and magnitude. To the best of our knowledge, there is no previous method that deals with 3D+time data. However, to provide a comparison for CI on 3D+time data as well, we extended the approach to 3D.

The evaluation is based on a temporally short prototype pattern (25 frames), see figure 2.1a and figure 2.11c. The prototype model is learned from the training sequences, and anomaly detection is performed on all 29 test sequences. For comparison, we evaluate two versions of Wu et al. (2010): 1) chaotic invariants only (position invariant), we denote by *CI* and 2) absolute position added to *CI*, we denote by *CI+pos*. We perform anomaly detection as described in Wu et al. (2010). For comparability, we provide the same optical flow to both methods (large displacement optical flow (Brox and Malik, 2011)). Since the approach by Wu et al. (2010) does not include model detection, it performs anomaly detection on the whole frame. We use the segmentation of juggling relevant motion patterns to provide the method by Wu et al. (2010) with segmentation information. To obtain a framewise anomaly score we use the maximum measure, whereas chaotic invariants (Wu et al., 2010) use the mean measure.

Results

Supertrajectory representation Figure 2.11 gives insights into the applied supertrajectory representation. Supertrajectories are plotted for increasing split levels using the example of the juggling pattern prototype. The raw trajectory pattern in figure 2.11a shows bundles of highly redundant trajectories tracing the ball parabolas (at the top) and circular hand and arm movements (at the bottom). Note that for visualizing the 4D patterns (3D+time), we have to resort to 3D projections, such as xyz-plots (miss temporal dynamics) and xyt-plots (miss spatial z-dimension). With increasing level, specified by distance d , the hierarchical cluster tree gets split at higher levels (review illustration in figure 2.3a). In this process, an increasingly larger set of raw trajectories, with larger spatio-temporal context, contributes to each supertrajectory. One can observe a more compact representation with fewer trajectories for ball, hand and arm motions in figure 2.11b-e. Apart from spatial merging, trajectories also get merged and extended in temporal dimension, specifically consider the merged hand and ball motion in figure 2.11d-e (light blue trajectory). While raw trajectories are too detailed, redundant and inefficient for computations and high level representations, such as those shows in figure 2.11d-e, are potentially too coarse, we used the intermediate level of representation in our experiments, see figure 2.11c. We chose the level at distance $d = 8$ cm, which is the diameter of a juggling ball, and thus results in single supertrajectories for each ball motion.

Elastic registration To give insights into the elastic registration of juggling patterns we first show exemplary results in figure 2.12 for two cases, the registration to a normal and an abnormal instance of the test pattern. Second, the effect of varying smoothness parameters is shown in figure 2.13.

The results in figure 2.12 first of all show, that a robust spatio-temporal rigid initialization is achieved in case of a normal as well as an abnormal instance of the test pattern, see figure 2.12a,e. These initializations are obtained during test pattern reconstruction, described in section 2.6.2. Rigid transformations are initialized from transformation hypotheses and subsequently refined by rigid and elastic registration. Results in figure 2.12a-b and figure 2.12e-f show that after elastic registration prototype trajectories fit well to the corresponding test trajectories, such that these test pattern points yield low residual distances and receive low anomaly scores. In contrast, abnormal trajectories are not fit well, due to structural differences or elastic deformations that exceed the elasticity of the prototype model. In particular, consider the high ball throw and subsequent wide arm motion to catch the ball, which

²We thank the authors for providing essential code pieces to reimplement their method.

2 Spatiotemporal Deformable Prototypes for Motion Anomaly Detection

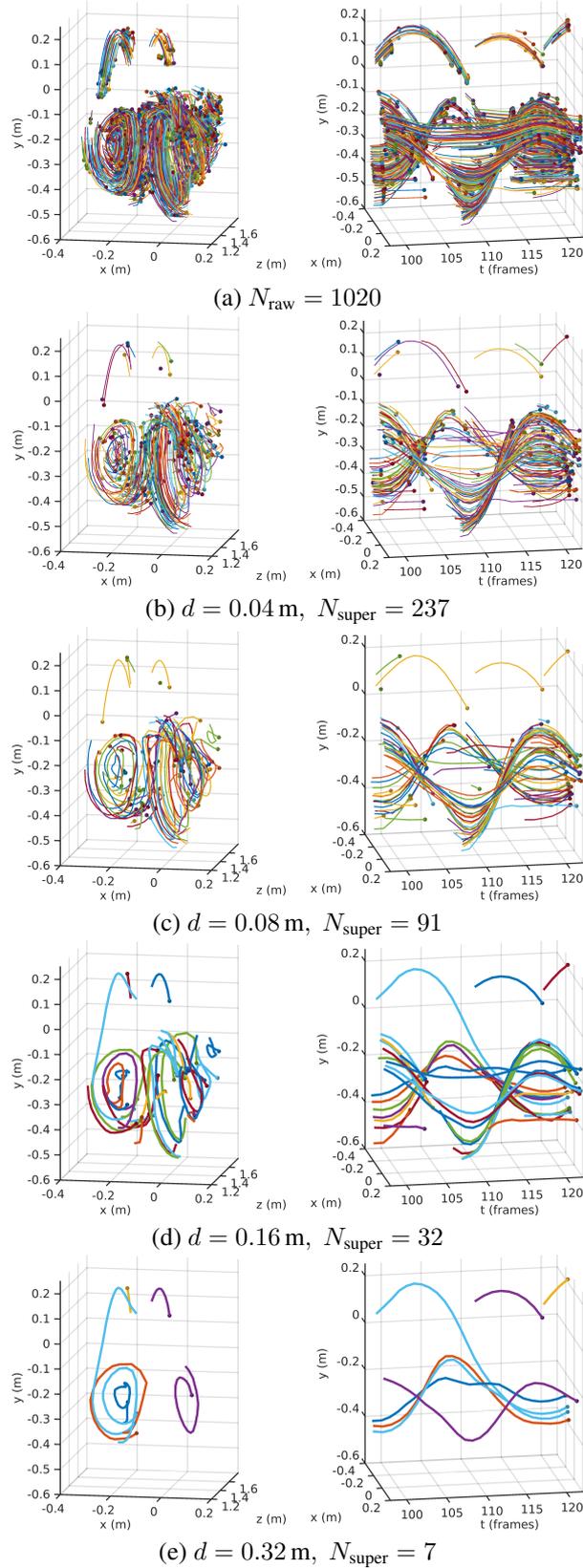


Figure 2.11: Supertrajectories obtained at increasing hierarchical levels illustrated using the juggling pattern prototype. **(a)** Raw trajectory pattern, with bundles of redundant trajectories. **(b-e)** Supertrajectories at increasing split levels (distance d) yield an increasingly compact representation with fewer supertrajectories. **(c)** Level of representation used in the experiments. (*left column*) xyz-view (*right column*) xyt-view; Trajectory end points are marked with a dot.

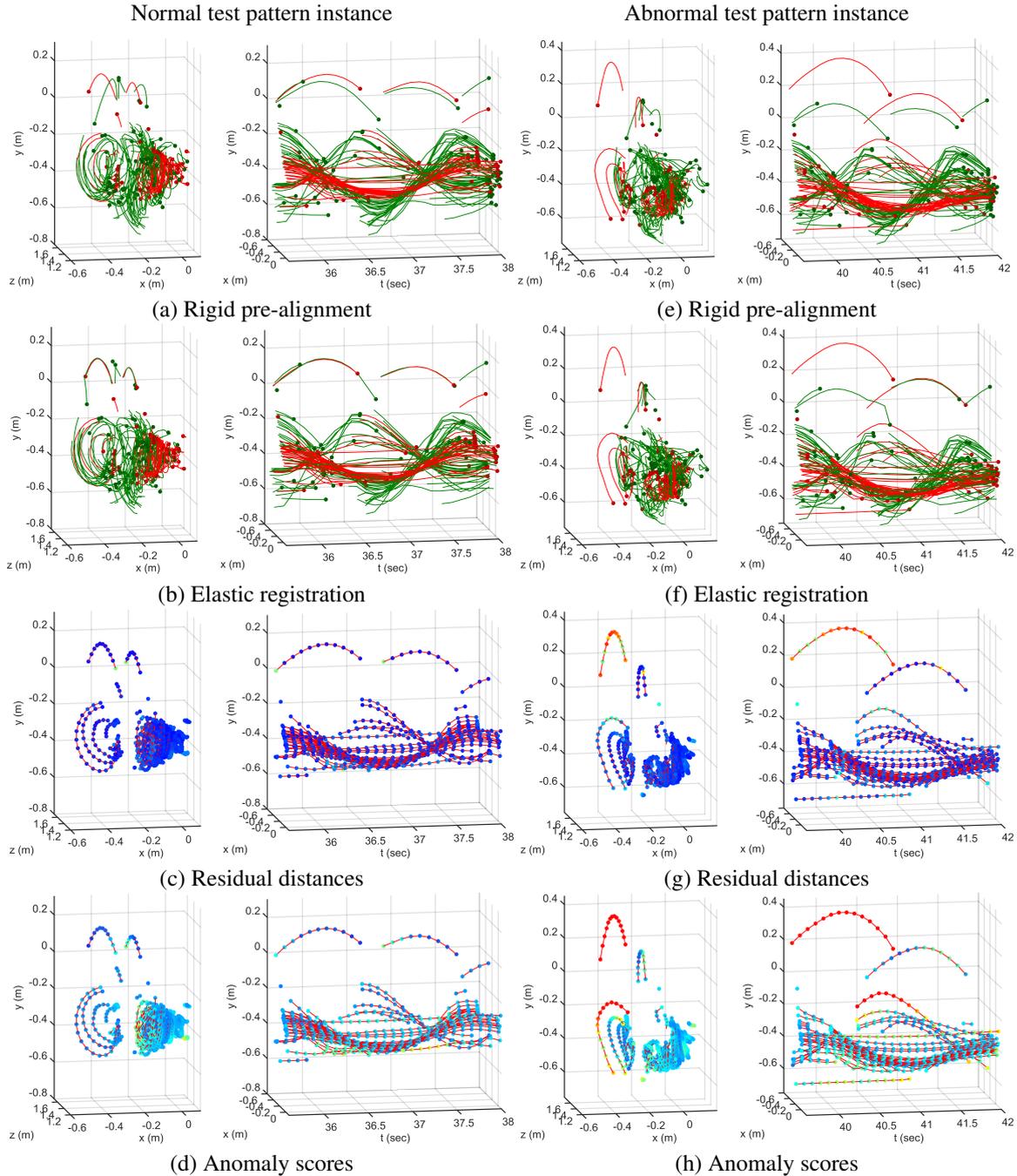


Figure 2.12: Elastic registration of juggling patterns using the example of a normal and abnormal instance of the test pattern. **(a-d)** Registration of the prototype pattern (*green*) to a normal instance of the test pattern (*red*). **(e-h)** Registration to an abnormal instance of the test pattern; Both instances originate from a long test sequence (depicted in figure 2.15, time markers at $t = 37$ sec and $t = 41$ sec). For visualization xyz -plots (*left columns*) and xyt -plots (*right columns*) of super-trajectories are used. **(a,e)** Rigidly pre-aligned prototype pattern. *Dots* mark trajectory end points. **(b,f)** Elastically registered prototype pattern. **(c,g)** Residual distances for each test pattern point, color-coded from *blue* to *red* for the range 0 to 50 cm. Only the test pattern is plotted. **(d,h)** Anomaly scores for each test pattern point, color-coded from *blue* to *red*.

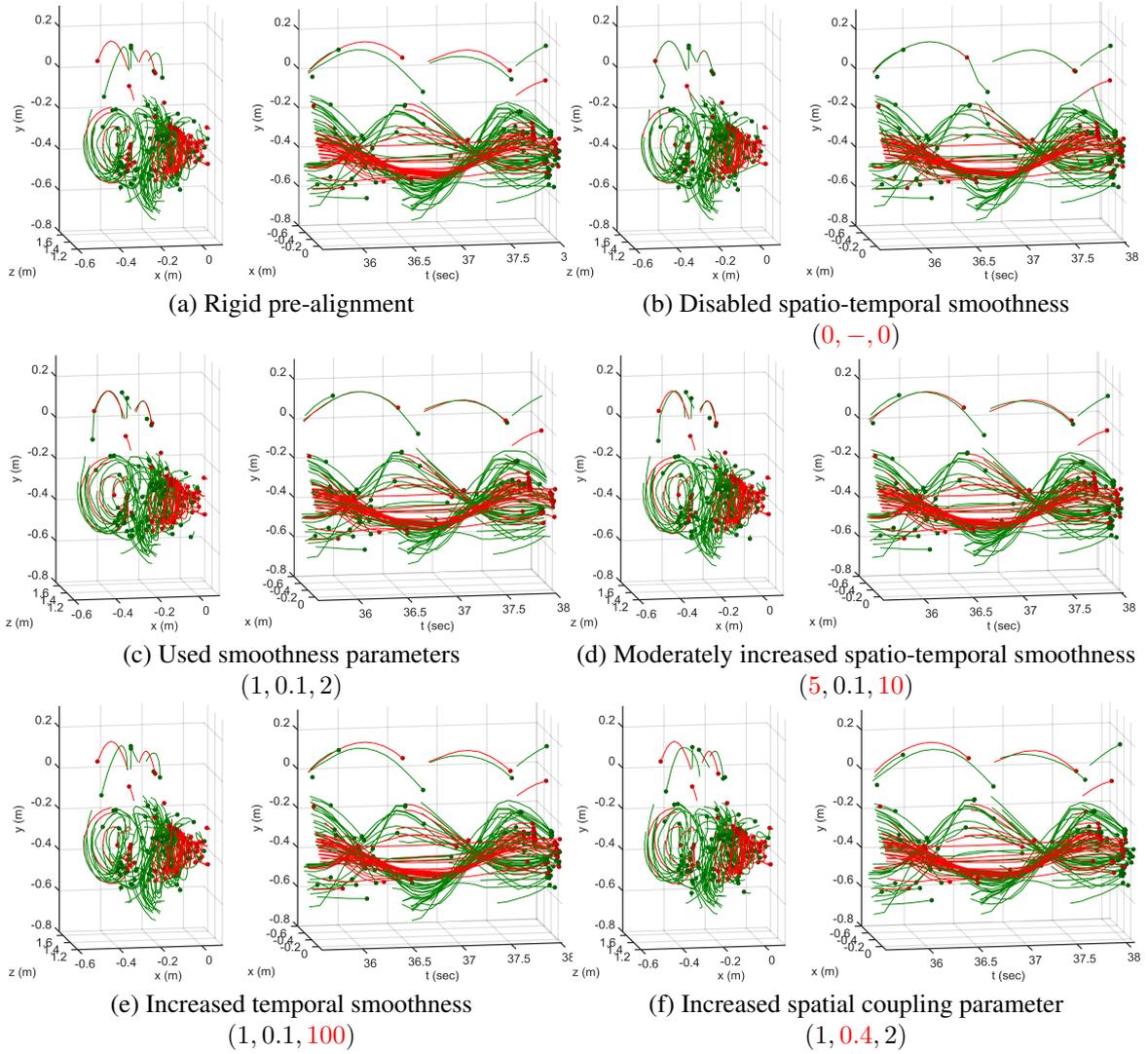


Figure 2.13: Effect of varying smoothness parameters on the elastic registration of juggling patterns. The parameter triple $(\alpha_{\text{spatial}}, \sigma_{\text{spatial}}, \alpha_{\text{temp}})$ is considered. **(a)** Rigidly pre-aligned prototype pattern (green) and test pattern (red). Dots mark trajectory end points. **(b-f)** Registration results for varying spatial and temporal parameters. The parameters used in our experiments are applied in (c).

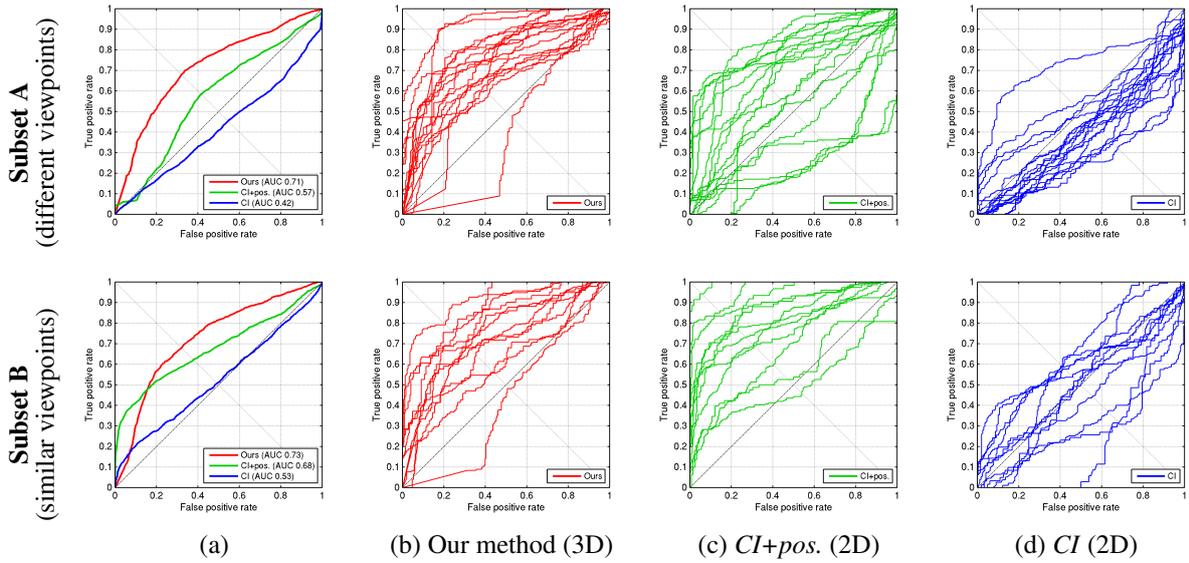


Figure 2.14: Anomaly detection ROC curves. Comparison of our method and the chaotic invariants (CI) (Wu et al., 2010) for the two subsets A (top row) and B (bottom row). (a) Average performance of all methods. (b-d) Performance on single sequences for each method.

receives high residual distances and high anomaly scores in figure 2.12f-g. Note that, although the deviation of the wide arm motion is not very strong (color-coded in light blue) anomaly scores are still very high. This is attributed to fact that deviations are evaluated relatively to the locally observed deviations during training, which is rather low in case of arm motions.

Figure 2.13 demonstrates the effect of varying spatial and temporal smoothness parameters on elastic registration. When smoothness is disabled, especially temporally unsmooth results are obtained, see the registered ball trajectories in figure 2.13a. Applying the smoothness parameters we used in our experiments, *i.e.* ($\alpha_{\text{spatial}} = 1, \sigma_{\text{spatial}} = 0.1, \alpha_{\text{temp}} = 2$), results in good registration, see figure 2.13c. Registration results seem to be robust in a certain range, see the result with moderately increased smoothness parameters in figure 2.13d. Strong increase of temporal smoothness results in temporally stiff and trajectory-wise rigid transformations (figure 2.13e). Increase of the spatial coupling parameter results in spatially stiff transformations (figure 2.13f). Both cases yield too stiff transformations that are not able to fit the test pattern well, which can be observed especially well for the ball trajectories.

Anomaly detection We split the juggling pattern test set into two subsets: Subset A contains 17 sequences with different viewpoints, *i.e.* the patterns are differently aligned compared to the training data. In contrast, subset B contains 12 sequences with similar viewpoints and well aligned patterns in 2D. ROC curves in figure 2.14 show the comparison against the original 2D version of CI. The average performance for each subset and the performance on single sequences is shown. Table 2.2 lists the average performance for both subsets and all sequences, and includes results for the 3D version of CI.

The results in figure 2.14 and table 2.2 show that *CI* has problems with the strong variation between the training and test data caused by different jugglers. *CI+pos.*, which uses the absolute position, achieves a competitive performance to our approach for subset B, where the absolute position is a valuable feature. However, on subset A, where the test data includes different viewpoints, the absolute position is a rather weak feature and our method, which is invariant to changing viewpoint, clearly outperforms both *CI* and *CI+pos.*

To emphasis yet again, our approach outperforms *CI* due to its robustness to strong variations in viewpoint, which is provided by normalization through detection and elastic registration of the proto-

Table 2.2: Anomaly detection results of our method and the chaotic invariants (CI). The average performance is given by the area under curve (AUC).

Method	Subset A	Subset B	All sequences
<i>CI</i> (2D) (Wu et al., 2010)	0.42	0.53	0.46
<i>CI+pos.</i> (2D) (Wu et al., 2010)	0.57	0.68	0.62
<i>CI</i> (3D)	0.42	0.41	0.42
<i>CI+pos.</i> (3D)	0.51	0.55	0.53
Our method (3D)	0.71	0.73	0.72

type pattern that compensates these variations. Moreover, *CI* do not seem to be able to differentiate between normal and abnormal juggling patterns, *e.g.* an normal versus an abnormal ball trajectory, probably because both share similar features in a local context. However, our approach exploits the global context of the prototype pattern to distinguish normal and abnormal trajectories.

Also the use of 3D trajectories in *CI* does not improve results. While using 3D in *CI* adds another dimension compared to 2D, it still does not add the capabilities of our approach, such as normalization by detection and elastic registration, and detecting anomalies in the global context of the prototype pattern. Our method outperforms all tested variants of *CI* to its accurate modeling of the normal pattern variation.

Anomaly localization We evaluate anomaly localization qualitatively from results on entire juggling sequences. We utilize 2D+time projections as illustrated in figure 2.15 on the one hand, and supertrajectories directly rendered into juggling video sequences as shown in figure 2.15d and figure 2.16 on the other hand. Supertrajectories are plotted (time-delayed in case of videos) and color-coded from blue to red by anomaly scores ranging from zero to one.

Both figure 2.15 and figure 2.16 show several examples and scenarios that demonstrate a precise localization of anomalous motion trajectories. Moreover, results appear robust in the presence of strong viewpoint variations and spatio-temporal deformations of the juggling pattern. Furthermore, the approach appears robust to the inter-subject variation from different persons performing the same juggling pattern (see figure 2.16). In addition, figure 2.15a gives an example of a full juggling pattern sequence plotted in xyt-view, and the direct relation between residual distances and anomaly scores is shown in figure 2.15b-c. Further qualitative anomaly detection and localization results are contained in figure 2.6b-g and in figure 2.12c-d,g-h.

Video results are contained in movie 2.17 that complement the results presented in this section. The movie shows the prototype and training sequences and anomaly detection results for five test sequences. Anomaly detection results are rendered into the sequences and give a good impression of the anomaly localization capability. They include the scenario of a moving juggler and two persons juggling side-by-side.

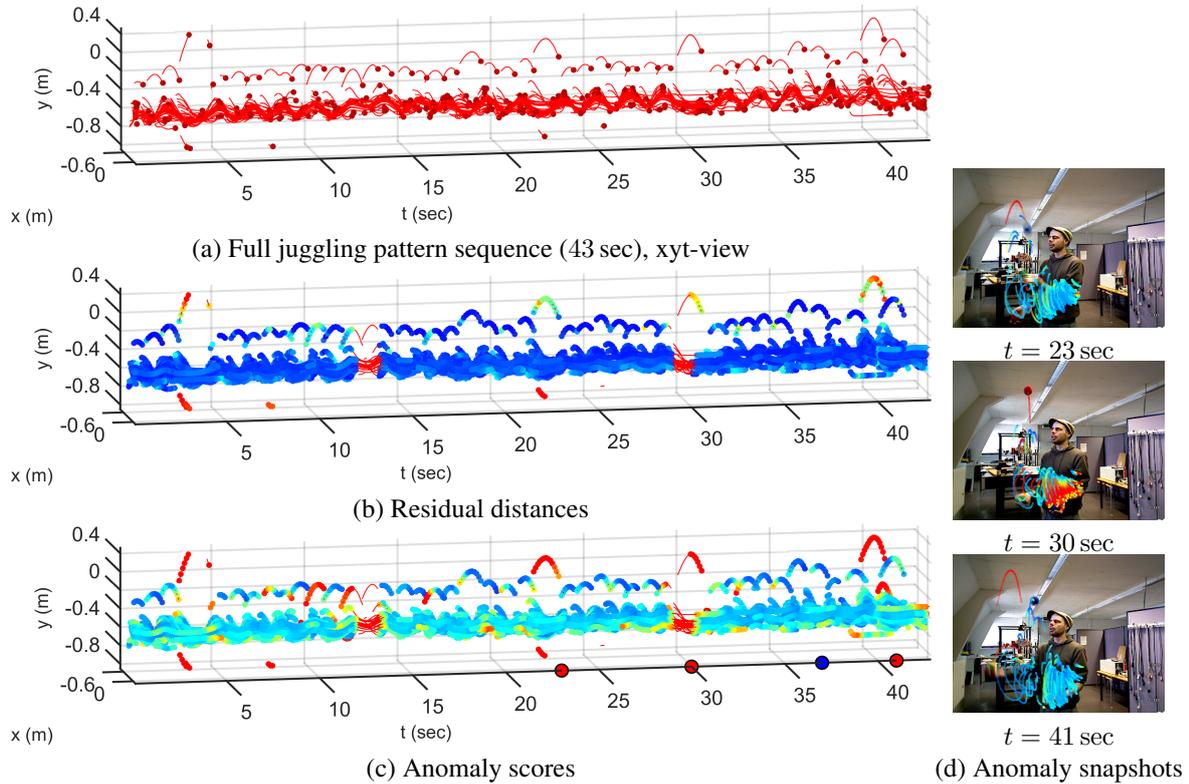


Figure 2.15: Anomaly localization results show that anomalous motion trajectories can be precisely localized within the entire course of a long sequence of repeating juggling cycles. Corresponds to "Example Sequence 2" in movie 2.17. **(a)** 2D+time projection (xyt-view) of a full juggling pattern sequence. Ball parabolas including anomalous throws (*top*), dense arm and hand motion (*bottom*). Dots mark trajectory end points. **(b)** Residual distances for associated test pattern points, color-coded from *blue* to *red* for the range 0 to 50 cm. **(c)** Anomaly scores for associated test pattern points, color-coded from *blue* to *red*. **(d)** Anomaly video snapshots, time points marked on time-axis in (c) by *red markers*. Supertrajectories plotted time-delayed (≈ 1.5 sec) and color-coded (anomaly scores range from *blue* to *red*).



Figure 2.16: Anomaly localization results show that anomalous motion trajectories can be precisely localized within the juggling patterns. Supertrajectories are plotted time-delayed (45 frames, ≈ 1.5 sec) and color-coded (anomaly scores range from *blue* to *red*). **(a)** The anomalous high and outside ball throws are correctly detected in various test patterns. **(b-e)** Anomalies are correctly identified in various complex scenarios.



Movie 2.17: Additional anomaly detection results. The movie shows the prototype and training sequences and anomaly detection results for five test sequences. Anomaly detection results are rendered into the sequences and give a good impression of the anomaly localization capability. They include the scenario of a moving juggler and two persons juggling side-by-side. (Available online at SpringerLink: http://www.edge-cdn.net/video_1073989?playerskin=37016)

2.7.3 Anomaly Detection in Biological Motion Patterns

Application

We show an application of our motion anomaly detection approach on real biological motion patterns. We use 3D+time trajectory data showing global endodermal cell dynamics in the early development of zebrafish embryos (Schmid et al., 2013), see figure 2.18. During this early developmental stage (referred to as "gastrulation" in developmental biology) cells, being located on a spherical surface, "perform a random walk to spread on the surface, followed by convergence and extension movements to gather along the embryonic axis" (Schmid et al., 2013). The data consists of two groups, the wild-type patterns, see figure 2.18a-b, which represent the normal developmental motion pattern, and the *cxcr4a* morphant patterns, see figure 2.18c-d, which result from genetically modified embryos. The goal is to identify, describe and quantify significant differences with regard to the wild-type that result from genetic modification. In Schmid et al. (2013) extensive analysis and quantification is presented comparing both groups. Here, we apply our approach to the data and investigate differences in terms of motion anomalies.

Dataset

The data consists of 12 wild-type (WT) embryos (normal patterns) and 12 *cxcr4a* morphant (MO) embryos (genetically modified). Figure 2.18 shows exemplary motion patterns for each group. The 3D+time trajectory data is plotted as 3D rendering with time color-coded (from blue to red), see figures 2.18a,c, and as 2D Mercator projection respectively, see figures 2.18b,d. The trajectory data is given as 3D point sequences, in which 3D points represent cell locations. All points are located on a spherical surface.

Implementation

All datasets are temporally realigned using manual time point markers (from Schmid et al. (2013)), such that the three developmental stages denoted as *60% epiboly*, *75% epiboly* and *tailbud stage* are matching. We resample all trajectory data to 100 time points. To learn a prototype model of normal motion, we select one WT pattern as the prototype and learn the model using 7 further WT patterns. We exclude 2 WT patterns from model construction, because they obviously constitute outliers. The

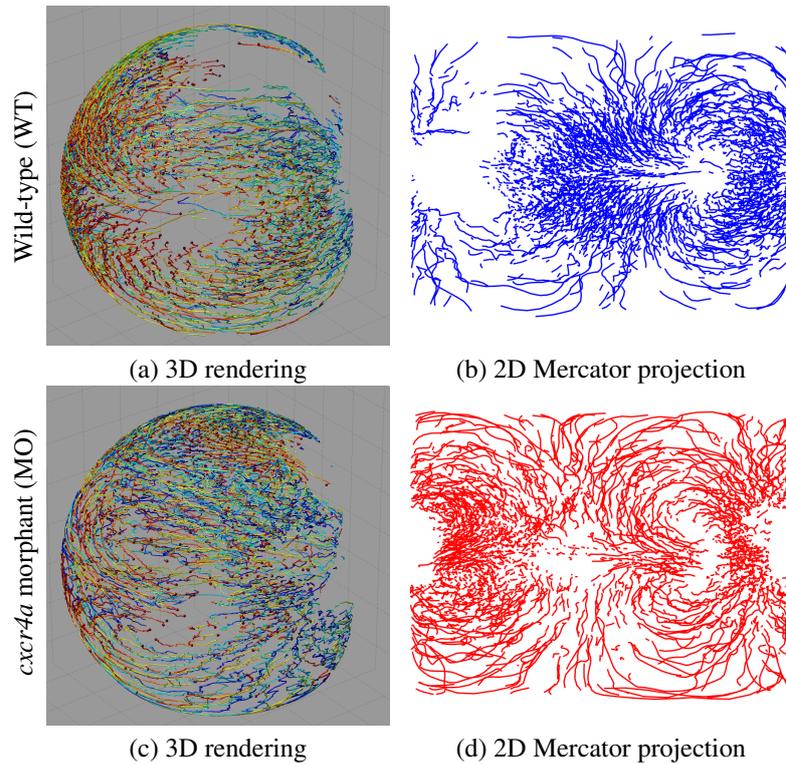


Figure 2.18: Biological motion patterns of cell dynamics in the early development of zebrafish embryos. **(a-b)** A wild-type (WT) pattern. **(c-d)** A *cxcr4a* morphant (MO) pattern. **(a-c)** 3D rendering with time color-coded (from *blue* to *red*). **(b-d)** 2D Mercator projection.

two remaining WT patterns are kept for validation in the test phase. For testing and evaluating motion anomalies we use the 12 MO patterns.

Results

We evaluate motion anomalies from different perspectives and on different scales. Our results concerning global, temporal and spatio-temporal motion anomaly detection are presented in the following.

Global anomaly detection By computing global anomaly scores for each pattern we are able to quantify significant differences between the WT and MO motion patterns. Figure 2.19a,b depicts global anomaly scores for all tested patterns (12x MO, 2x WT).

The tested WT and MO patterns can be clearly distinguished when using the 99%tile for computing global anomaly scores from all pointwise anomaly scores, see figure 2.19a. Compared to using the maximum value, using an upper percentile gives a more robust measure in presence of noise and outliers. Using the 98%tile yields a more sensitive result that better reflects variations within the MO patterns, see figure 2.19b. In figure 2.19c we show how global anomaly scores change with varying percentile value, *i.e.* with varying sensitivity threshold. A clear difference between the tested MO and WT patterns can be observed. The corresponding curves for the training patterns are plotted in figure 2.19d. Excluding 1% outliers, *i.e.* using the 99%tile, yields a clear separation of WT and MO patterns, as can be seen in figures 2.19c,d and figure 2.19a.

Temporal anomaly localization We analyze the temporal course of motion anomalies by computing anomaly profiles. For each time point all pointwise anomaly scores are projected to a single score using the 98%tile, again we use a robust measure as for computing the global anomaly scores.

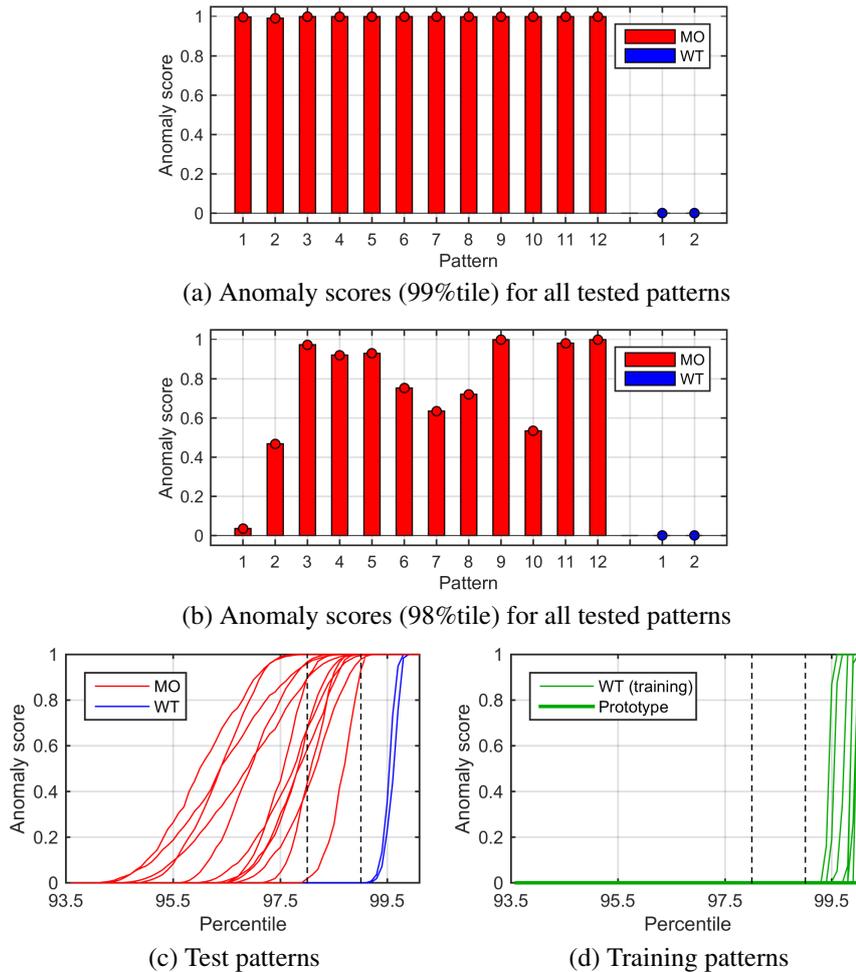
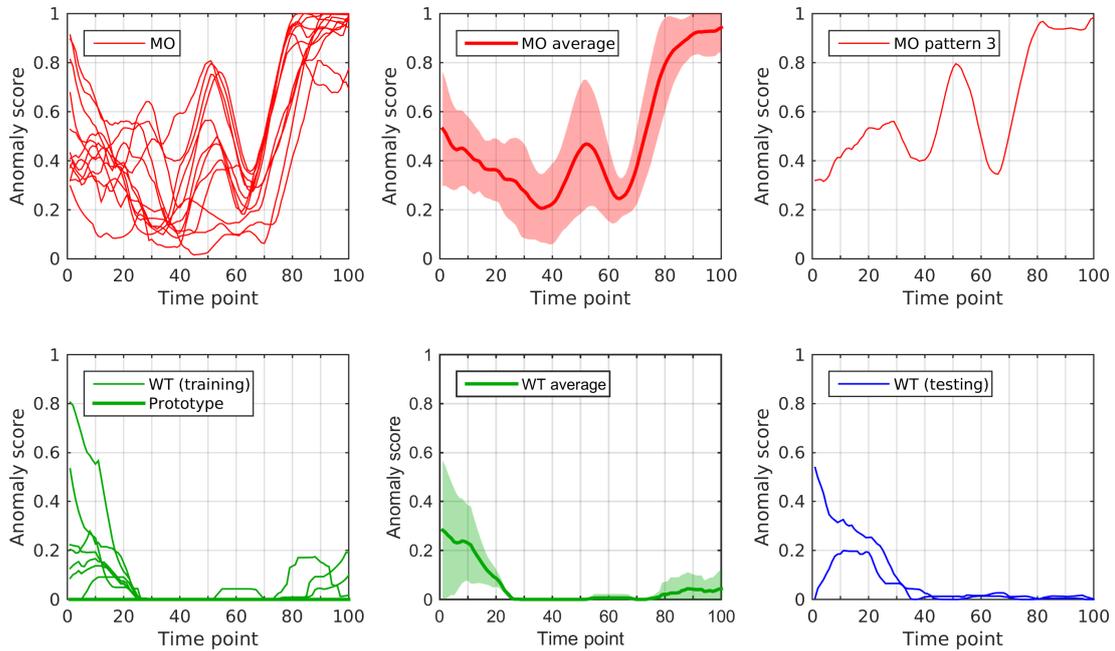


Figure 2.19: Global anomaly scores clearly distinguish WT and MO patterns. **(a)** Using the 99%tile clearly distinguishes WT and MO patterns. **(b)** Using the 98%tile yields a more sensitive result that better reflects variations within the MO patterns. **(c-d)** Variation of anomaly scores with varying percentile value. **(c)** Tested WT and MO patterns are clearly separated. **(d)** Both training and test patterns suggest an outlier level of 1% for WT patterns. The corresponding 99%tile, as well as the 98%tile, are indicated by *dashed lines*.

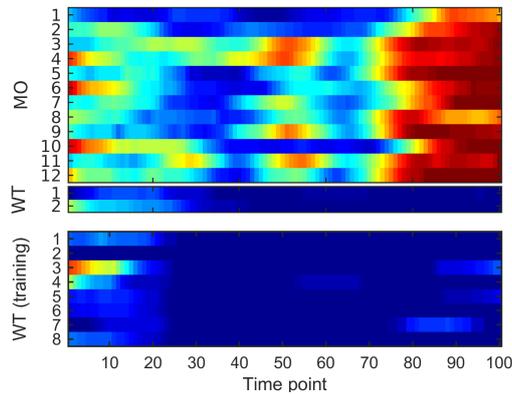
The resulting anomaly profiles are plotted in figure 2.20 and reconfirm a clear difference between WT and MO patterns.

Moreover, an analysis of the temporal course of motion anomalies is possible. Figure 2.20a(bottom row) illustrates anomaly profiles for all WT training patterns (in green). Overall WT anomaly profiles for both test and training patterns have very low scores over the main time course. However, on average (see WT average plot) an anomaly score of about 0.3 which decreases to zero within the time window $[0, 30]$ can be observed. Furthermore, there is a strong variation among the WT patterns in this time window.

This result can be explained by the following fact: At the beginning of the observed cellular motion, cells are rather randomly distributed over the spherical shell, however with distinct distributions for WT and for MO patterns. Moreover, cells first perform a random walk to spread on the surface, then cellular motion transforms into a more distinct pattern of convergence and extension movements. Our model is encoded using a prototype pattern, for which regular variations are learned. This however cannot represent rather random variations as they appear at the beginning of cellular motion. However, when the motion pattern gradually turns from random walk into a distinct motion, the variations can



(a) Anomaly profiles (line plots).



(b) Anomaly profiles (color-coded anomaly scores from blue to red).

Figure 2.20: Temporal anomaly localization is shown by anomaly profiles that plot the anomaly score along time dimension. **(a)** Anomaly profiles are given as line plots for MO (top row, in red) and WT patterns (bottom row, training patterns in green, test patterns in blue). All single profiles for the MO and the WT training patterns are given (left column). The average and 95%tiles over all MO patterns and WT training patterns (excluding the prototype pattern) are shown (middle column). A single representative MO pattern profile and the two WT test pattern profiles are given (right column). **(b)** All anomaly profiles are additionally shown using color-coded anomaly scores (from blue to red) for better visualization, especially for highly overlapping line plots (see MO patterns in the time window $[0, 60]$).

be explained by our model, and anomaly scores accordingly decrease, see time window $[0, 30]$. Only the prototype pattern itself gives a zero anomaly profile, when evaluated with the model, see prototype zero line plot (in bold green) or the color-coded profile plot of WT (training) pattern 2, respectively.

Figure 2.20a (top row) shows the profiles for MO patterns. Again a high variation between the patterns can be observed, especially in the time window $[0, 60]$, which is conform with the mentioned higher embryo-to-embryo variability for MO embryos. The profiles seem to cluster into different groups, which is easier to recognize with color-coded anomaly profile visualization, see figure 2.20b. On average, see the MO average plot, the time course can be divided into three intervals: T1 ($[0, 39]$), T2 ($[30, 69]$) and T3 ($[60, 99]$). During the first interval, the average profile follows a similar slope compared to the WT average profile. Again the same explanation as for the WT patterns holds. Anomaly scores decrease gradually when random motion turns into a distinct pattern. Compared to WT patterns, an offset of about 0.2 for average anomaly scores in the interval T1 indicates a motion anomaly. The second interval is characterized by an anomaly peak on average. As can be seen from the color-coded profiles most MO patterns have an anomaly peak at this location, however there are a few outliers, specifically MO patterns 1, 2 and 10. In the third interval anomaly scores are consistently increasing to a very high score for all MO patterns and the variation decreases. A single profile plot is given for MO pattern 3, which is a representative for the cluster of profiles that have a strong peak at around time point 50.

Spatio-temporal anomaly localization We further compute the anomaly density on the spherical shell and accumulate densities for time interval T1-T3 to localize motion anomalies on a spatio-temporal scale, see figure 2.21. The anomaly density is computed by accumulating anomaly scores into bins on the spherical surface, parametrized by spherical coordinates (polar angle θ and azimuthal angle φ). The values are finally normalized by the area of the spherical bins. Figure 2.21 shows the anomaly density in comparison to the cell density in the three consecutive intervals T1-T3. Both the average over all MO patterns, as well as a single pattern (MO pattern 3) is shown. In interval T2 a clear motion anomaly spot is localized on the left side of the plot (anterior side of the embryo). It corresponds to the temporal peak observed in figure 2.20. The spot is increasing to a larger motion anomaly region in interval T3, which corresponds to high anomaly scores at the end of the anomaly profiles in figure 2.20.

Overall, the proposed approach quantifies significant differences between wild-type and morphant motion patterns and localizes motion anomalies in the spatio-temporal domain.

2.8 Conclusion

We have presented a new approach to motion anomaly detection in complex motion patterns. The problem is formulated as a reconstruction task, in which a test sample is explained using prototype patterns. Prototype patterns may appear anywhere in a video and in any orientation. Therefore, our approach starts with a robust detection that is invariant to rigid transformations, followed by a spatiotemporal elastic registration of the prototype pattern to the test pattern. The precise alignment of the patterns allows us to detect and localize subtle anomalies, as demonstrated by experiments on a 3D motion anomaly dataset. An important application area for our approach is in biomedical image analysis, with complex developmental and growth patterns to be compared.

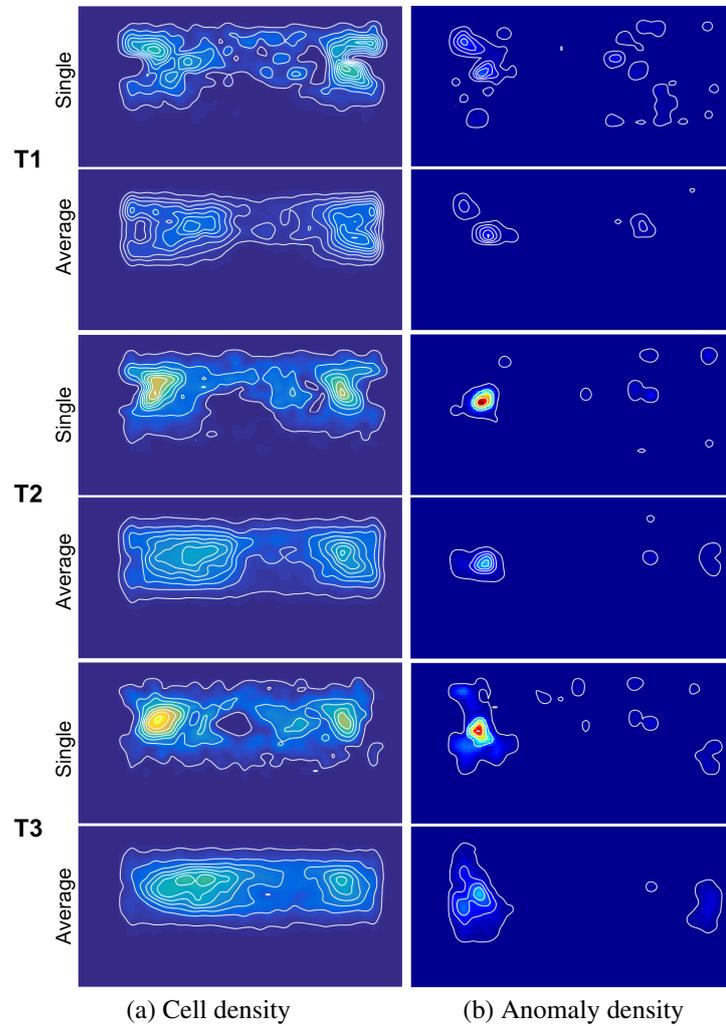


Figure 2.21: Spatio-temporal motion anomaly localization from anomaly density and cell density plots. The three time intervals $T1$, $T2$ and $T3$ are considered, and for each the average over all MO patterns and a single MO pattern (pattern 3) is shown. **(a)** Cell density is color-coded (from *blue* to *yellow*) and indicated by *isolines*. **(b)** Anomaly density is color-coded (from *blue* to *red*) and indicated by *isolines*. (All plots represent a map of the spherical surface, with spherical coordinate φ on the x-axis and θ on the y-axis. The equator is located on the *mid-horizontal line*.)

3 Spatiotemporal Geometrical Models for Motion Pattern Detection

In this chapter, we present a method for motion pattern detection using a spatiotemporal geometrical model. The method was developed to detect specific motion patterns in dense cellular motion in 4D (3D+time), namely *cell intercalations*. Cell intercalations occur when a cell enters the space between adjacent cells, as depicted in figure 3.1. Cell intercalation is an important mechanism in tissue formation in developmental biology, and thus of great interest in that field.

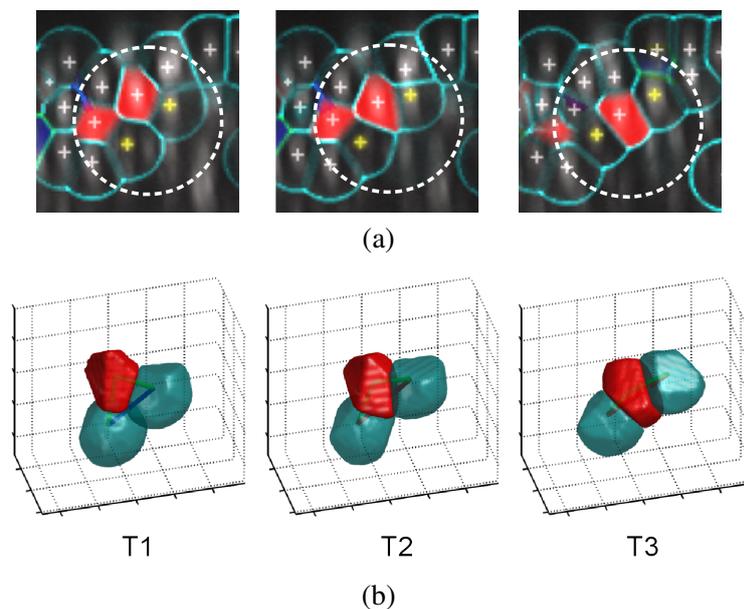


Figure 3.1: Cell intercalation in developmental biology using the example of embryonic development in zebrafish (from [Bensch et al. \(2013\)](#) with modifications). **(a)** Lateral views from 3D time-lapse recordings show a cell (marked in *red*) intercalating between two adjacent cells (*yellow crosses*) in three stages (T1 to T3). The raw data contains the nuclei fluorescence (in *grey*). Tracked cell nuclei positions are marked with *crosses* and calculated cell boundaries are highlighted in *cyan*. **(b)** 3D rendering of the cell triple highlighted in (a). More details are given in figure 3.9.

We use sparse motion trajectories of tracked single cells to represent these motion patterns. Cell intercalations are modeled using a spatiotemporal geometrical model, which is coarsely illustrated in figure 3.2 and described in detail in section 3.2.

Detecting motion patterns that are similar to the defined intercalation model is challenging for several reasons: 1) Position measurements of cellular motion from biological data are often very noisy or interrupted, 2) There is a strong variability within the class of possible intercalations, 3) Compared to planar intercalations in 2D+time, intercalations in 3D+time are more complex. We developed a method to robustly detect cell intercalations, which is presented in section 3.3.

The method presented here originates from our work in the field of developmental biology which we published in *Biology Open* ([Bensch et al., 2013](#)). The workflow of image analysis is shown in figure 3.3. There the method was applied for the detection of cell intercalations but the description of

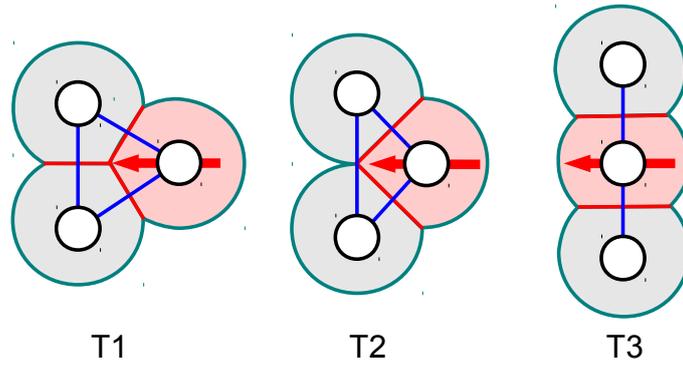


Figure 3.2: Spatiotemporal geometrical model of cell intercalation for a cell triple in three stages (T1 to T3). The *active* cell (marked in *red*) intercalates between two neighboring cells by pushing through in the indicated direction (*red arrow*). In the initial stage (T1) the cell centers are arranged in an equilateral triangular configuration. When the active cell pushes through, the neighboring cells lose contact in stage T2. The cell triple approaches a linear configuration in stage T3. More details are given in figure 3.5.

the method itself was limited by the scope of the journal. Therefore, in this work, we present details and focus on the methods for modelling and detecting cell intercalations (highlighted in green in figure 3.3).

In the following, we first introduce the biological motivation and related work in section 3.1. Sections 3.2 and 3.3 describe the intercalation model and the detection of intercalations. In section 3.4 we summarize the experiments conducted and the results obtained in Bensch et al. (2013), together with application specific method and implementation details. Section 3.4 is published in Bensch et al. (2013) in large parts.

3.1 Biological Motivation and Related Work

In the development of biological organisms, cellular reorganization is important for reshaping tissue and thus driving morphogenesis, which “refers to the cell and tissue movements that give the developing organ or organism its shape in three dimensions” (Slack, 2005, p. 6). Specific mechanisms exist that promote morphogenesis. Cell intercalation is one of them and denotes the process of cells entering the space between neighboring cells. This results in a local cellular reorganization. As soon as a larger number of cells performs coordinated cell intercalation, tissue thinning, for example, can occur. Several examples to intercalation patterns and corresponding types of tissue reshaping can be found in Keller et al. (2000).

Figure 3.4 illustrates the importance of the cell intercalation mechanism in developmental biology using the example of the embryonic development of the zebrafish. Radial cell intercalation occurs “with a concomitant thinning of the deep cell layer” (see figure 3.4b), (Bensch et al., 2013, p. 845). The process called *doming* initiates epiboly and states an important stage during the embryonic development of the zebrafish (marked in green and blue in figure 3.4).

In Bensch et al. (2013), we conducted a quantitative comparison of cell intercalations and their motion statistics between wildtype and mutant embryos of zebrafish. Due to the complexity of the 4D cellular motion patterns at hand—more than 1000 cells per recorded embryo perform complex intermingled motion patterns, see supplementary material movies 3.20 and 3.21—a manual analysis is impractical. Moreover, a manual analysis is always affected by a user specific bias and is not repeatable. For these reasons, an automated method for quantification is required, specifically automated detection of cellular events in this case.

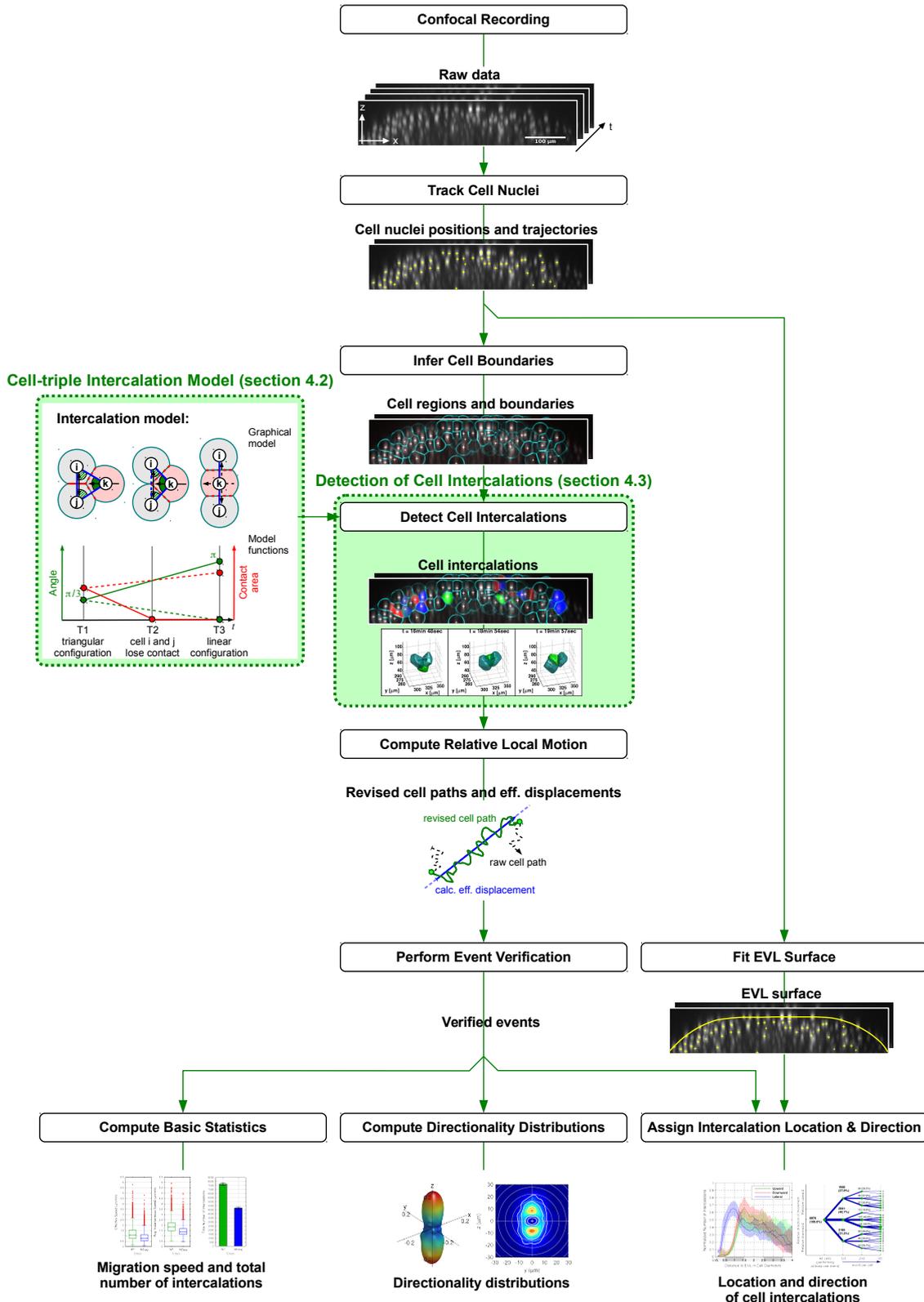


Figure 3.3: Workflow of image analysis (from Bensch et al. (2013) with modifications). The workflow of image analysis is schematically presented: Starting from the raw data input (*top*) the successive steps performing image analysis and data evaluations are presented. Details on the algorithms highlighted in *green* are contained in section 3.2 and section 3.3. Details on the remaining algorithms are contained in section 3.4.2 in the subsections with the corresponding headings. In the data flow intermediate results are illustrated that serve as input for the subsequent algorithm. Final results are presented at the *bottom*.

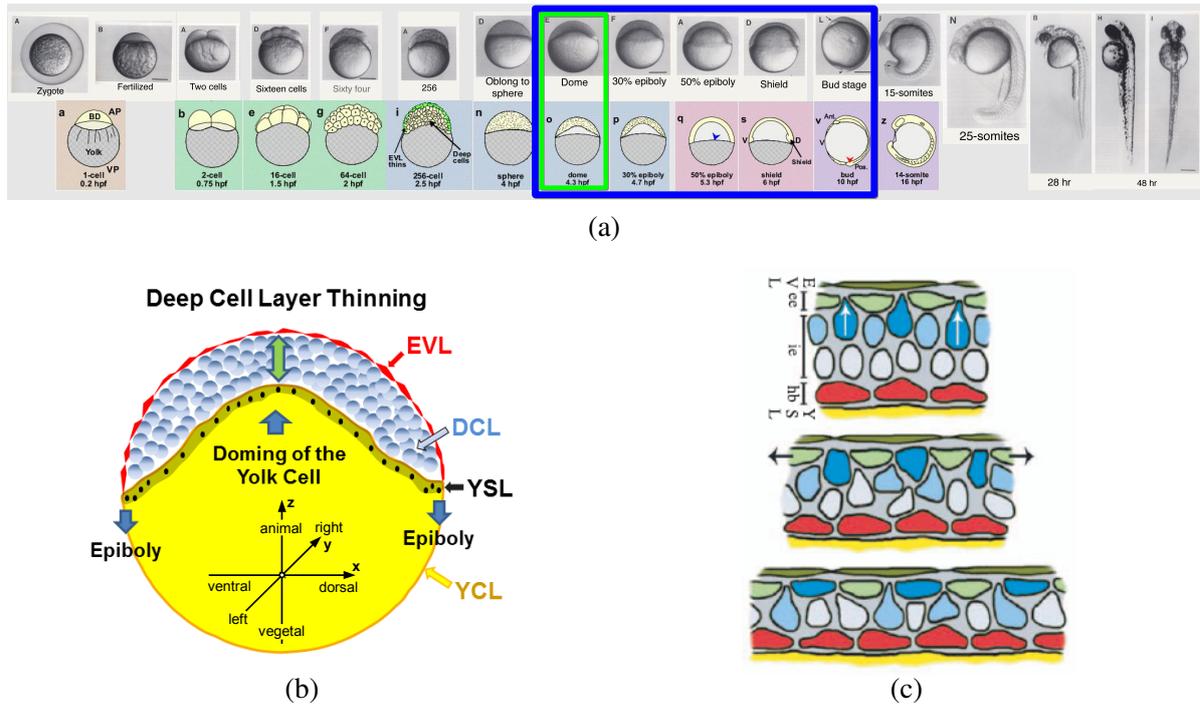


Figure 3.4: Importance of the cell intercalation mechanism in developmental biology. **(a)** Stages of embryonic development of the zebrafish (from Kimmel et al. (1995); Webb and Miller (2007) with modifications). From single-cell stage (left) to dome stage (middle, marked in green), epiboly stages (marked in blue) and a recognizable fish embryo (right). **(b)** Animalward doming of the yolk cell initiates epiboly of the zebrafish embryo (from Bensch et al. (2013)). **(c)** Radial intercalation and cell layer thinning during epiboly. Cells of the interior layer of the epiblast (blue, light blue, white) intercalate between the cells of the exterior layer (light green) (from Kane et al. (2005)).

Early works, such as Irvine and Wieschaus (1994), had to rely on manual analysis from video playback. Photographs were taken at intervals from the video monitor and cell rearrangements were reconstructed by following individual cells during playback of videotapes and marking their positions onto the photographs. Fortunately, today's methods are more elaborate. In a recent work by Lienkamp et al. (2012) for example, a method is presented that automatically detects cellular rosette structures in vertebrate kidney morphogenesis. In this developmental process rosette formation is a characteristic event when remodelling of 4-cell junctions leads to tissue elongation. Other recent works that analyse cell intercalations in developing tissue, present methods that measure local tissue deformation. In Blanchard et al. (2009) a kinematic framework is presented that measures domain deformation in terms of the relative motion of cell positions and the evolution of their shapes. Local strain rates defined by a domain translation vector and a tissue velocity gradient tensor are computed. Also in Heller et al. (2014) quantitative analysis of tissue deformation is performed. For determining cell-cell coordination a Delaunay triangulation is used to connect cells to their nearest neighbors and the cosine similarity between velocity vectors of cell pairs is computed.

In contrast to existing approaches that mainly deal with 2D (planar) intercalations, we present a method that is able to detect cellular events in $3D+time$. Obviously, this is very important because developmental processes occur in 3D by nature and often cannot be represented in 2D. Furthermore, our approach allows to define an explicit prototype model representing the class of events to be detected, see figure 3.5 for example. Based on this, our method is able to detect all similar event in 3D space and time under spatiotemporal variations from noisy input data (cellular motion trajectories).

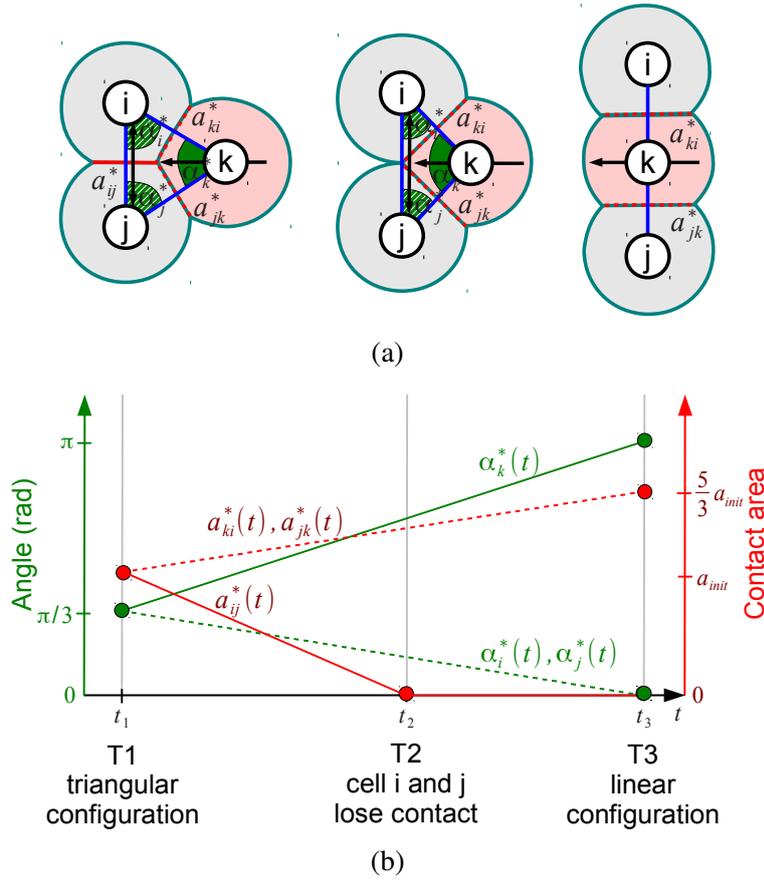


Figure 3.5: Model of cell intercalation for a cell triple (i , j , and k) in three stages (T1 to T3) (from Bensch et al. (2013) with modifications). **(a)** Geometrical model: The cell k intercalates between the neighboring cells (i and j). Pairwise distances (blue), enclosing angles (green), contact areas (red) and motion directions (black arrows). **(b)** Model functions: Enclosing angles (green) and contact areas (red) across the three stage (T1 to T3). In the initial configuration (T1) all angles as well as all contact areas are equal. When cell k pushes through, cell i and j lose contact in configuration T2, i.e. contact area a_{ij}^* (red solid) decrease to zero. When approaching the linear configuration (T3) angle α_k^* (green solid) converges to π .

3.2 Cell-triple Intercalation Model

Cell intercalations are modeled based on the intercalation model depicted in figure 3.5. It is applied in the overall workflow as shown in figure 3.3. This model describes an intercalation in terms of the angles (α_i^* , α_j^* and α_k^*) and contact areas (a_{ij}^* , a_{jk}^* and a_{ki}^*) of three cells (i , j , and k). The three cells start in a triangular configuration (time point T1) and end in a linear configuration (time point T3). Time point T2 models the point when cells i and j lose their contact. The start and end values of the six features for an ideal intercalation were manually defined by inspecting several clear intercalation events in the data set. Furthermore, we assume a linear transition for all features from the start to the end values, except for a_{ij}^* , which drops to zero between T1 and T2 and stays zero between T2 and T3. This function models the characteristic event, when cells i and j lose their contact at time point T2. We point out that using angles and contact areas as local features provides invariance to 3D translation and rotation in this detection task.

Formally, the model functions are defined as one-dimensional continuous functions, with the angles being defined by $\alpha : \mathbb{R} \rightarrow \mathbb{R} : t \mapsto \alpha(t)$ and the contact areas being defined by $a : \mathbb{R} \rightarrow \mathbb{R} : t \mapsto a(t)$.

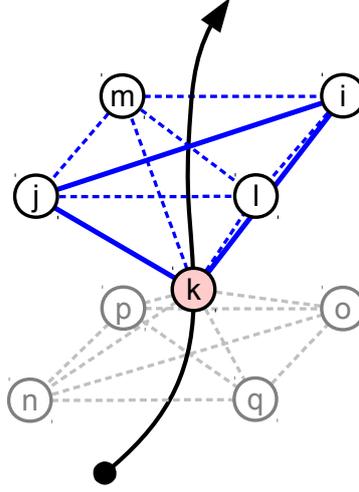


Figure 3.6: Cell intercalation in 3D. Cell k (marked in red, cell boundaries are omitted) moves through an 8-neighborhood of cells i, j, l and m . This involves intercalations, not only in a single cell-triple, e.g. (i, j, k) (blue edges), but in five further cell-triples (dashed blue edges). In this example, 28 cell-triples have to be considered in total. For clarity, only 12 cell-triples are marked.

Specifically, the model functions of the angles are defined by

$$\alpha_k^*(t) = \frac{1}{(t_3 - t_1)} \frac{2}{3} \pi \cdot (t - t_1) + \frac{\pi}{3}; \quad t \in [t_1, t_3] \quad (3.1)$$

$$\alpha_i^*(t) = \alpha_j^*(t) = -\frac{1}{(t_3 - t_1)} \frac{\pi}{3} \cdot (t - t_1) + \frac{\pi}{3}; \quad t \in [t_1, t_3] \quad (3.2)$$

and the model functions of the contact areas are defined by

$$a_{ij}^*(t) = \begin{cases} -\frac{1}{(t_2 - t_1)} a_{\text{init}} \cdot (t - t_1) + a_{\text{init}} & \text{if } t \in [t_1, t_2] \\ 0 & \text{if } t \in [t_2, t_3] \end{cases} \quad (3.3)$$

$$a_{ki}^*(t) = a_{jk}^*(t) = \frac{1}{(t_3 - t_1)} (r - 1) a_{\text{init}} \cdot (t - t_1) + a_{\text{init}}; \quad t \in [t_1, t_3], \quad (3.4)$$

corresponding to the functions shown in figure 3.5b. The parameter $a_{\text{init}} \in \mathbb{R}$ denotes the contact area when the cells are arranged in the initial configuration (T1). For detecting intercalations, contact area functions are normalized, such that a_{init} is no free parameter, see section 3.3 equation 3.6. The parameter $r \in \mathbb{R}$, with $r > 1$, models how much cell contact areas increase, when cells get compressed when pushing through each other. We set $r = 5/3$ in our experiments.

3.3 Detection of Cell Intercalations

For detecting cell intercalations in an unseen data set, cell nuclei are tracked and cell boundaries and contact areas are inferred to extract the features necessary for fitting the intercalation model. The different steps of processing are depicted in the workflow in figure 3.3. In this section, we assume that the trajectories of labeled cell nuclei and the cell contact areas are already given. The experimental section 3.4.2 gives more details about how cell nuclei positions, trajectories and cell contact areas are computed in particular.

When cells intercalate through neighboring cells in 3D, as depicted in figure 3.6, not only a single, but multiple cell-triple intercalations are involved simultaneously. In the example, a cell moves through

an 8-neighborhood of cells and starts intercalating through four cells, which immediately involves six cell-triple intercalations. However, in total 28 cell-triples have to be considered in this example, which shows that cell intercalation in 3D is more complex than 2D planar cell intercalation.

In the following, we explain how we detect 3D cell intercalations based on the cell-triple intercalation model. Basically, for each cell k we fit the intercalation model to all involved cell triples. This yields intercalation scores for all contributing cell triples, which are combined to a single intercalation score for each cell k . The algorithm groups into two steps. First, the intercalation model is fit to every possible cell triple configuration (i, j, k) , where k is the intercalating cell (described in section 3.3.1). Second, for each cell an intercalation score is computed from all contributing cell triples (described in section 3.3.2).

3.3.1 Fitting the Intercalation Model

The intercalation model is fit to every possible cell triple configuration (i, j, k) and to every possible combination of start time t_1 , intermediate time t_2 and end time t_3 . Throughout this and the next section (section 3.3.2) figure 3.7 serves as a practical example. It shows measured contact areas and angles for a single cell triple (see figure 3.7a) and gives concrete examples of the model fitting functions derived in the following.

Energy formulation

We formulate the following total energy

$$E(i, j, k, t_1, t_2, t_3) = E_{\text{Angles}}(i, j, k, t_1, t_3) + \gamma E_{\text{Areas}}(i, j, k, t_1, t_2, t_3)$$

expressing how well the intercalation model fits to a cell triple (i, j, k) with time points (t_1, t_2, t_3) , compare to figure 3.5. The energy consists of two parts, the model fitting energy in terms of the angles and in terms of the contact areas of the intercalation model. The parameter γ controls the influence of the second term. We set $\gamma = 3.8$ in our experiments.

Linear regression

As mentioned before, trajectory measurements from biological data can be very noisy and interrupted. To provide robustness, we fit linear functions to the measured data and compare these to the model functions instead. We denote the functions representing the measured contact area and the measured angle by a and α respectively. The linear functions \hat{a} and $\hat{\alpha}$ are fit by linear regression within the interval $[t_1, t_3]$, such that

$$\begin{aligned} \hat{a}(t) &= \hat{b}_a + \hat{m}_a \cdot t; t \in [t_1, t_3] & \text{with } (\hat{m}_a, \hat{b}_a) &= \arg \min_{m,b} \int_{t_1}^{t_3} (a(t) - b - mt)^2 dt \quad \text{and} \\ \hat{\alpha}(t) &= \hat{b}_\alpha + \hat{m}_\alpha \cdot t; t \in [t_1, t_3] & \text{with } (\hat{m}_\alpha, \hat{b}_\alpha) &= \arg \min_{m,b} \int_{t_1}^{t_3} (\alpha(t) - b - mt)^2 dt. \end{aligned}$$

Concerning model function a_{ij}^* , which is defined piecewise in the interval $[t_1, t_3]$, linear regression is performed for each interval separately, such that

$$\hat{a}_{ij}(t) = \begin{cases} \hat{b}_{a,12} + \hat{m}_{a,12} \cdot t & \text{if } t \in [t_1, t_2) \\ \hat{b}_{a,23} + \hat{m}_{a,23} \cdot t & \text{if } t \in [t_2, t_3]. \end{cases}$$

Figure 3.7b shows an example where linear functions are fit to the measurements of a single cell triple within a certain time interval.

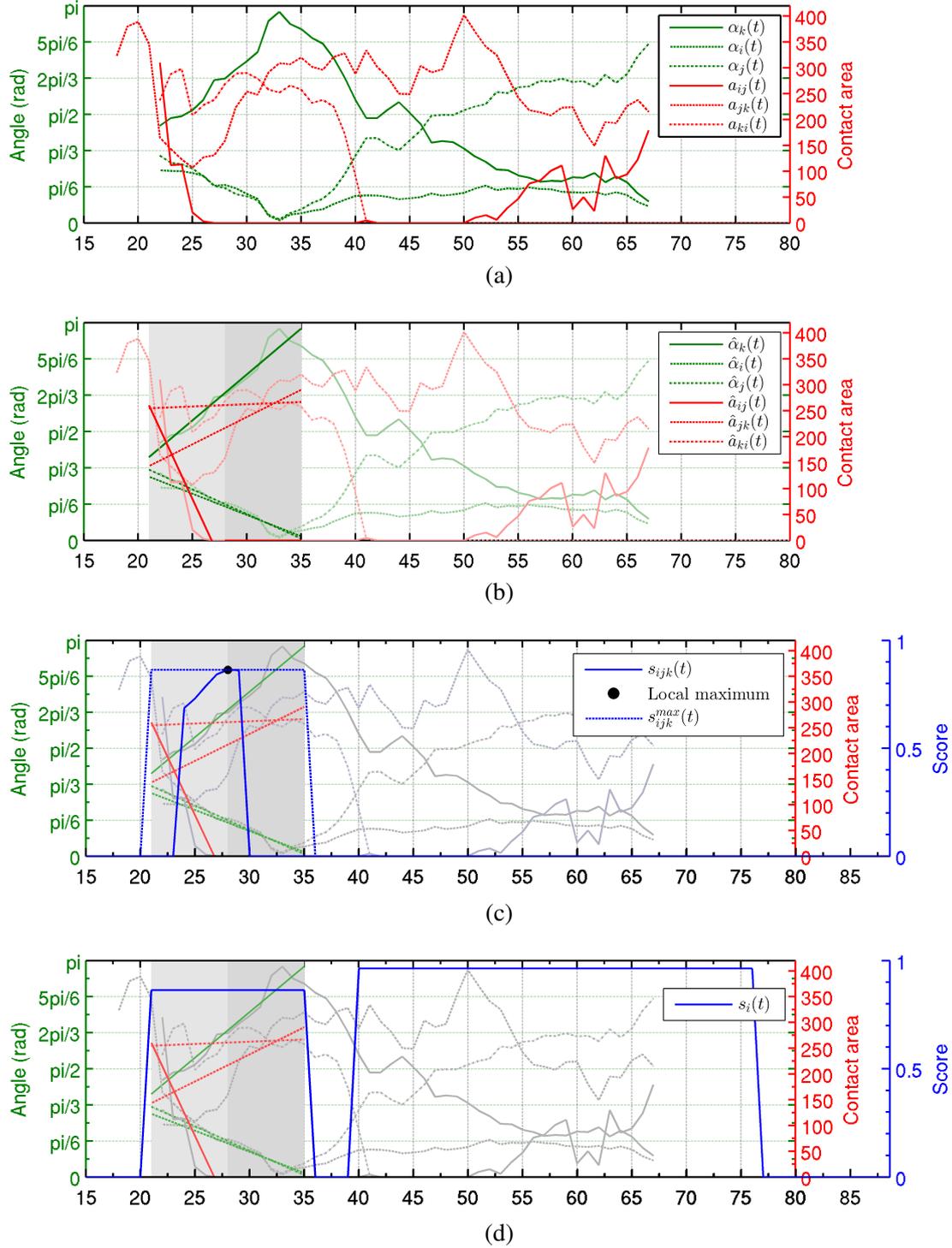


Figure 3.7: Fitting the cell intercalation model, illustrated at a concrete example of a single cell triple. **(a)** Measured contact areas (red) and angles (green). Time (in frames) on the x-axis. **(b)** Linear functions fit by linear regression in the highlighted time intervals (the intervals $[t_1, t_2]$ and $[t_2, t_3]$ are highlighted in two shades of gray). **(c)** Score function (blue, see equation 3.9), detected local maximum at $t = 28$ (black dot) and maximum score function (dashed blue, see equation 3.10). **(d)** Intercalation score function for cell i (blue, see equation 3.11). Note, this also contains information from other cell triples that are not shown here.

First energy term (angles)

The energy for fitting the angles part of the intercalation model is defined by

$$E_{\text{Angles}}(i, j, k, t_1, t_3) = \frac{1}{(t_3 - t_1)} \frac{1}{\pi^2} \frac{1}{3} \int_{t_1}^{t_3} (\hat{\alpha}_k(t) - \alpha_k^*(t))^2 + (\hat{\alpha}_i(t) - \alpha_i^*(t))^2 + (\hat{\alpha}_j(t) - \alpha_j^*(t))^2 dt$$

The energy computes a linear combination of the mean squared errors between the linear model functions and the linear functions fit to the measured data. This energy term is independent of time point t_2 .

Second energy term (contact areas)

The energy for fitting the contact areas part of the intercalation model is defined by

$$E_{\text{Areas}}(i, j, k, t_1, t_2, t_3) = \underbrace{w_{a_{ij}} \left(\frac{1}{(t_2 - t_1)} \int_{t_1}^{t_2} (\hat{n} \hat{a}_{ij}(t) - n^* a_{ij}^*(t))^2 dt + \frac{1}{(t_3 - t_2)} \int_{t_2}^{t_3} (\hat{n} \hat{a}_{ij}(t) - n^* a_{ij}^*(t))^2 dt \right)}_{\text{Energy for fitting piecewise model function } a_{ij}^*} + \underbrace{(1 - w_{a_{ij}}) \frac{1}{(t_3 - t_1)} \left(\frac{1}{2} \int_{t_1}^{t_3} (\hat{n} \hat{a}_{ki}(t) - n^* a_{ki}^*(t))^2 dt + \frac{1}{2} \int_{t_1}^{t_3} (\hat{n} \hat{a}_{jk}(t) - n^* a_{jk}^*(t))^2 dt \right)}_{\text{Energy for fitting model functions } a_{ki}^* \text{ and } a_{jk}^*}. \quad (3.5)$$

It combines the energies for fitting the piecewise model function a_{ij}^* and the model functions a_{ki}^* and a_{jk}^* to the underlying measured data. The energy is computed by linear combination of the mean squared errors between the linear model functions and the linear functions fit to the measured data. The parameter $w_{a_{ij}}$ weights the importance of the model function a_{ij}^* , that contains the characteristic event of cells i and j losing contact. In our experiments we set $w_{a_{ij}} = 0.4$. Note that the second part of this energy is independent of time point t_2 .

We add a normalization of the contact area functions because we do not want a scaling of the cells to affect the model fitting. The contact area model functions are normalized by

$$n^*(i, j, k, t_1, t_2, t_3) = \frac{1}{\int_{t_1}^{t_3} a_{ij}^*(t) + a_{ki}^*(t) + a_{jk}^*(t) dt} \quad (3.6)$$

and the linear functions fit to the data are normalized by

$$\hat{n}(i, j, k, t_1, t_2, t_3) = \frac{1}{\int_{t_1}^{t_3} \hat{a}_{ij}(t) + \hat{a}_{ki}(t) + \hat{a}_{jk}(t) dt} \quad (3.7)$$

respectively. To simplify the notation in equation 3.5 we omitted the arguments of the normalization functions.

Temporal filtering

Given a cell triple configuration (i, j, k) we define a temporal filtering by function $f_{ijk} : \mathbb{R} \rightarrow \mathbb{R} : t \mapsto f_{ijk}(t)$, such that

$$f_{ijk}(t) = \min_{t_1, t_3} E(i, j, k, t_1, t, t_3). \quad (3.8)$$

3 Spatiotemporal Geometrical Models for Motion Pattern Detection

For each t , where t corresponds to the intermediate time point t_2 in the intercalation model, the optimal model fit is computed by minimizing the energy w.r.t. the event start and end time (t_1, t_3) . Consequently, for each t the temporal filtering computes the minimum energy corresponding to the best model fit.

Correspondingly, for each t we denote the arguments (t_1, t_3) that minimize the energy by

$$(t_{1,ijk}(t), t_{3,ijk}(t)) = \arg \min_{t_1, t_3} E(i, j, k, t_1, t, t_3).$$

We further map the values of f_{ijk} to a score in the range $[0, 1]$ by

$$s_{ijk}(t) = \exp(-\lambda f_{ijk}(t)). \quad (3.9)$$

A high score is obtained for a good model fit and a low score for a bad model fit, respectively. We set $\lambda = 1$. In case of incomplete measurements, when a single cell is not tracked within a time interval, for example, the score is defined to be zero. Figure 3.7c gives an example to this case.

Optimization

In our experiments the time domain is discretized by the temporal sampling of the biological recordings. For the optimization in equation 3.8 we use exhaustive search, *i.e.* we compute the energy for each combination of t_1 and t_3 on the set of discretized time points, and extract the minimum. We further restrict the event duration $(t_3 - t_1)$ by defining lower and upper bounds T_{\min} and T_{\max} , such that $T_{\min} \leq (t_3 - t_1) \leq T_{\max}$.

3.3.2 Detecting Cell Intercalations

Cell-triple intercalation events

To detect intercalation events local maxima of the score function in equation 3.9 are computed. We denote detected intercalation events for a cell-triple (i, j, k) by a four-tuple

$$(t_p, s_{ijk}(t_p), t_{1,ijk}(t_p), t_{3,ijk}(t_p)),$$

where t_p is the intermediate time point, $t_{1,ijk}(t_p)$ and $t_{3,ijk}(t_p)$ are the corresponding start and end time points and $s_{ijk}(t_p)$ is the score. We discard events whose time window $[t_{1,ijk}(t_p), t_{3,ijk}(t_p)]$ overlaps with the time window of another event with higher score. A threshold s_{thr} is used to suppress low scored events with $s_{ijk}(t_p) < s_{\text{thr}}$. In our experiments we set $s_{\text{thr}} = 0.85$.

Corresponding to the score function s_{ijk} in equation 3.9, we define a maximum score function, such that

$$s_{ijk}^{\max}(t) = \begin{cases} s_{ijk}(t_p) & \text{if } \exists p : (t_{1,ijk}(t_p) \leq t \leq t_{3,ijk}(t_p)) \wedge (s_{ijk}(t_p) \geq s_{\text{thr}}) \\ 0 & \text{else.} \end{cases} \quad (3.10)$$

At each time point t that is contained in a detection time window it yields the score of the detected event. Figure 3.7c gives an example.

Single-cell intercalation scores

Finally, to compute an intercalation score function for each cell individually, intercalation score functions from all contributing cell triples are combined by maximum pooling, such that

$$s_i(t) = \max_{j,k} s_{ijk}^{\max}(t) \quad (3.11)$$

defines the single cell intercalation score for cell i . In figure 3.7d an example to this is given.

3.4 Experiments

The presented method for the detection of cell intercalations enabled us to perform quantitative comparisons of cell intercalations and their motion statistics between wildtype (WT) and mutant (MZ*spg*) embryos in the early development of zebrafish (Bensch et al., 2013).

To better understand the conducted experiments, please reconsider the biological motivation we give in section 3.1. Especially recall figure 3.4 that illustrates the developmental stages called *doming* and *epiboly*, and deep cell layer thinning. As we mention in Bensch et al. (2013), “epiboly is the first coordinated cell movement in most vertebrates and marks the onset of gastrulation. During zebrafish epiboly, enveloping layer (EVL) and deep cells spread over the vegetal yolk mass with a concomitant thinning of the deep cell layer.” (p. 845).

We tested “a prevailing model [suggesting] that deep cell radial intercalations directed towards the EVL would drive deep cell epiboly.” (Bensch et al., 2013, p. 845) by analyzing “global deep cell migratory behavior in WT and MZ*spg* mutant embryos. MZ*spg* mutant embryos are deficient in the Pou5f1 (homolog of mammalian Oct4) transcription factor, and develop a severe delay in epiboly, while emboly proceeds similar to WT (Lunde et al., 2004; Reim and Brand, 2006; Lachnit et al., 2008).” (Bensch et al., 2013, p. 847).

To this end, “we have globally recorded 3D cell trajectories for zebrafish blastomeres between sphere and 50% epiboly stages, and [...] determined intercalation events, intercalation directionality, and migration speed for cells at specific positions within the embryo. This framework uses Voronoi diagrams to compute cell-to-cell contact areas, defines a feature-based spatiotemporal model for intercalation events and fits an anatomical coordinate system to the recorded datasets. We further investigate whether epiboly defects in MZ*spg* mutant embryos devoid of Pou5f1/Oct4 may be caused by changes in intercalation behavior.” (Bensch et al., 2013, p. 845).

The following experimental descriptions in sections 3.4.1–3.4.3 are published in Bensch et al. (2013) in the section “Material and Methods”. The results, including figures, and the biological discussion presented in sections 3.4.4 and 3.4.5 are published in Bensch et al. (2013) in the sections “Results” and “Discussion”.

3.4.1 Datasets

Zebrafish Maintenance and Image Acquisition AB/TL strain was used as WT control. For embryos devoid of Pou5f1 function maternal and zygotic *spg*^{m793} mutants (MZ*spg*^{m793}) were used. 3D time-lapse recording of global blastomere migration was performed by Song et al. (2013). All nuclei were labeled by microinjection of *nls-tomato* mRNA (50 pg) at the one-cell stage. The 3D time-lapse stacks were recorded using a LSM5 Live Duo confocal microscope (Zeiss, Jena) with a Zeiss LD LCI Plan-Apochromat 25x/0.8 objective lens. The laser wavelength used was 532 nm together with the filter BP 560–675.

Volumetric Time-lapse Datasets For the analysis 6 wildtype and 6 mutant datasets were used. The 3D stack size was $1024 \times 1024 \times 81$ voxels with voxels of size $0.525 \times 0.525 \times 1.367 \mu\text{m}^3$. The time-lapse was recorded in 120 time steps with 1.05 min intervals, *i.e.* a total duration of 126 min.

3.4.2 Image Quantification Details

All analyses presented here depend on the position of the nuclei recorded in the primary data. However, virtual cell boundaries were used to define neighborhoods of cells and provided features for the detection of cell intercalations. While the full image analysis pipeline (see figure 3.3), the intercalation model and the detection of cell intercalations have already been described in the previous sections (see

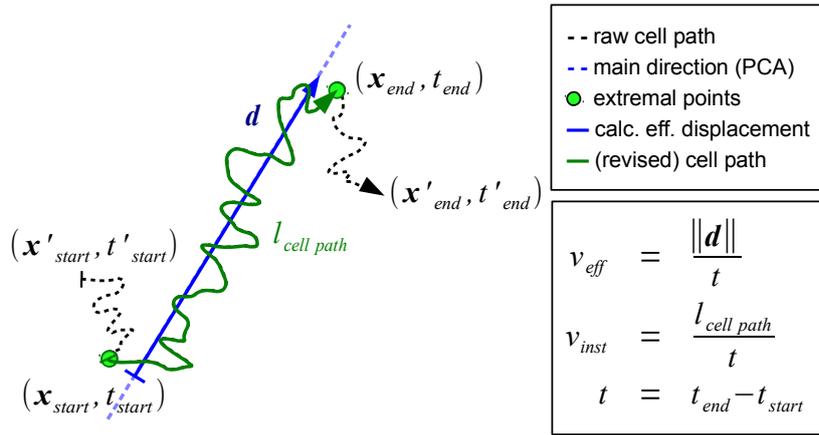


Figure 3.8: Measurement of migration direction and path length of cells undergoing intercalation events (from Bensch et al. (2013)). A schematic drawing of the raw cell path and the extracted features is shown (complementary to figure 3.13).

sections 3.2 and 3.3), the application specific method and implementation details are described in the following paragraphs.

Track Cell Nuclei

Trajectories of labeled cell nuclei were obtained by applying the track spot tool of Imaris software (Bitplane, v7.3) using the following parameters: estimated diameter: $7\mu\text{m}$, background subtraction, tracking algorithm: Brownian motion, MaxDistance: $12\mu\text{m}$, MaxGapSize: 2.

Infer Cell Boundaries and Contact Areas

Putative individual cell regions and outer boundaries (“membranes”) were estimated by a 3D Voronoi diagram using the nuclei positions as seeds. The maximal size of a cell was limited to a sphere with $20\mu\text{m}$ radius to obtain reasonable cell regions at the borders. With this approach, the size of bordering cells (especially the flat EVL cells) is overestimated in outward direction. However, inner cell boundaries are estimated well and the conducted analysis described in the following was not affected. Figures 3.1 and 3.9 and movies 3.20 and 3.21 give a good impression of the inferred cell boundaries.

Compute Relative Local Motion

The motion of intercalating cells *relative* to the local tissue was computed by successive temporal registration of local groups of cells. This compensates for translational motion from both global motion and local growth motion of the tissue. The resulting *relative* raw cell path is depicted schematically in figure 3.8 (dashed black path). The figure furthermore shows the main motion direction we computed from this raw cell path, the calculated effective displacement (abbreviated as calc. eff. displacement) and the revised cell path. The main motion direction was found by principle component analysis (PCA) and represents the best fit with respect to the raw cell path. The extremal points in this main direction were used to refine the start and end point and the time window of the intercalation event, resulting in the calculated effective displacement and the revised cell path.

Perform Event Verification

To increase the quality of intercalation detection we subsequently applied a verification step that discards events that are not likely to be intercalations. This reduces false positive detections.

Referring to the intercalation model in figure 3.5, intercalating cells are likely to show an effective displacement in the range of a cell diameter when moving from configuration T1 to T3. Therefore, we discarded detected events with less than $6\mu\text{m}$ absolute effective displacement, which is below the observed range of cell diameters. Furthermore, intercalating cells are likely to perform a rather directed motion through neighboring cells. Therefore, we discarded events that show an undirected motion. We calculate the directedness of motion as the ratio of the dominant eigenvalue to the sum of eigenvalues obtained from principle component analysis (PCA) of the revised cell path (green path in figure 3.8). In our experiments we discarded events with motion directedness below value 0.85.

Compute Basic Statistics

From the verified intercalation events, basic motion statistics were extracted. The calculated effective displacement (blue vector in figure 3.8) yielded the absolute effective displacement (supplementary material figure 3.18a) and the effective speed (figure 3.13b). The revised cell path (green path in figure 3.8) was used to quantify the cell path length (supplementary material figure 3.18b) and the average instantaneous speed (figure 3.13c).

Compute Directionality Distributions

3D directionality distributions were computed by density estimation, *i.e.* accumulating measurements into a single 3D distribution with the starting point of the event centered at the origin (figure 3.11a–f). High density peaks are obtained when many events show similar direction. To describe isotropy and polarity, the 3D directionality distributions were projected onto the unit sphere and modeled by spherical harmonics (Rose, 1995) basis functions Y_{lm} (supplementary material figure 3.16a). The resulting expansion coefficients c_{lm} with zero order ($m = 0$) in every band l (*i.e.* c_{10}) describe the distribution of the signal from North to South pole (averaged along the latitudes). For the present application, especially the second coefficient c_{20} is important (supplementary material figure 3.16a-b). It is positive, if the signal is located at the poles, negative if the signal is located at the equator, and zero, if the signal is homogeneously distributed.

Fit EVL Surface

An anatomical embryo coordinate system was defined by fitting a smooth surface to EVL cell nuclei for each time point, which is depicted in figure 3.3 (see “EVL surface”, a 2D section of the 3D surface is plotted in yellow).

Assign Intercalation Location and Direction

The fitted EVL surface was used as an anatomical reference. The direction of intercalations was obtained by measuring the angle of the calculated effective displacement (blue vector in figure 3.13a) to the surface normal. The location of intercalations was obtained by measuring the distance to this surface. For representation, the distance is discretized into cell diameters (figure 3.10d-e; supplementary material figure 3.14 and figure 3.17a,c). A reference cell diameter was used and estimated from the nearest-neighbor distances of all cells in all time steps and datasets. We used the median value $d = 16.1095\mu\text{m}$ as an estimate. Directions were classified into “lateral”, if the displacement component in the direction of the EVL was within $\pm\sqrt{3}/4 d$ ($\approx 7\mu\text{m}$) and “upward” or “downward” otherwise.

3.4.3 Statistical Analysis

Statistical significance in terms of directionality of intercalation in WT and *MZspg* (figure 3.11) was evaluated using the non-parametric Wilcoxon rank sum test (Gibbons and Chakraborti, 2011; Hollander and Wolfe, 1999). The test was based on the expansion coefficient c_{20} describing isotropy and polarity (supplementary material figure 3.16, $n = 6$ samples per class). For the description of WT and *MZspg*, measurements were averaged (figure 3.11, figure 3.12c) or summed over 6 datasets per class (figure 3.10, figure 3.12a-b, figure 3.12e-f, figure 3.13d; supplementary material figures 3.14, 3.15, 3.17). Standard MATLAB boxplots were used to plot class distributions (figure 3.13b-c; supplementary material figure 3.18b-d): the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considering outliers, and outliers are plotted individually (red crosses). The medians are significantly different at the 5% significance level if their comparison intervals (notches) do not overlap. In figure 3.10a-e, figure 3.12a-b, figure 3.12e,f, figure 3.13d, supplementary material figures 3.14 and 3.17 the error is given by the 95% confidence intervals assuming Poisson noise for counting intercalation events.

All computations and statistical evaluations were performed in MATLAB (The MathWorks Inc.) and C++.

3.4.4 Results

Furthermore, “quantification of radial and lateral intercalation dynamics of blastomeres reveals that radial intercalation is symmetric along the animal–vegetal axis of the embryo, which is not in line with the prevailing model of directed radial intercalation driving deep cell epiboly (Kane et al., 2005; Málaga-Trillo et al., 2009). Instead speed and migration efficiency of blastomeres appear to be crucial for the deep cell epiboly” (Bensch et al., 2013, p. 847).

Zebrafish Gastrulation is Initiated with Symmetric Radial Intercalation of Blastomeres

To investigate intercalation mechanism during zebrafish early gastrulation, we analyzed the trajectories of blastoderm cell nuclei in embryos labeled with NLS-tomato (Tomato fluorescent protein with nuclear localization signal) between sphere and 50% epiboly stage (Song et al., 2013).

We determined upward (into more exterior level), downward (into more interior level), and lateralward (intra-level) intercalation events of blastomeres (figure 3.9a-b; supplementary material movies 3.20–3.22). To obtain a quantitative understanding of cell behavior during epiboly, we analyzed the total number of intercalations in each WT embryo dataset (figure 3.10a). Surprisingly, the total number of upward and downward intercalations was in the same range, with slightly more downward intercalations. This observation does not support the prevailing model that suggests asymmetric radial intercalation of epiblast cells, *i.e.* inserting predominantly from an interior level into a more exterior level, drive DCL flattening (Kane et al., 2005). We further analyzed whether the epiboly delay phenotype of *MZspg* embryos may correlate with different intercalation behavior. The total and relative number of upward and downward intercalations was significantly lower in *MZspg* embryos than in WT, while the ratios between the upward and downward intercalations of blastomeres were balanced both in WT and *MZspg* embryos (figure 3.10a-c). However, factors other than the total number of intercalations in a specific direction may affect epiboly progression, including directional bias in subsequent intercalations of individual cells, and dynamic aspects of cell movement. We further investigated both possibilities in detail.

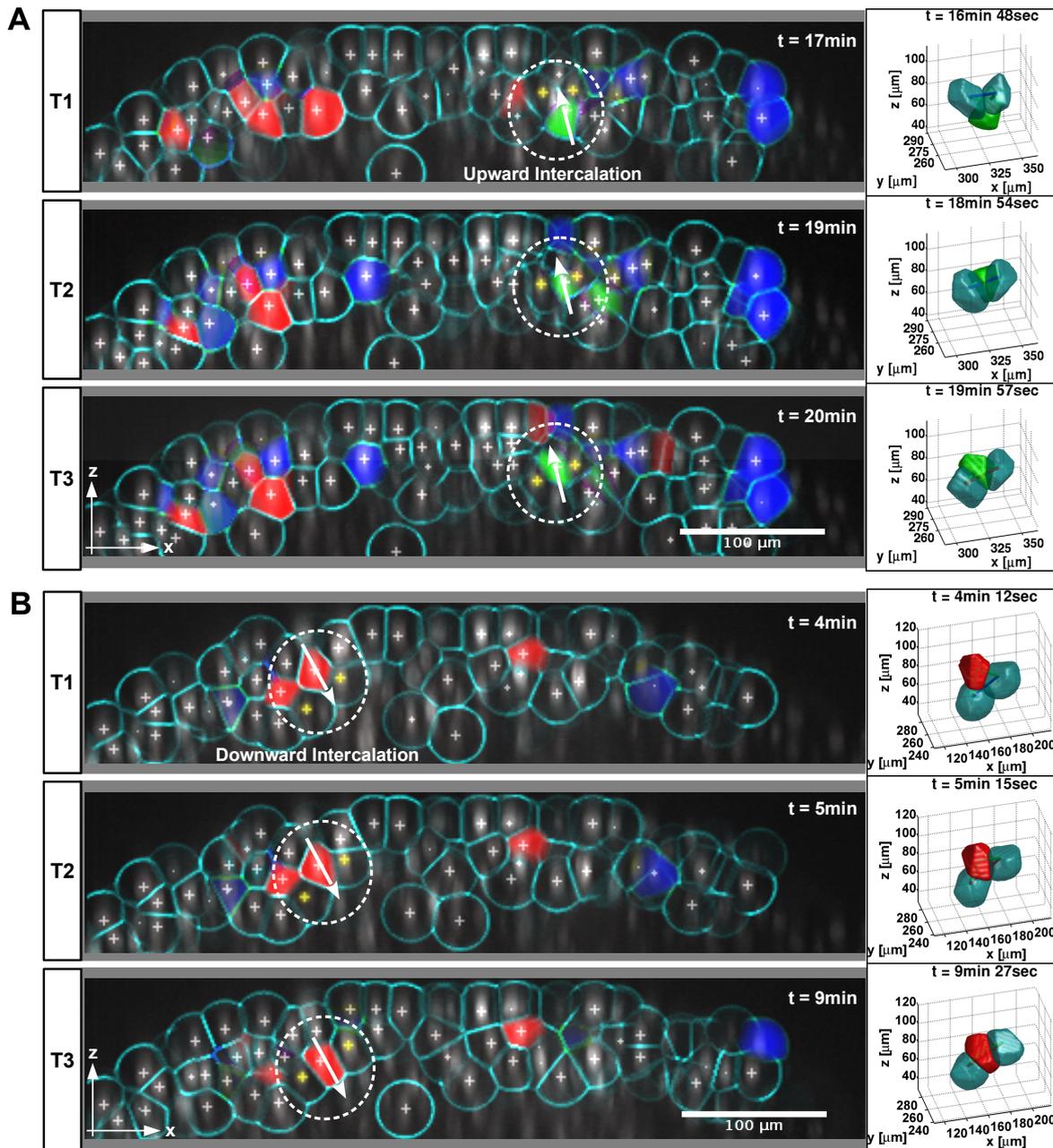


Figure 3.9: Automated detection of intercalating zebrafish blastomeres. **(a-b)** Computational detection and classification of radial intercalations from 3D time-lapse recording (supplementary material movies 3.20, 3.21). Embryo stages: sphere to 50% epiboly. The rendering shows lateral views (animal pole at top) with raw nuclei fluorescence (*grey*), tracked nuclei positions (*crosses*) and calculated cell boundaries (*cyan*). *Arrows* indicate direction of cell migration. Upward (*green*), downward (*red*), and lateralward intercalations (*blue*) were detected along an $18\mu\text{m}$ thick animal–vegetal oriented sheet transecting the embryo along its dorsoventral axis (shown here as *y*-projection representing $18\mu\text{m}$ orthogonal to the *z*-stack). In the *circled areas*, a blastomere intercalates between two neighboring cells (*yellow crosses*) located in adjacent more exterior level (a) or in adjacent more interior level (b). These two groups of cells were separately rendered in 3D (*right*). Scale bars: $100\mu\text{m}$.

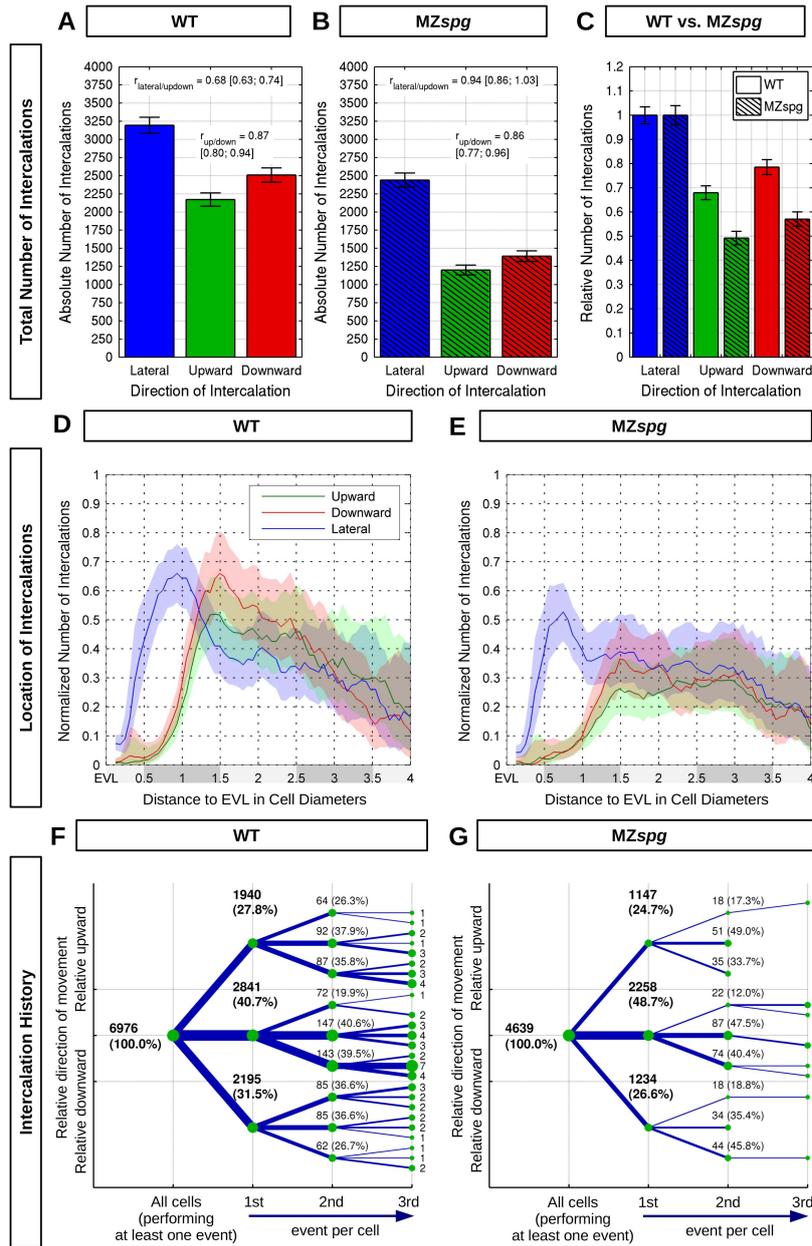


Figure 3.10: Quantification of radial and lateral intercalation events. **(a-b)** Absolute number of lateralward, upward, and downward intercalations in WT and MZspg (summed over 6 embryos for WT and MZspg each). Ratios between up- and downward, and between lateral and up-/downward intercalations are shown in each graph. **(c)** Relative number of upward or downward intercalations normalized to the number of lateralward intercalations. **(d-e)** Quantification of WT (d) and MZspg (e) blastomeres performing upward, downward, or lateralward intercalations in each depth level. Depth levels (*shaded grey* along x-axis) were numbered and distance was measured starting from the EVL in vegetal direction. To be able to compare different depth levels, the absolute number of intercalations (summed over 6 embryos for WT and MZspg each; supplementary material figure 3.14) was normalized by the total number of cells observed for each distance. The x-axis is truncated at 4.0, where the number of measured intercalations starts becoming too small to provide meaningful results. **(f-g)** Summarized intercalation history of all individual cells (sum over six embryos for each genotype). The graph presents up to three successive intercalations of individual blastomeres, indicating upward, downward, or lateralward directions. The root node (*leftmost*) denotes all cells performing the first intercalation event. The absolute number and relative fraction of intercalations is given at each node. Errors are given by 95% confidence intervals assuming Poisson noise.

Location of Intercalations and Intercalation History

We next examined the number of intercalations along depth levels. Upward and downward intercalations were mainly distributed between the first and third DCL level in both WT and *MZspg* embryos (figure 3.10d-e; supplementary material figure 3.14). However, in *MZspg* embryos, the distribution of intercalations was shifted towards deeper levels compared to WT. This may be mainly due to significantly reduced thinning of the DCL in *MZspg* embryos during the two-hour time window (supplementary material movie 3.20). Therefore, in WT, intercalations can only be detected in the first three levels at the end of the two-hour recording time, while in *MZspg* embryos the deep cell layer is still thicker and intercalations can be detected in all levels. Furthermore, the first DCL depth level shows the highest number of lateralward intercalations both in WT and *MZspg* embryos. This is reminiscent of previous reports that cells in the first DCL level are connected to EVL by E-cad mediated adherens junctions, suggesting them to be dragged by EVL during epiboly (Shimizu et al., 2005). However, cells in the first DCL level frequently moved back into the deeper levels both in WT and *MZspg* embryos during the two-hour observations, suggesting that adherens junctions were dissociated.

We next measured the number of blastomeres performing subsequent intercalation events towards the three different directions (up-, down-, and lateralward; figure 3.10f-g). During the first intercalation event WT blastomeres performed 28% upward, 41% lateralward, and 31% downward intercalations (figure 3.10f). In *MZspg* embryos 25% upward, 49% lateralward, and 26% downward intercalations were measured (figure 3.10g). Following their first intercalation many blastomeres performed a second and third intercalation in any of those three directions. These data clearly indicate that most blastomeres have the ability to perform subsequent intercalations in any direction with similar directionality distribution as in previous events. Supplementary material figure 3.15 shows the intercalation history of cells grouped by the depth level position of their first intercalation event, confirming that there is no directional bias. Similar observations were made for *MZspg* mutant embryos. In summary, cells have no directional intercalation bias based on previous intercalation events, which we interpret to exclude that some extrinsic signal may initiate irreversible cell-intrinsic processes that would determine directionality.

Quantification of Intercalation Directionality

To investigate differences in global radial intercalation rates, we compared average motion directionality and cell speed during intercalation events between WT and *MZspg* embryos. Blastomeres intercalate laterally in a rotationally symmetric distribution around the animal–vegetal axis, both in WT and *MZspg* embryos (figure 3.11). In contrast, radial intercalation directions are distributed polarized in animal as well as vegetal direction along the animal–vegetal axis both in WT and *MZspg* embryos (figure 3.11b-c,g and figure 3.11e-f,h). Strikingly, 3D directionality distributions indicate a stronger polarization in animal–vegetal direction in WT embryos than in *MZspg* embryos, represented by the significantly higher relative number of intercalations toward the animal and vegetal pole of the WT embryos (North Pole (NP): $0.23 \pm \begin{matrix} 0.06 \\ 0.04 \end{matrix}$; South Pole (SP): $0.22 \pm \begin{matrix} 0.09 \\ 0.08 \end{matrix}$) than in *MZspg* (NP: $0.14 \pm \begin{matrix} 0.06 \\ 0.05 \end{matrix}$; SP: $0.17 \pm \begin{matrix} 0.07 \\ 0.03 \end{matrix}$) (figure 3.11g-h). Accordingly, the ratio between lateralward intercalations and up- or down- intercalations is higher in *MZspg* embryos (0.94) than in WT (0.68) (figure 3.10a-c). These data reveal that in *MZspg* mutant embryos the directionality of total intercalation events may be affected by modulation of motility of blastomeres (Song et al., 2013), but may also be caused by a delay in the onset of epiboly.

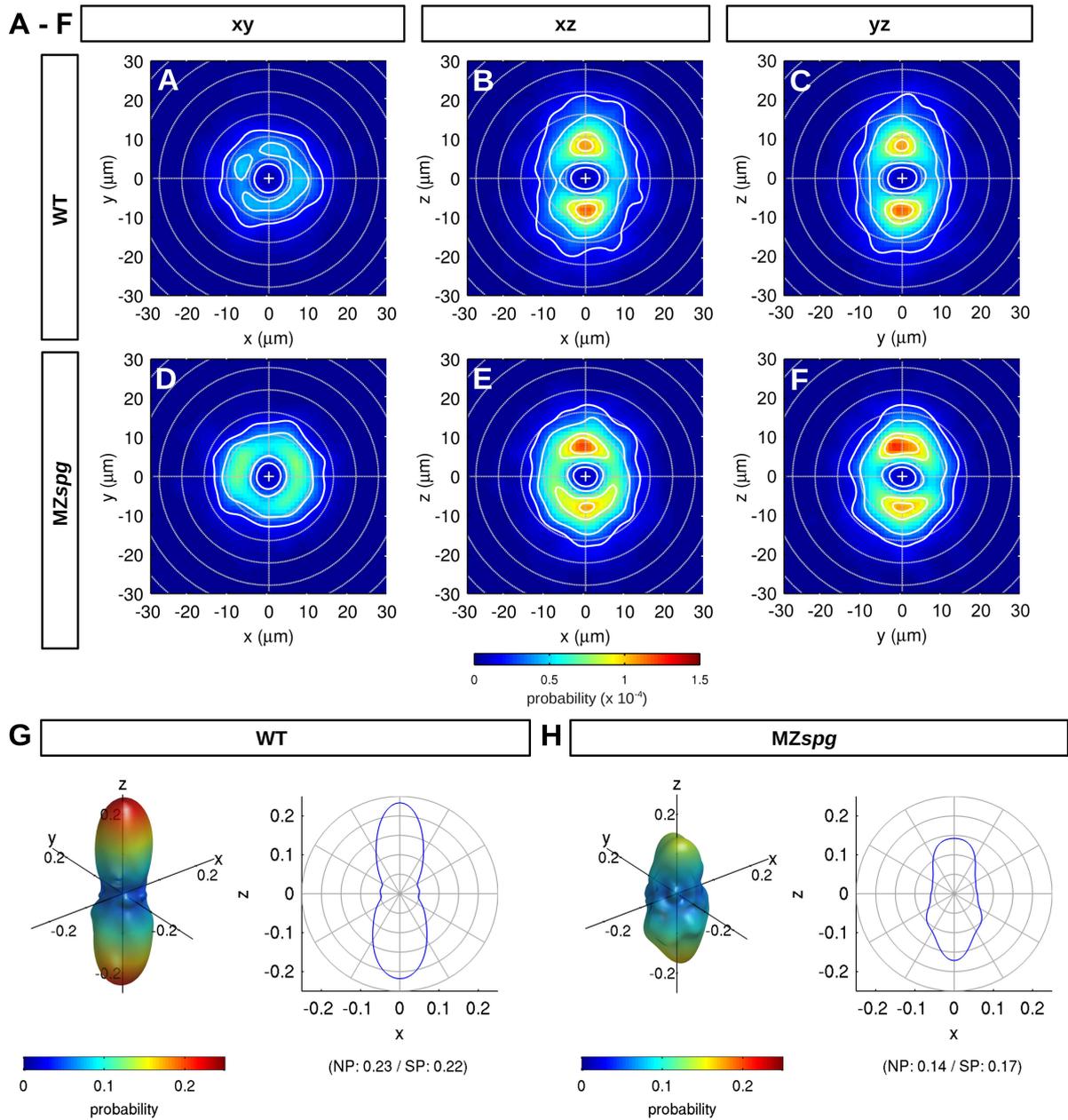


Figure 3.11: Motion directionality of intercalation events. **(a–f)** Average motion directionality analyzed for WT (a–c) and *MZspg* (d–f) embryos. The occurrence probability for an intercalation with certain migration direction and displacement is indicated by *color*. Isocontours (*white*) denote lines of equal probability. Cross-sections of 3D directionality distributions are given: x–y plane (a,d), perpendicular to animal–vegetal axis, x–z plane (b,e) and y–z plane (c,f), perpendicular to left–right and dorsoventral axis, respectively. **(g–h)** 3D reconstruction of the average motion directionality for WT (g) and *MZspg* (h) embryos (modeled by spherical harmonics of degree $l = 0 \dots 10$). *Left*: 3D rendering visualizing occurrence probability of intercalation directions using both *color* (*blue*=low and *red*=high probability) and *shape* (extension in each direction corresponds to probability). *Right*: 2D plot visualizing occurrence probability of intercalation directions averaged along the latitudes. ($P < 0.01$; $n = 6$ embryos each WT and *MZspg*).

Analysis of Spatial and Temporal Patterns of Intercalation

The time window of our analysis spans from sphere to 50% epiboly stages, a period during which different forces may contribute to epiboly. Doming of the yolk may affect cells in a different manner as compared to progress of epiboly between 30% and 50% epiboly. Therefore, we reanalyzed our data in three time windows covering the first 42 minutes roughly equivalent to doming, the time from 42 to 84 minutes equivalent to early epiboly stages, and from 84 to 126 minutes equivalent to 30–50% epiboly. The precision in determining developmental stages is estimated to be in the range of ten minutes between different embryo recordings, which argued against analysis of even shorter time windows. We find that during doming, in WT there are significantly less total intercalations (figure 3.12a), but the ratio between lateral, up and downward intercalations is not much different from the later two time windows (figure 3.12b). We also compared the total number of intercalations between WT and *MZspg* in each time window, and find that while there are significantly less intercalations in *MZspg* during time windows T1 and T2, the intercalation rate is similar between both genotypes in time window T3 (supplementary material figure 3.17b). The later onset of epiboly in *MZspg* thus contributes to the differences in intercalation behavior between the genotypes.

We also analyzed the depth distribution of radial intercalations in WT in each time window (supplementary material figure 3.17a), and found that the depth profile changes slightly from the first time window, when up, down and lateral intercalations appear at similar rates at depth levels from 1.5 to 3 cells distant from EVL, to the last time window, when the normalized number of intercalations in these three directions is higher in depth layers closer to the EVL than in deeper layers.

We further analyzed whether cells in the animal central portion of the blastoderm at the animal pole may behave differently from those located more towards the vegetal margin. We defined an inner sector S1 representing the central animal cells, and an outer sector S2 representing the more vegetally located and marginal cells (figure 3.12d). Given that S1 contained less cells than S2, we normalized the number of intercalations in each sector to the cell number. The inner sector S1 has higher up- and downward intercalation rates, with the downward intercalations nearly as strong as the lateral intercalations (figure 3.12e-f). In sector S2 the number of lateral intercalations significantly exceeds the downward intercalations. When analyzing the depth distribution of intercalation rates, inner sector S1 cells have a similar distribution of intercalation directions from layers 1.5 through 3.5, while in sector S2 the normalized number of intercalations per cell decreases in deeper layers (supplementary material figure 3.17c).

Quantification of Radial Intercalation Dynamics

Given the importance of effective movement for intercalations, we investigated the influence of loss of *Pou5f1* activity on the dynamics of cell behavior during radial intercalation by measuring effective speed and average instantaneous speed of blastomeres undergoing intercalations during early gastrulation (figure 3.13a). Both median effective and median average instantaneous speed was significantly higher in WT embryos than in *MZspg* embryos (figure 3.13b-c), suggesting *Pou5f1*-dependent mechanisms are important for control of migration speed of intercalating blastomeres. Supplemental to the analysis of migration speed during intercalation, measured absolute effective displacements and cell path lengths are given in supplementary material figure 3.18. We also determined the total number of intercalations, which was significantly higher in WT embryos than in *MZspg* (figure 3.13d). These data together indicate that *Pou5f1* affects the total number of intercalations of blastomeres by controlling cell motility, especially migration speed of cells.

3 Spatiotemporal Geometrical Models for Motion Pattern Detection

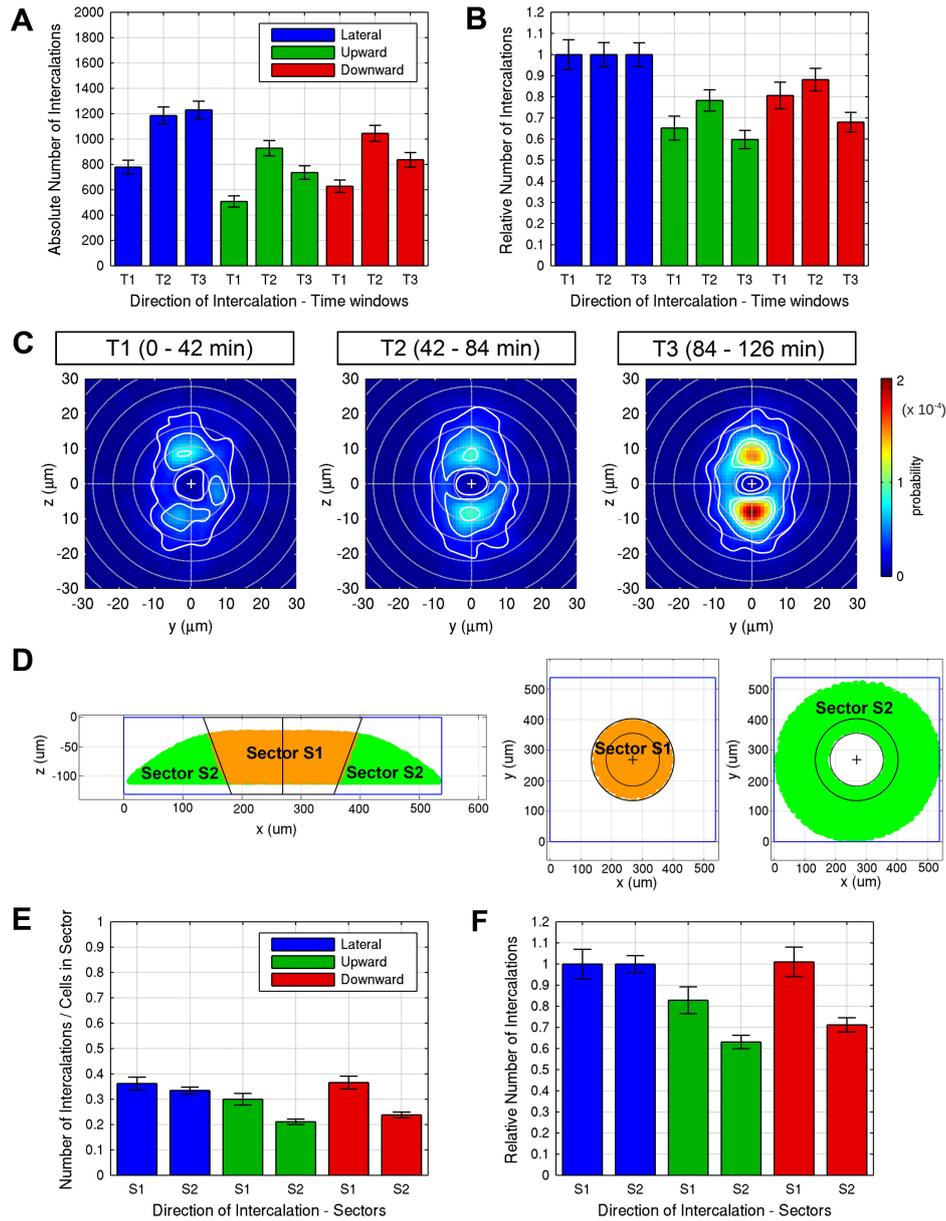


Figure 3.12: Analysis of spatial and temporal patterns of intercalation. **(a-b)** Absolute number (a) of lateralward, upward, and downward intercalations in WT for the time windows T1 (0–42 min), T2 (42–84 min), and T3 (84–126 min). The data are summed over 6 embryos each. (b) Relative number of upward or downward intercalations in each time window normalized to the number of lateralward intercalations. **(c)** Average motion directionality analyzed for WT embryos for each of the time windows T1 to T3. The occurrence probability for an intercalation with certain migration direction and displacement is indicated by *color*. Isocontours (*white*) denote lines of equal probability. Cross-sections of 3D directionality distributions are given for the *y*–*z* plane. **(d)** To analyze potential differences in intercalation behavior between an inner sector located centrally at the animal pole and an outer sector encompassing more marginal and vegetal cells, the 3D space of the image data stack was separated into an inner sector S1 (*orange*) and an outer sector S2 (*green*), visualized in lateral (*left*) and animal pole views (*right*). **(e-f)** Number (e) of lateralward, upward, and downward intercalations in WT for the sectors S1 and S2 normalized by the number of cells for each sector. The data are summed over 6 embryos each. (f) Relative number of upward or downward intercalations in each sector normalized to the number of lateralward intercalations.

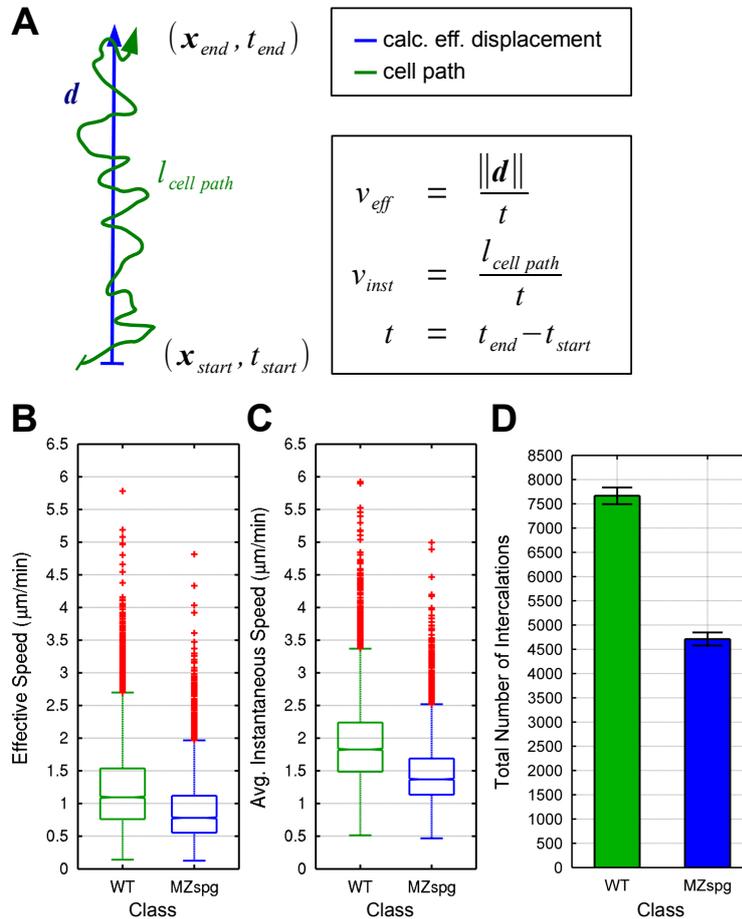


Figure 3.13: Migration speed of intercalating cells. **(a–c)** Radial intercalation dynamics. The effective and average instantaneous speeds of intercalating blastomeres were quantified. **(a)** Schematic drawing of cell path (*green*) and calculated effective displacement (*blue*) during intercalation. Calculation and comparison of effective **(b)** and average instantaneous speed **(c)** for WT and MZspg embryos ($P < 0.05$; $n = 6$ embryos each WT and MZspg; Standard MATLAB boxplots). **(d)** Quantification of the total absolute number of intercalations for WT and MZspg embryos ($P < 0.05$; summed over $n = 6$ embryos each WT and MZspg). Error bars show 95% confidence intervals assuming Poisson noise.

3.4.5 Biological Discussion

Gastrulation is an excellent model to study mechanisms controlling coordinated movements of large numbers of cells. However, even for the earliest gastrulation movement, epiboly, there is little understanding of the mechanisms that regulate this movement spatially and temporally throughout the embryo. Here, we used the zebrafish for a detailed analysis and description of intercalation cell behavior during the first two hours of zebrafish gastrulation, from sphere stage to doming of the yolk and epibolic spreading of cells up to 50% epiboly. We aimed to record most cell movements based on the position of their cell nuclei in one coherent data stack, which limited our analysis to about 50% epiboly stage, as we were not able to image throughout the embryo with a confocal laser scanning microscope at later stages. While other techniques have enabled whole embryo documentation (Keller et al., 2008) and analysis of surface movement of cells, such data have not been analyzed for cell behavior orthogonal to the surface, which is essential for analysis of radial cell intercalations. Our global cell intercalation study therefore focused on early epiboly stages, while previous analyses of small regions of the embryo investigated cell behavior at late epiboly stages from 70 to 90% epiboly (Kane et al., 2005).

We used a mathematical three point model to analyze intercalation cell behavior, which enabled us to apply image analysis algorithms to automatically detect and characterize cell intercalations throughout the 3D data volume and the two hour time-lapse recording. The results showed that upward, downward and lateral intercalations occur throughout the deep cell layers, and surprisingly revealed similar rates of up- and downward intercalations with regard to the EVL surface, which argues against intercalations directed towards the EVL to be the major force to shape spreading and thinning of deep cells during epiboly. Observing individual cells also revealed no long-term bias in intercalation directionality: following a first intercalation, cells that performed a second intercalation did not show any bias in up-, down-, or lateralward direction. Thus, it appears that cells during early epiboly do not appear to become intrinsically programmed to intercalate in a defined direction only.

We also investigated changes in cell behavior in three time windows for doming, early epiboly and 30–50% epiboly. We found that in WT embryos, the intercalation directionality is not very prominent during doming, but a clear animal–vegetal directional bias is established during early epiboly, with similar upward versus downward intercalation distribution along this axis. We also found that the depth distribution of intercalations changes as epiboly progresses: while during dome stage up-, downward and lateral intercalations appear at similar low frequencies in layers one to three cell diameters away from the EVL, during mid-epiboly a profile is established in which the frequency of intercalating cells is higher in the upper cell layers compared to deeper ones.

The cell intercalation data raise the question whether they are sufficient to explain blastoderm thinning and epibolic spreading towards the vegetal pole. First, it appears counter-intuitive that a high number of both up- and downward intercalations has been detected. Because, if they occur in the same layers, this would effectively eliminate any net effect on expanding the DCL. However, the depth profile of intercalation events normalized to cell numbers in each depth layer reveals that from dome stage, when the profile is even, a gradient of intercalation rates establishes with higher intercalation rates in deep cell layer one to two, as compared to layers three to four. Together with the slightly higher propensity for downward intercalations, this may effect a net redistribution of cells by intercalation to promote thinning and spreading of the DCL. We attempted to evaluate quantitatively the number of cells exiting the inner sector S1 in comparison to the number of radial intercalations (supplementary material figure 3.19). To visualize temporal changes we performed the analysis in eight 15 minute time windows. We also determined the ratio of the number of cells leaving the sector S1 and the number of radial intercalation events. Supplementary material figure 3.19 reveals that the number of cells leaving sector S1 trails behind the number of radial intercalations, until shortly before 50% epiboly, when exiting cells and radial intercalations occur at a ratio of approximately one. This analysis would be consistent with up- and downward intercalations partially compensating each other during doming and early epiboly,

while intercalations may drive epiboly more effectively when the directionality (figure 3.12c) and the steeper radial profile of intercalations (supplementary material figure 3.17a) are established at 50% epiboly. However, it has been impossible for us to exactly quantitate the contribution of intercalation events to epibolic spreading. The reason is that in the time window analyzed, also the number of cells approximately doubles as the asynchronous thirteenth cell cycle progresses, and towards the end of the recording cells also have only about half of the volume each as compared to sphere stage.

Two models have been put forward about the forces that drive DCL epiboly (Keller et al., 2003). In one model, radial intercalation of deep cells is the driving force to spread the DCL. Here, directional cues would have to orient the intercalation behavior. Adhesion gradients, specifically of E-cad, have been proposed to direct radial intercalation during late epiboly to predominantly occur in the direction towards the EVL (Kane et al., 2005). Our analysis reveals that such a directional intercalation cannot be detected during early epiboly stages. In the second model (Keller et al., 2003), radial intercalation may be a more indirect effect of migrational spreading over yolk cell or EVL surfaces. In this model, for zebrafish epiboly of the YSL and EVL would open a space into which deep cells migrate. The prevalence of intercalation orthogonal to the EVL surface observed here may be caused by this type of intercalation effectively filling space opened up by EVL/YSL epiboly. Here, the dynamics and effectiveness of deep cell migration would be crucial for DCL epiboly progress, which is confirmed by our measurements. This is also consistent with the changes observed in *MZspg* mutant embryos, in which E-cad trafficking and adhesion is affected in a way to reduce effective cell movements (Song et al., 2013).

3.5 Conclusion and Outlook

We have presented a new approach to motion pattern detection in 4D biological data. First, an explicit model of the target event is defined. It consists of piecewise linear functions modelling feature transitions of the spatio-temporal model. Here, we presented a model to detect cell intercalations. The proposed detection method fits linear models to the measured data and thus is robust to very noisy and even interrupted data. Furthermore, the detection is invariant to 3D translation and rotation on the one hand, and temporal scaling and feature scaling on the other hand.

“Our study provides a new approach to investigate dynamic behavior and intercalations of individual cells within a tissue during embryo development. This method may also be exploited in other fields such as cancer research to quantify epithelial–mesenchymal transitions *in vivo*” (Bensch et al., 2013, p. 853). For future work we suggest to generalize the presented approach, such that it is able to learn or infer a model from given examples in an unsupervised manner, in contrast to an a priori defined explicit model.

3.6 Supplementary Material

This section contains figures and movies supplemental to the results in section 3.4.4. They are published in Bensch et al. (2013) in the “Supplementary Material”.

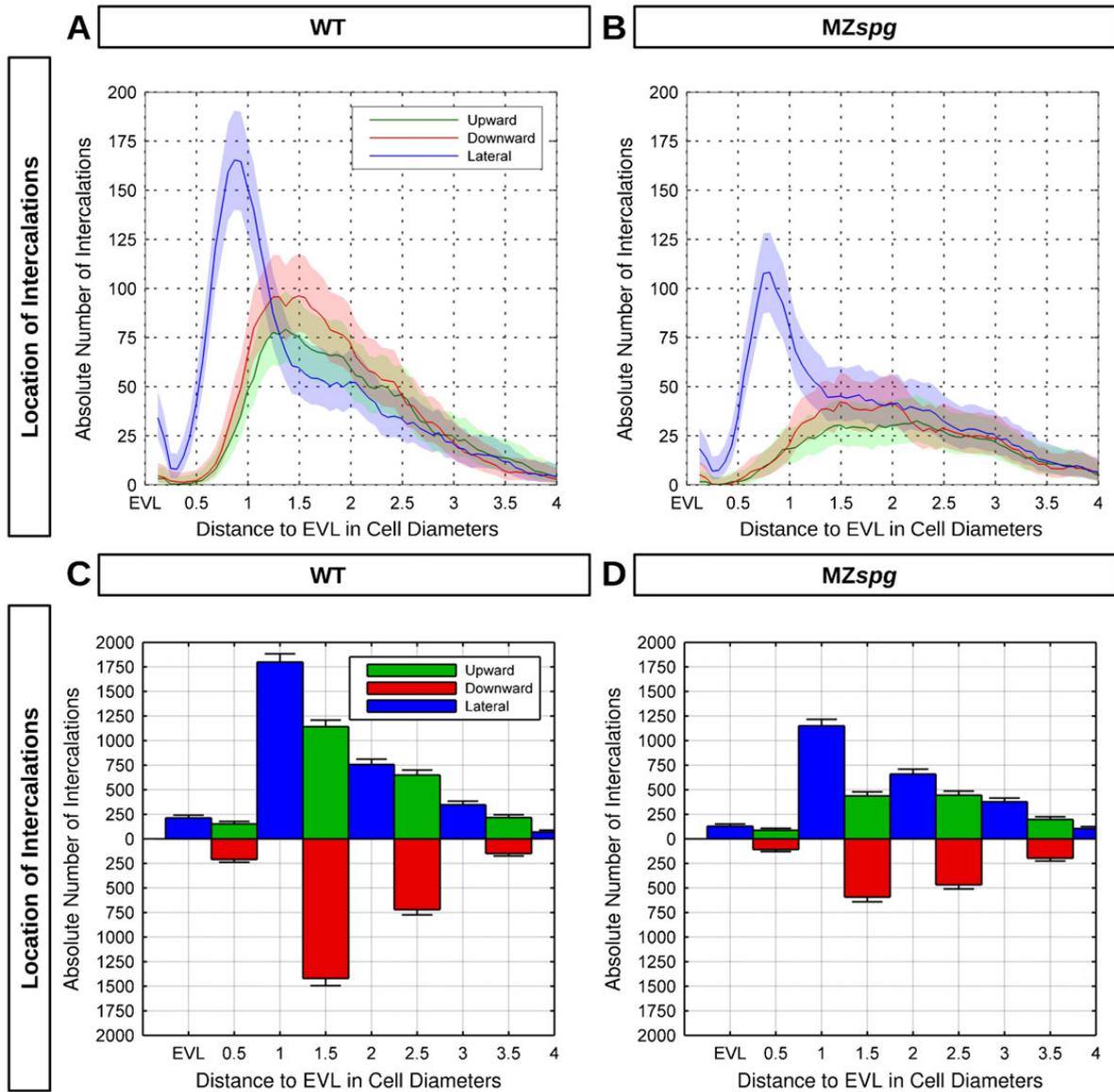


Figure 3.14: Quantification of location of radial and lateral intercalation events. Quantification of WT (a,c) and MZspg (b,d) blastomeres performing upward (green), downward (red), or lateralward (blue) intercalations in each depth level. Depth levels were numbered and distance was measured starting from the EVL at the animal pole towards the vegetal pole of the embryos. Compared to figure 3.10d-e the absolute (non-normalized) number of intercalations is shown (summed over all datasets). Besides the continuous representation (a-b) measurements are given discretized into bins of one cell diameter size (c-d) centered on integer distances to EVL for lateralward and in between for radial intercalations ($n = 6$ embryos each for WT and MZspg). The x-axis is truncated at 4.0, where the number of measured intercalations becomes too small to provide meaningful results. Errors are given by the 95% confidence intervals assuming Poisson noise.

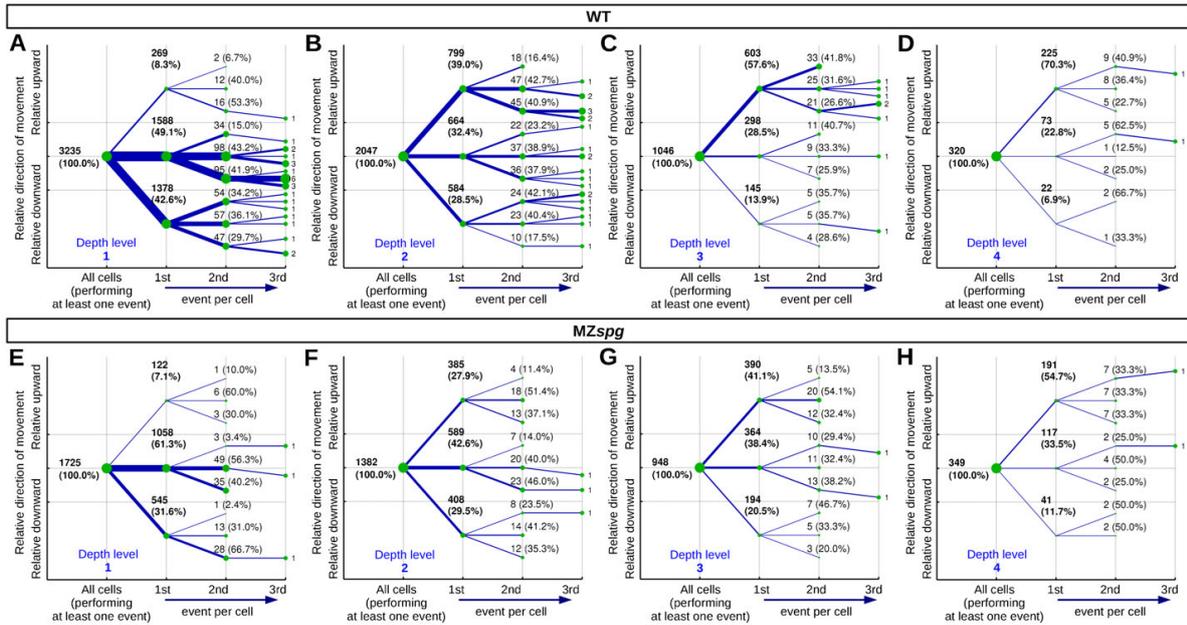


Figure 3.15: Quantification of intercalation history of radial and lateral intercalation events. Summarized intercalation history of all individual cells for WT (a–d) and *MZspg* (e–h) embryos (sum over six embryos for each genotype). The graph presents up to three successive intercalations of individual blastomeres, indicating upward, downward, or lateralward directions. The root node (*leftmost*) denotes all cells performing the first intercalation event. Compared to figure 3.10g the intercalation history is split according to the depth level where the first intercalation event is localized. From *left* to *right* the depth level increases by one. The absolute number and relative fraction of intercalations is given at each node.

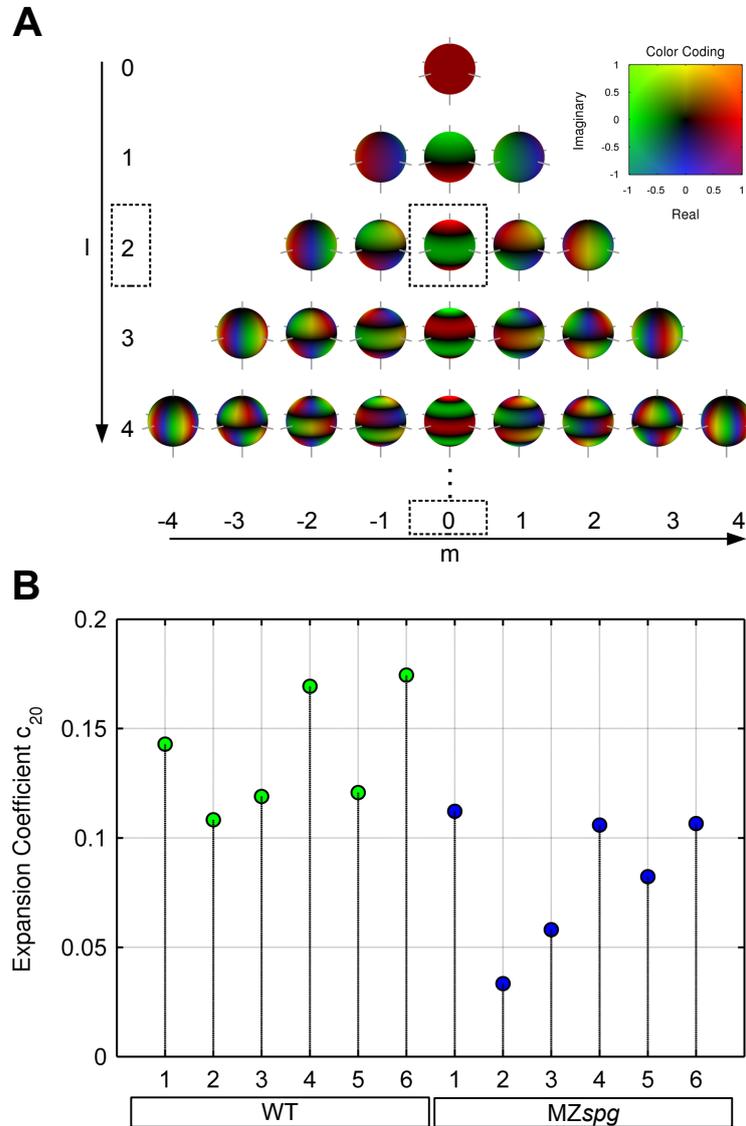


Figure 3.16: Motion directionality of intercalation events: spherical harmonics and expansion coefficients c_{20} . **(a)** Spherical harmonics basis functions Y_{lm} , plotted for degree $l = 0 \dots 4$ and order $m = 0 \dots 4$. For the present application, especially Y_{20} (degree 2, order 0) is important. It represents signals that are polarized at the North and South pole. **(b)** Expansion coefficients c_{20} for all WT and MZspg datasets representing the contribution of spherical harmonics basis function Y_{20} to their directionality distributions. Larger positive values represent stronger polarization at the North and South pole, smaller positive values represent weaker polarization respectively. WT shows significantly increased polarization compared to MZspg embryos ($P < 0.01$; $n = 6$ embryos each for WT and MZspg), see also figure 3.11.

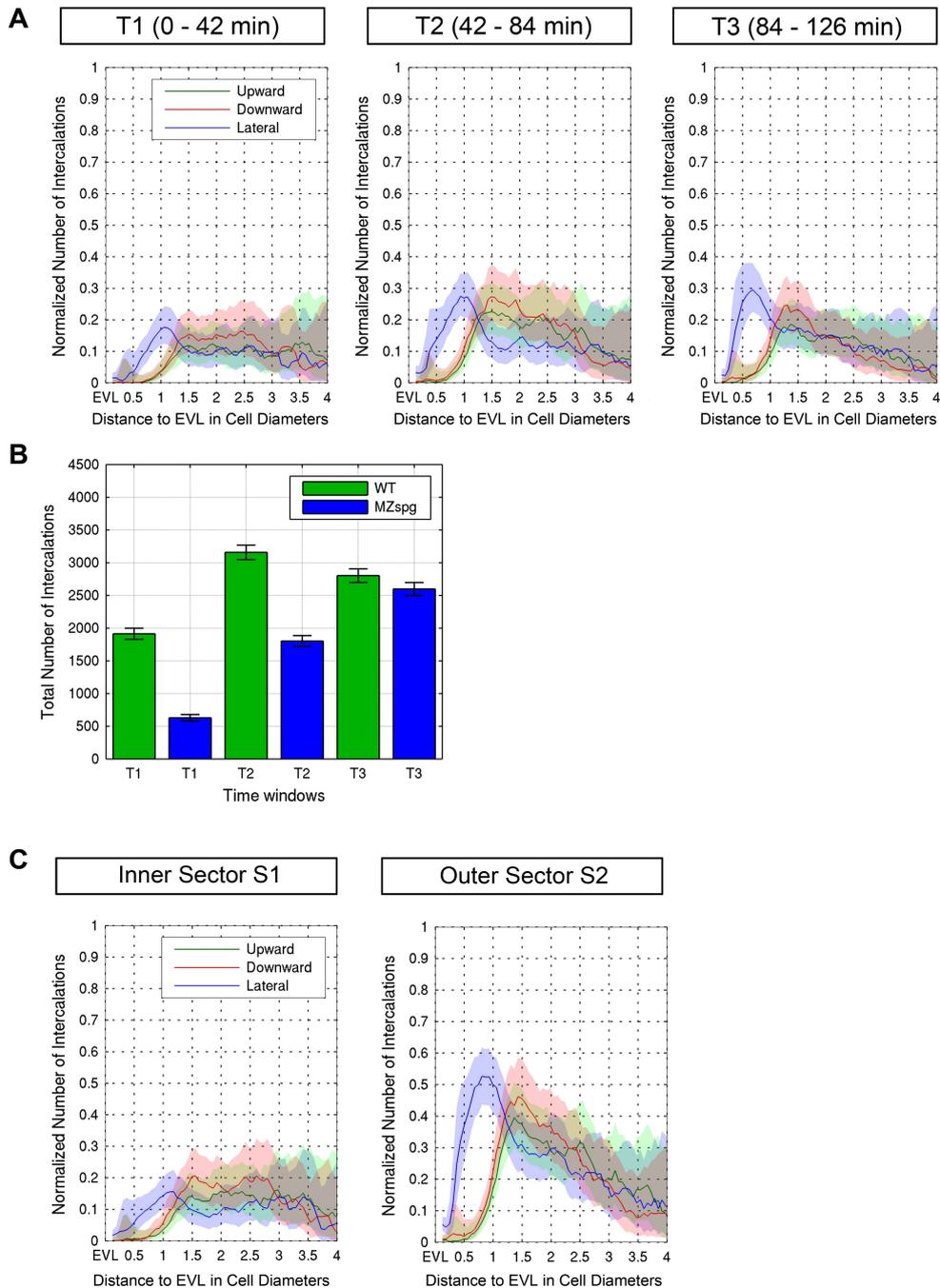


Figure 3.17: Analysis of spatial and temporal patterns of intercalation in wildtype and *MZspg*. **(a)** Quantification of WT blastomeres performing upward, downward, or lateralward intercalations in each depth level for the time windows T1 to T3. Depth levels as distance in average cell diameters were numbered and distance was measured starting from the EVL in vegetal direction. To be able to compare different depth levels, the absolute number of intercalations (summed over 6 embryos) was normalized by the total number of cells observed for each distance. The x-axis is truncated at 4.0, where the number of measured intercalations becomes too small to provide meaningful results. **(b)** Quantification of the total absolute number of intercalations for WT and *MZspg* embryos in each of the three time windows T1 to T3 ($P < 0.05$; summed over $n = 6$ embryos each WT and *MZspg*). Error bars show 95% confidence intervals assuming Poisson noise. **(c)** Quantification of WT blastomeres performing upward, downward, or lateralward intercalations in each depth level for the inner (S1) and outer (S2) sectors defined in figure 3.12. For details see panel A.

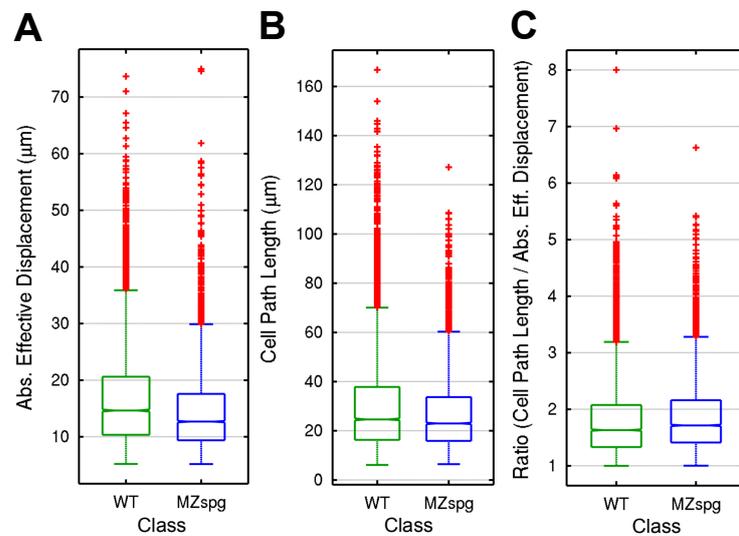


Figure 3.18: Measurement of migration direction and path length of cells undergoing intercalation events. **(a-c)** Radial intercalation dynamics (complementary to figure 3.13). **(a-b)** Complementary to figure 3.13b-c, the measured absolute effective displacements and cell path lengths are shown. **(c)** The ratios of cell path lengths (c) and absolute effective displacements (a) are shown. The ratio indicates how effective or directed cellular motion is during intercalation. A value of 1.0 indicates straight linear motion, while larger numbers indicate less effective or directed motion on a non-linear path.

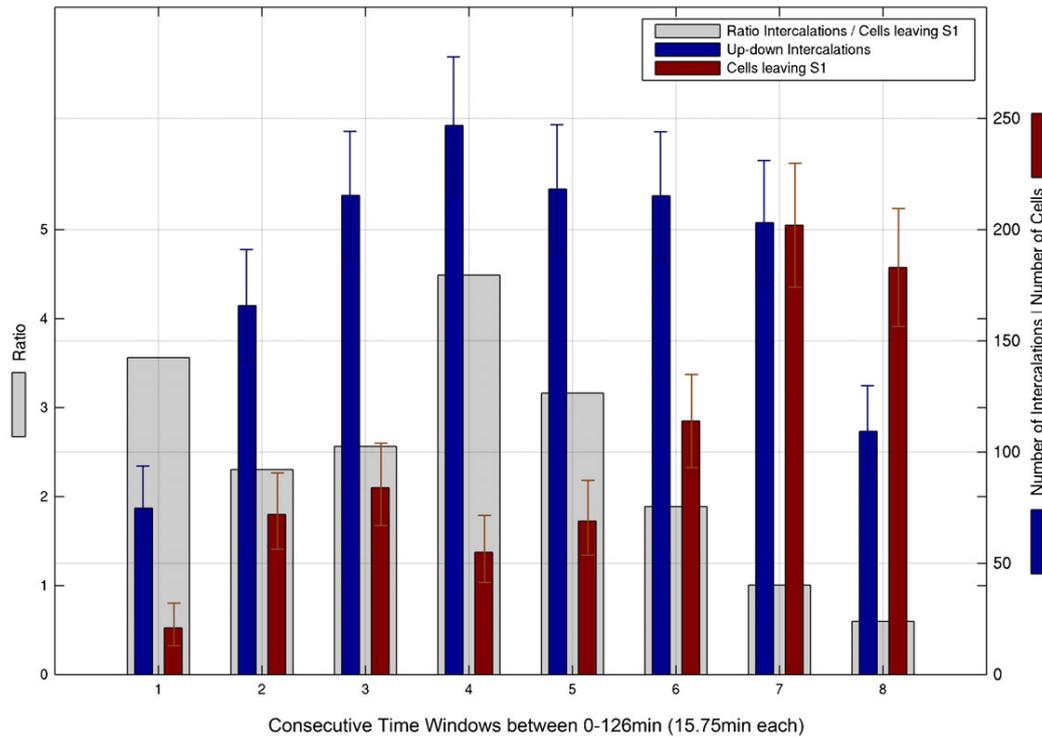
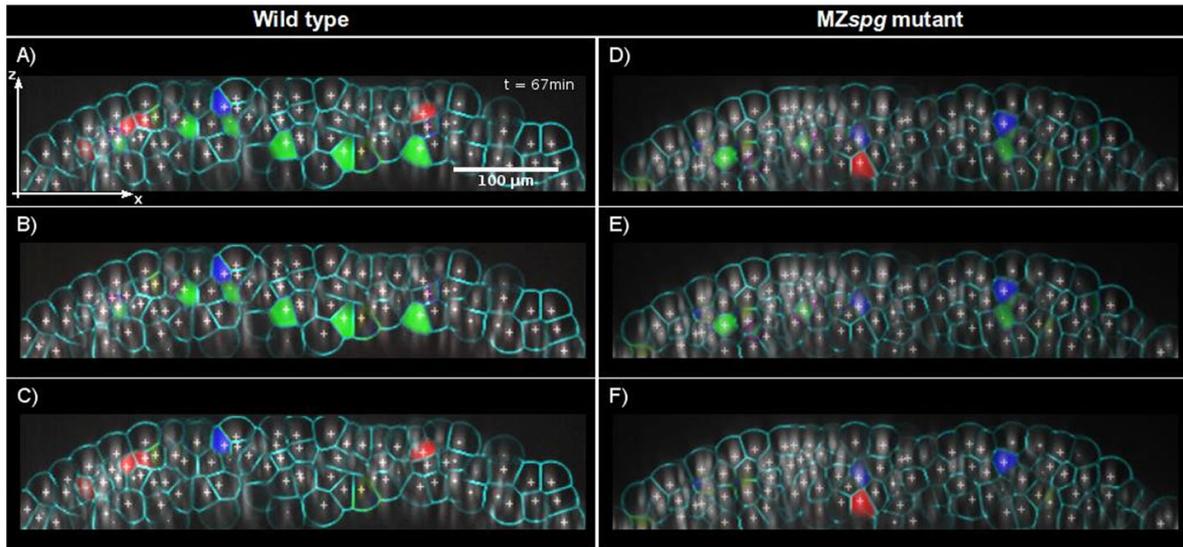
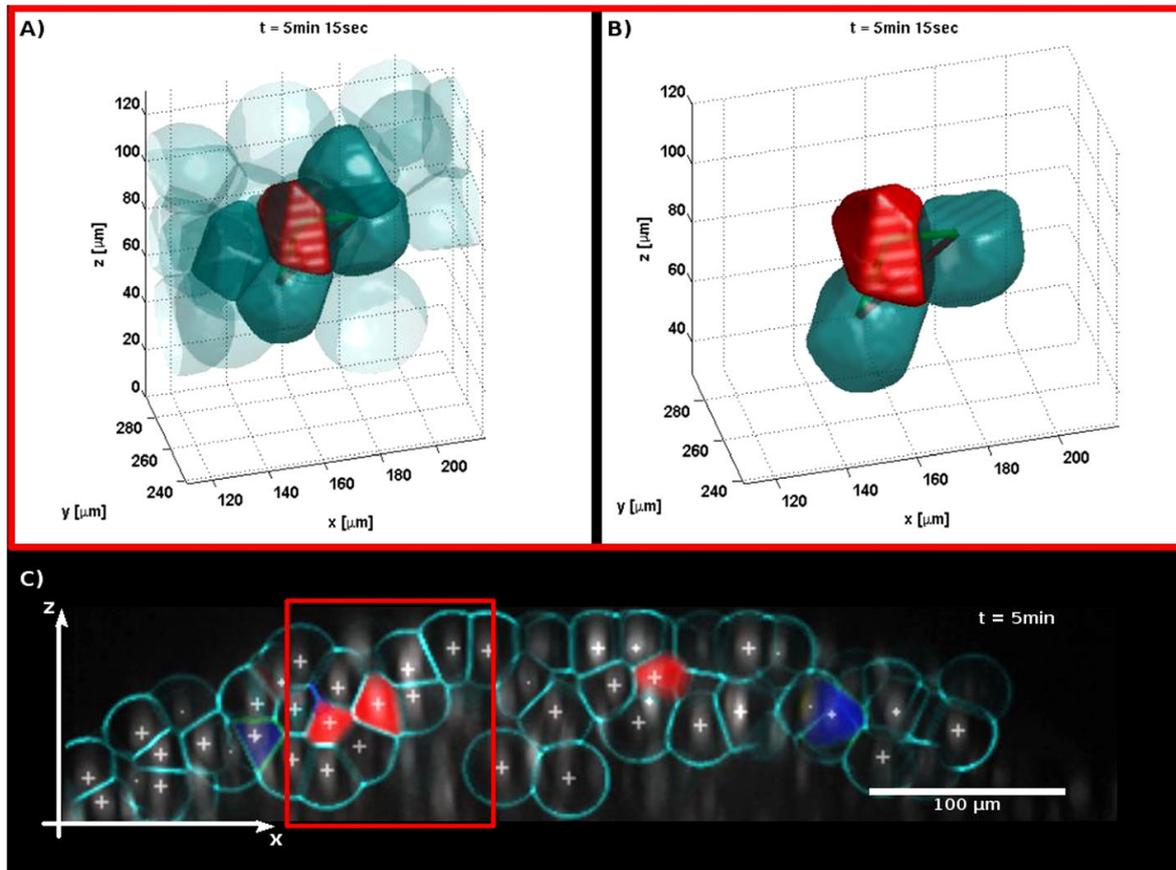


Figure 3.19: Quantification of a potential contribution of radial intercalation to epibolic spreading in WT embryos. To determine whether the number of intercalations correlates with the number of cells leaving the inner sector S1 during shorter time windows, we defined eight consecutive time windows 1 to 8 (15.75 min each during the total observation time of 126 min), and quantified for each time window the number of radial intercalations and the number of cells leaving sector S1 (summed over $n = 6$ embryos). The number of summed upward and downward intercalations in sector S1 (*blue bars*) is compared to the number of cells leaving sector S1 (*brown bars*). The ratio between both quantities is plotted in *gray bars*. Error bars show 95% confidence intervals assuming Poisson noise.

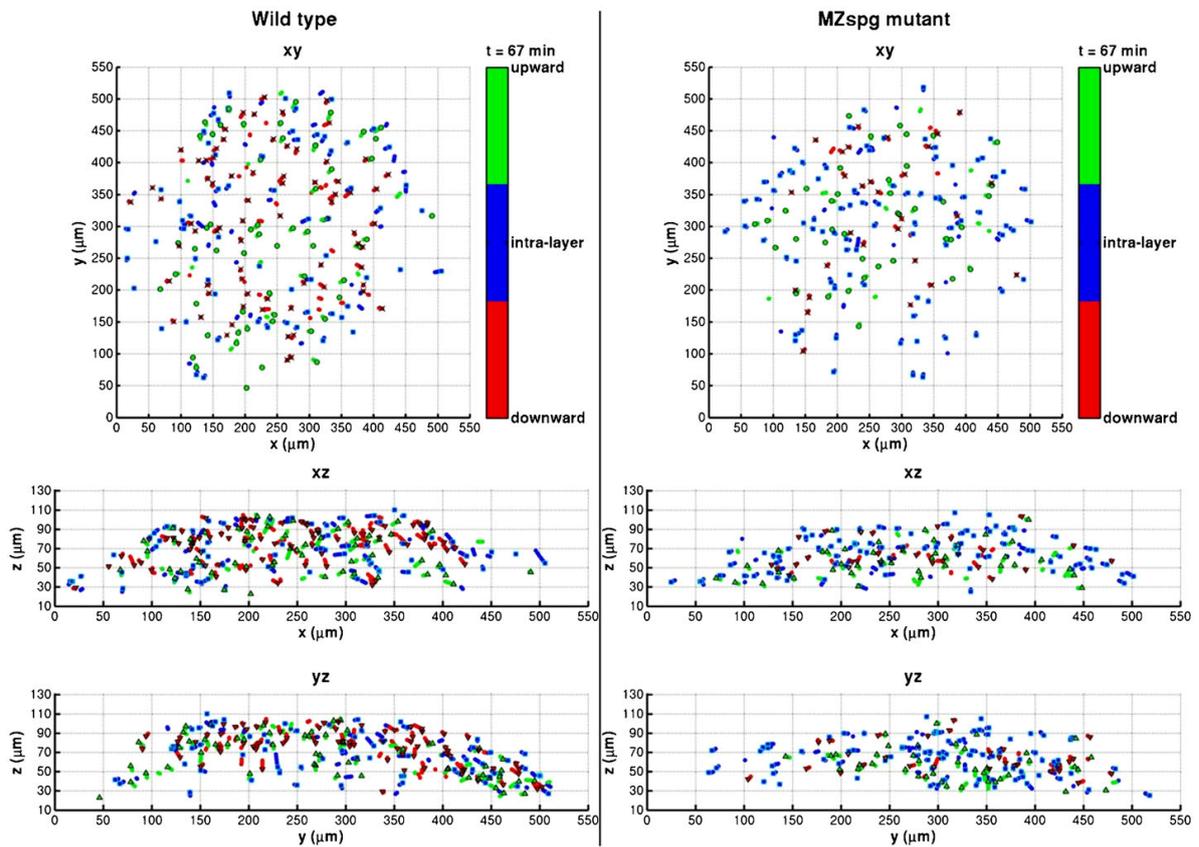


Movie 3.20: Radial intercalation events during blastomere movement. Global blastomere migration analyzed by computational detection and classification of radial intercalation events from two hour 3D time-lapse recording of WT (**a–c**) and *MZspg* (**d–f**) embryos. Embryo stages: sphere to 50% epiboly. Confocal stacks ($109\mu\text{m}$) were recorded from animal pole EVL into the margin of blastoderm over the doming yolk of the embryos. The renderings show lateral views (animal pole to the top) with raw nuclei fluorescence (*grey*), tracked nuclei positions (*crosses*) and calculated cell boundaries (*cyan*). Upward intercalations into more exterior levels (*green*), downward intercalations into more interior levels (*red*), and lateralward intra-level intercalations (*blue*) were detected along an $18\mu\text{m}$ thick animal–vegetal oriented sheet transecting the embryo along its dorsoventral axis. (**a,d**) All intercalation events. (**b,e**) Upward and lateralward intercalations. (**c,f**) Downward and lateralward intercalations. Scale bar: $100\mu\text{m}$. (Available online: <http://bio.biologists.org/content/suppl/2013/06/03/bio.20134614.DC1/bio.20134614-s2.mp4>)



Movie 3.21: Examples for intercalation events within a group of blastomeres. Rendering of individual upward and downward intercalation events from 3D timelapse recording of WT embryo (supplementary material movie 3.20a–c). **(c)** Lateral view rendering (animal pole to the top) with raw nuclei fluorescence (*grey*), tracked nuclei positions (*crosses*) and calculated cell boundaries (*cyan*). Upward intercalations into more exterior levels (*green*), downward intercalations into more interior levels (*red*), and lateralward intra level intercalations (*blue*) were detected along an $18\ \mu\text{m}$ thick animal–vegetal oriented sheet transecting the embryo along its dorsoventral axis. Downward intercalation is shown exemplarily inside the *red rectangle* during the first 10 minutes of the movie. Upward intercalation is shown exemplarily inside the *green rectangle* during the following 20 minutes of the movie. **(a–b)** 3D rendering of corresponding individual cells performing upward or downward intercalation within a group of blastomeres. Scale bar: $100\ \mu\text{m}$. (Available online: <http://bio.biologists.org/content/suppl/2013/06/03/bio.20134614.DC1/bio.20134614-s3.mp4>)

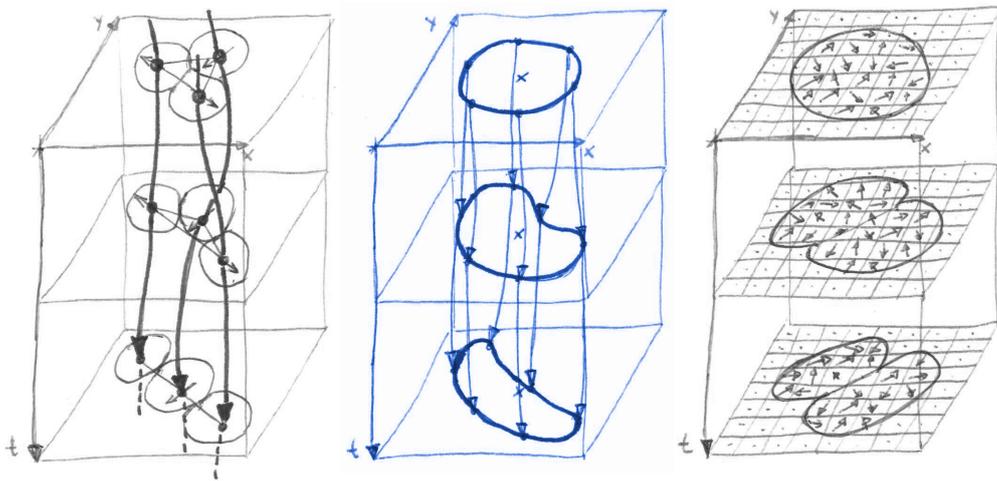
3 Spatiotemporal Geometrical Models for Motion Pattern Detection



Movie 3.22: Directional analysis of migrating blastomeres. The calculated effective displacement (figure 3.13a) of all detected intercalations from WT and MZspg embryos from two hour 3D time-lapse recording (supplementary material movie 3.20) were visualized. Colors indicate the classified direction (upward: green, downward: red, and lateralward: blue). Three different views are shown: x–y plane (animal top view of the embryo), x–z plane and y–z plane (lateral views of the embryo). Motion direction is indicated in each view additionally by different symbols (lateralward: square; top view upward: circle, downward: cross; side views up-/downward: up-/downward pointing triangle). Point traces are shown for the past 5 minutes preceding each frame only. (Available online: <http://bio.biologists.org/content/suppl/2013/06/03/bio.20134614.DC1/bio.20134614-s4.mp4>)

Part II

Contour-based Methods



4 Asymmetric Graph Cut for Cell Segmentation in Phase Contrast Images

In this chapter, we propose a new robust, effective, and surprisingly simple approach for the segmentation of cells in phase contrast microscopy images. The key feature of our algorithm is that it strongly favors dark-to-bright transitions at the boundaries of the (arbitrarily shaped) segmentation mask. The segmentation mask can be effectively found by a fast min-cut approach. The small but essential difference to standard min-cut based approaches is that our graph contains directed edges with asymmetric edge weights. Combined with a simple region propagation our approach yields better segmentation results than the top ranked methods on the PhC-C2DH-U373 phase contrast dataset in the ISBI Cell Tracking Challenge 2014 and was the 2nd ranked method in the ISBI Cell Tracking Challenge 2015 winning against the previously 1st ranked method. We provide an open-source implementation for Matlab on our homepage.

The method has been published in an earlier version at the International Symposium on Biomedical Imaging (Bensch and Ronneberger, 2015). An oral presentation was given at the conference in New York 2015 where it was selected among the 18 finalists of the best student paper award¹. Our results achieved at the ISBI Cell Tracking Challenge 2015 were published in Nature Methods (Ulman et al., 2017). The results were also published on the challenge website in the category “Latest Results”² (our method acronym is: FR-Be-GE). The cell tracking challenge has been continued by the organizers and is open for online submissions since 2017. Up-to-date results are evaluated and published on the new challenge website (ISB).

4.1 Introduction

Since its invention around 75 years ago, phase contrast microscopy (Zernike, 1942) has become the premier choice to visualize thin transparent regions in living cells (figure 4.1a). The advantageous high contrast at the cell borders comes with several artifacts, like shade-off and halo patterns (Murphy, 2001), see figure 4.2, which complicate an automated segmentation. The shade-off effect increases

¹<http://biomedicalimaging.org/2015/student-paper-competition>

²http://www.codesolorzano.com/Challenges/CTC/Latest_Results.html (as of July 4th, 2018)

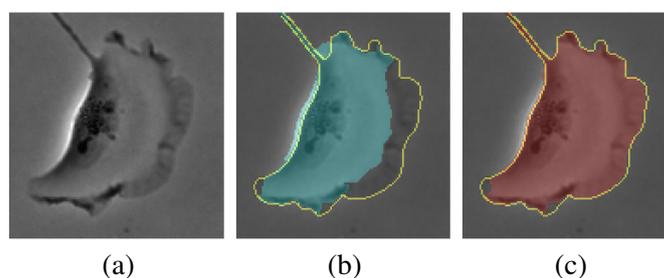


Figure 4.1: Cell segmentation for phase contrast images. **(a)** Raw image. **(b)** Segmentation result with traditional graph cut (*cyan*), and ground truth contour (*yellow*). **(c)** Segmentation result with the proposed method (*red*).

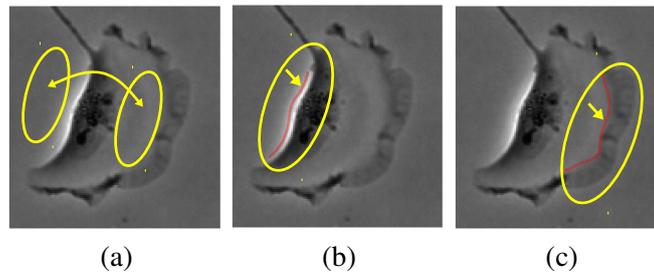


Figure 4.2: Artifacts in phase contrast microscopy. **(a)** Shade-off **(b)** Halo pattern **(c)** Strong edges inside and outside the cell. Artifacts are indicated in *yellow* and *red*.

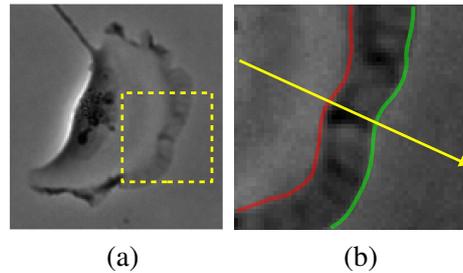


Figure 4.3: Cell segmentation approach: True cell borders in positive phase contrast microscopy appear as a dark-to-bright transition in outwards direction. **(a)** Raw cell image with difficult to segment cell border highlighted by region in *yellow*. **(b)** Closeup of region shown in (a). Cell outwards direction in *yellow*. True cell border in *green* and wrong cell border in *red*.

the inner cell intensity to the same level as the surrounding medium, which hamper the application of simple region-based approaches. Additionally both effects can introduce strong edges inside and outside of the cell, which then guide standard edge-based algorithms (see figure 4.1b) to the wrong positions.

In this work we propose a principled solution that makes use of the fact that the true cell borders in positive phase contrast microscopy always appear as a dark-to-bright transition in outwards direction, see figure 4.3. *I.e.* all borders with an inverse transition (bright-to-dark) are definitely not the sought cell borders. For simple morphologies, like roundish or star-shaped cells, the wrong borders could be easily suppressed in a pre-processing step. However, for more complicated morphologies (see figure 4.9) the outwards-direction depends on the local border-normal of the resulting segmentation mask, which is not available in advance. We solve this problem by minimizing an energy functional that searches for a segmentation mask and simultaneously favors dark-to-bright transitions at its boundary. Discretization of this functional yields a combinatorial optimization problem that can be solved efficiently by a min-cut approach (see figure 4.1c). The important difference to the usual application of min-cuts in image segmentation is the use of a graph with directed edges and asymmetric edge costs.

We show that this approach results in a large improvement regarding quality and robustness in phase contrast images. At the same time, it inherits all the advantageous properties of min-cut segmentation, like global optimality, simultaneous optimization of region and boundary terms, and computational efficiency.

In combination with a simple segmentation propagation our approach yields better segmentation results than the top ranked methods on the PhC-C2DH-U373 phase contrast dataset in the ISBI Cell Tracking Challenge 2014. Furthermore, it is less complex and has fewer tuning parameters than the top ranked method by Magnusson et al. (2015). Our approach was the 2nd ranked method in the ISBI Cell Tracking Challenge 2015 winning against the previously 1st ranked method.

We provide the Matlab source code at <http://lmb.informatik.uni-freiburg.de/resources/opensource/CellTracking/>.

4.1.1 Related work

Cell segmentation in phase contrast images has recently been extensively studied by the Kanade group (*e.g.* see [Li and Kanade \(2009\)](#)). They propose a two step approach by first reconstructing the absolute phase image and then applying basic threshold techniques. This technique works only for completely transparent samples. It fails if the sample contains light absorbing structures, because absorption induced intensity changes and phase-based intensity changes are indistinguishable in standard phase-contrast or DIC microscopic images. [Ambühl et al. \(2012\)](#) propose a series of morphological image processing steps combined with level set approaches. They overcome the problem of the strong halo edges by changing the image during the evolution of the level sets. They apply a morphological top-hat filter to temporarily hide these edges until the contour has passed by. In [Dimopoulos et al. \(2014\)](#) a graph-cut based approach is presented that uses cross-correlation to define the edge weights. Cross-correlation is computed between a defined edge intensity profile and the image intensities in radial direction given the center of a seed point detected before. However, this approach does not naturally handle the segmentation of arbitrarily shaped cells especially non-convex shapes, in this case several seed points have to be considered, merged and additional seed points introduced. The approach by [Magnusson et al. \(2015\)](#) is currently the top ranked method on the ISBI Cell Tracking Challenge dataset (ISB). It mainly relies on a strong tracking approach using the Viterbi algorithm. To increase the performance for the ISBI challenge, they applied a segmentation algorithm based on bandpass filtering, thresholding and watershed transform, which requires several parameters to be adjusted. The U-net by [Ronneberger et al. \(2015\)](#) is a convolutional neural network for biomedical image segmentation which achieved superior performance in several challenges via deep learning. We participated with the U-net in the ISBI Cell Tracking Challenge as well, where it showed excellent performance in particular for DIC and phase contrast datasets.

The usage of asymmetric boundary costs in the min-cut segmentation was already proposed by Boykov in his original graph-cut segmentation paper ([Boykov and Funka-Lea, 2006](#)), but never found its way to the phase contrast microscopy.

4.2 Methods

4.2.1 Cell segmentation

Phase contrast microscopy allows to visualize transparent objects. It turns the invisible phase shifts of the light waves originating from the object into visible intensity changes by using interference with the 90° phase shifted illumination wave. Ideally this would result in an intensity decrease proportional to the object thickness. In reality other effects induce additional shade-off and halo patterns ([Murphy, 2001](#)), such that the intensity drop is only reliably found at the object borders to the surrounding medium.

We cast the segmentation as an energy functional for a mask $M : \Omega \rightarrow \{0, 1\}$ with $\Omega \subset \mathbb{R}^2$ and the given image $I : \Omega \rightarrow \mathbb{R}$. The functional contains a data cost $C_{\text{obj}} : \mathbb{R} \rightarrow \mathbb{R}$ that depends on the intensity, and an edge cost $C_{\text{edge}} : \mathbb{R} \rightarrow \mathbb{R}$ that depends on the intensity gradient at the mask border in outwards direction

$$E(M) = \lambda \int_{\Omega} M(\mathbf{x}) \cdot C_{\text{obj}}(I(\mathbf{x})) d\mathbf{x} + \int_{\Omega} C_{\text{edge}} \left(\langle \nabla M(\mathbf{x}), -\nabla I(\mathbf{x}) \rangle \right) d\mathbf{x}, \quad (4.1)$$

where we define ∇M to be a unit normal vector on the mask boundary and $\mathbf{0}$ elsewhere. The data cost for a gray value v is derived from the foreground intensity histogram $P(v|\mathcal{O})$ and background intensity histogram $P(v|\mathcal{B})$ from training regions. We define it as $C_{\text{obj}}(v) = (P(v|\mathcal{B}) - P(v|\mathcal{O})) / (P(v|\mathcal{O}) + P(v|\mathcal{B}))^3$. This differs from the commonly used negative log-likelihood term, which however comes

³Corrected equation, compared to [Bensch and Ronneberger \(2015\)](#).

with the assumption that the training and testing region histograms are the same or very similar. For example it does not allow to assign a pixel to foreground if its intensity has never been observed among foreground training pixels. In this case infinite region penalties are applied which results in a hard constraint. Our term yields bounded penalties and rather results in a soft constraint. The edge cost for the intensity derivative d is computed as

$$C_{\text{edge}}(d) = \begin{cases} \exp\left(-\frac{d^2}{2\sigma^2}\right) & \text{if } d > 0 \\ 1 & \text{else.} \end{cases} \quad (4.2)$$

I.e., the edge term in the energy functional favors dark-to-bright transitions at the mask borders.

To optimize this energy, we discretize the edge term into 8 directions and solve it by a min-cut as described in [Boykov and Funka-Lea \(2006\)](#). Compared to the “standard” min-cut segmentation approach, our approach results in a directed graph with asymmetric edge weights. In contrast, when using $\exp(-d^2/2\sigma^2)$ for both cases in equation 4.2, an undirected graph with symmetric edge weights is obtained.

Figures 4.4 and 4.5 illustrate the benefit of using asymmetric costs over symmetric costs at an example. Asymmetric costs allow to favor low boundary costs at characteristic dark-bright intensity transitions at cell boundaries in phase contrast images (figure 4.4). Symmetric costs however yield non-specific boundary costs, since also irrelevant bright-dark transitions receive low costs (figure 4.5).

4.2.2 Cell tracking

Our tracking algorithm consists of two parts. *Segmentation propagation* promotes temporally consistent segmentation by propagating segmentation information to subsequent frames, see figure 4.7. Each segmented object is assigned a unique label, and *label propagation* transfers the labels to subsequent frames using a greedy association, see figure 4.6.

Segmentation propagation

Each frame is segmented using min-cut, which yields a binary segmentation mask. To promote temporal consistency, we propagate segmentation information from frame t to frame $t + 1$ in two fashions:

Foreground propagation The eroded mask is set as hard foreground constraint for the min-cut segmentation in the next frame, see figure 4.7a. This adds robustness to the region term in case of insufficient foreground evidence. The size of erosion must be chosen at least as large as the expected motion of object boundary pixels between frames.

Non-merging constraint If it can be assumed that cells do not merge, it is reasonable to prevent separate objects from merging in the next frame. We achieve this by computing a distance transform on the segmentation mask and applying watershed transform seeded at the object locations. The boundaries of the herewith computed “support regions” of each object are set as hard background constraint, see figure 4.7b.

Label propagation

Each segmented object is assigned a unique label, which is propagated to subsequent frames using a greedy algorithm, see figure 4.6: Each segment in frame t transfers its label to the segment in frame $t + 1$ with the highest overlap (measured as intersection over union (IoU)). If a segment in frame $t + 1$ receives multiple labels, it prefers the segment in frame t with the highest overlap and discards the other labels. If a segment receives no label, a new label is assigned.

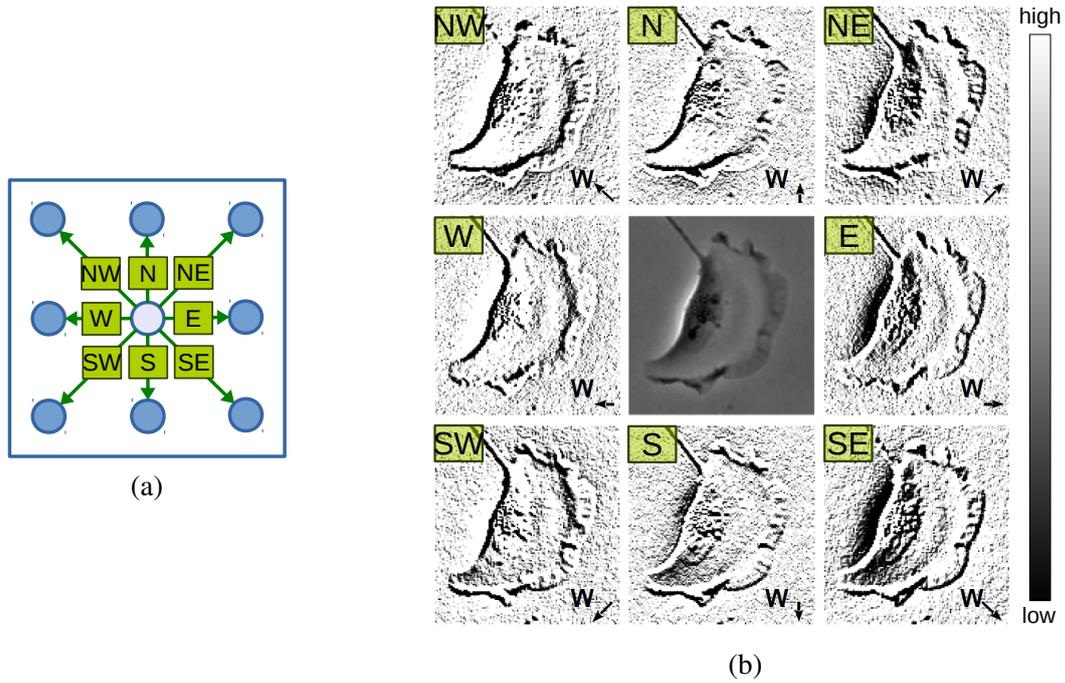


Figure 4.4: Asymmetric boundary costs promote low costs at characteristic dark-bright intensity transitions at cells boundaries in phase contrast microscopy. **(a)** 8-connected pixel neighborhood. Pixels in *blue*, edges and weights in each direction in *green* (only outwards edges shown). **(b)** Boundary costs at each pixel are shown in separate maps for each direction in an 8-connected pixel neighborhood (*arrows and orientation labels* indicate direction). Costs range from zero to one (*black to white*). The phase contrast image is shown in the *middle*.

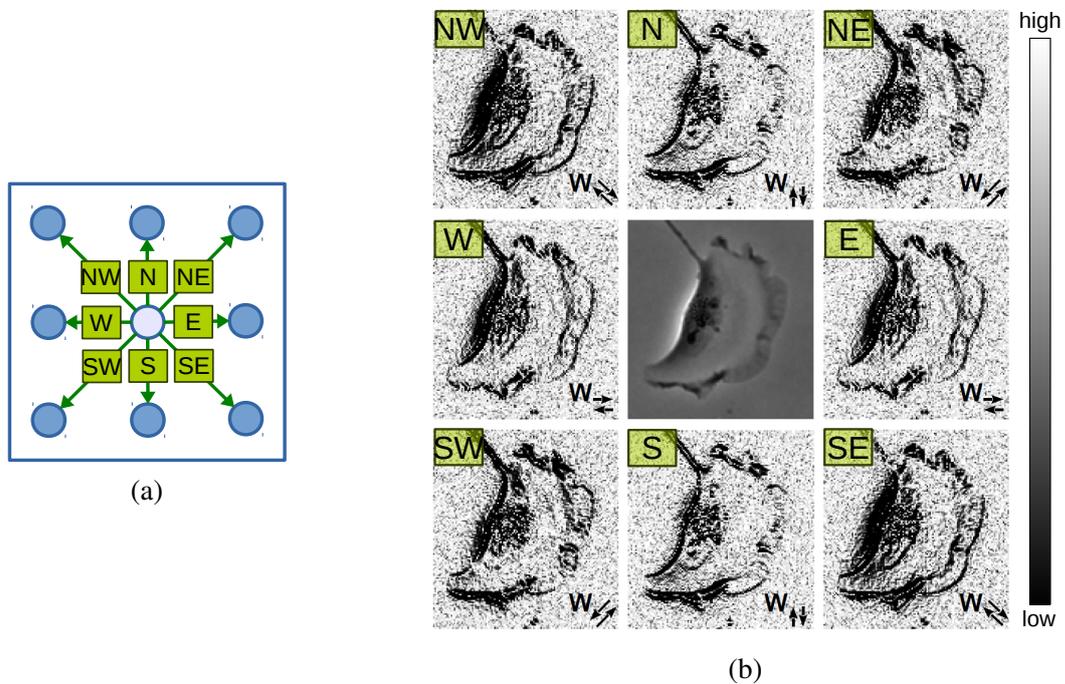


Figure 4.5: Symmetric boundary costs, in contrast to asymmetric costs (figure 4.4), also yield low costs at irrelevant boundaries at bright-dark intensity transitions. Note that the pairs of opposed boundary maps are redundant in the case of symmetric costs. (For full description see caption of figure 4.4.)

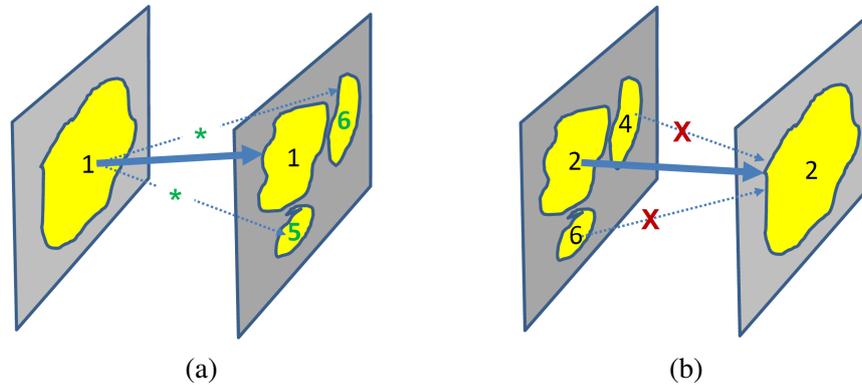


Figure 4.6: Label propagation between two consecutive frames. Propagate labels to overlapping segments using maximum intersection over union (IoU). **(a)** Resolve one-to-many correspondences: Start new tracks (with new label) for segments that do not receive a label from the previous frame. **(b)** Resolve many-to-one correspondences: Stop tracks for segments that do not propagate a label to the next frame. (Images created by Olaf Ronneberger)

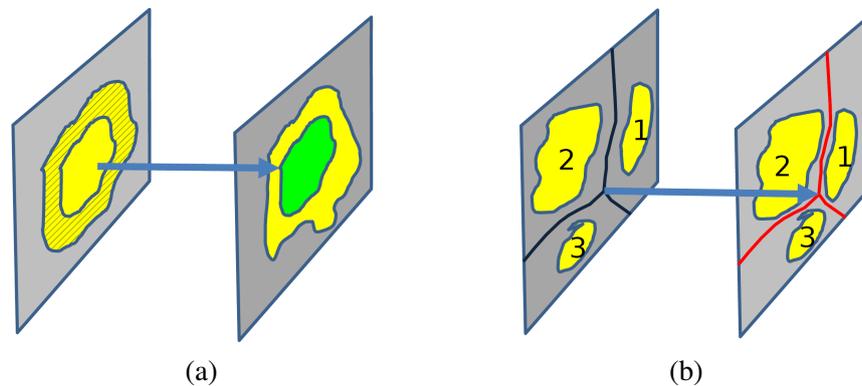


Figure 4.7: Segmentation propagation between two consecutive frames. **(a)** Propagate foreground segmentation information using the eroded segmentation mask (set mask in *green* as hard foreground constraint). **(b)** Propagate partitioning information using borders of „support regions“ (set borders in *red* as hard background constraint). (Images adapted from figure 4.6)

Boundary cost	Seq. 1			Seq. 2		
	F-meas.	Recall	Prec.	F-meas.	Recall	Prec.
Symm.	0.863	0.838	0.889	0.768	0.732	0.808
Asymm. (equation 4.2)	0.896	0.894	0.897	0.835	0.822	0.847

Table 4.1: Boundary detection results on the PhC-C2DH-U373 *training* dataset with 4 pixels tolerance. For comparing symmetric and asymmetric boundary costs, best performing parameters λ and σ (in terms of F-measure) have been chosen for each setting. Parameters were obtained by grid-search over the parameter space, shown in figure 4.8.

4.3 Experiments

4.3.1 Dataset

We evaluate our method on challenging phase contrast microscopy videos of moving cells published by the ISBI Cell Tracking Challenge (ISB; Maška et al., 2014; Ulman et al., 2017). We use the training dataset PhC-C2DH-U373 (provided by Dr. Sanjay Kumar from UC Berkeley). It contains two 2D sequences (115 frames each) of Glioblastoma-astrocytoma U373 cells on a polyacrylimide substrate. Cell segmentation and tracking ground truth is included, along with evaluation tools. Segmentation masks are available only for a subset of frames and cells.

4.3.2 Implementation details

Image intensities are normalized to the interval $[0, 1]$ first. Then, images are background corrected by subtracting the smoothed image (large Gaussian kernel with σ_{bgr}) from the original image. Region histograms for computing the data costs in equation 4.1, are obtained from manual foreground and background scribbles drawn by the authors in one frame of each sequence (that is not contained in the segmentation ground truth). For graph construction we use an 8-neighborhood. The min-cut (with parameters λ , σ) is computed using the maxflow algorithm MATLAB interface (Boykov and Kolmogorov, 2004). Small segments below pixel area a_{min} are discarded. The method starts with segmenting the first frame and then segments subsequent frames using segmentation propagation. Erosion for foreground propagation is computed using a disk-shaped structuring element with radius s_{erosion} . For evaluation, segmentation masks are post processed by a hole-filling algorithm. We set these parameters: $\sigma_{\text{bgr}} = 20\text{px}$, $a_{\text{min}} = 500\text{px}$, $s_{\text{erosion}} = 15\text{px}$. Best performing parameters λ and σ were found by grid-search. The method was implemented in MATLAB⁴.

4.3.3 Evaluation

Boundary detection results We compared the segmentation results obtained when using symmetric and asymmetric boundary costs in terms of boundary detection recall and precision. Recall measures the ratio of ground truth boundary pixels recalled by the computed boundary pixels within 4 pixels tolerance. We used the benchmark code from the Berkeley segmentation benchmark (Ber) to compute boundary detection results. Table 4.1 shows boundary detection results for both sequences and compares symmetric and asymmetric costs. The results show that asymmetric boundary costs perform better, especially in terms of recall. We also compared the stability of results when varying the min-cut parameters λ and σ . Figure 4.8 shows that using asymmetric costs also yields more stable results.

⁴The Matlab source code is available at <http://lmb.informatik.uni-freiburg.de/resources/opensource/CellTracking/>.

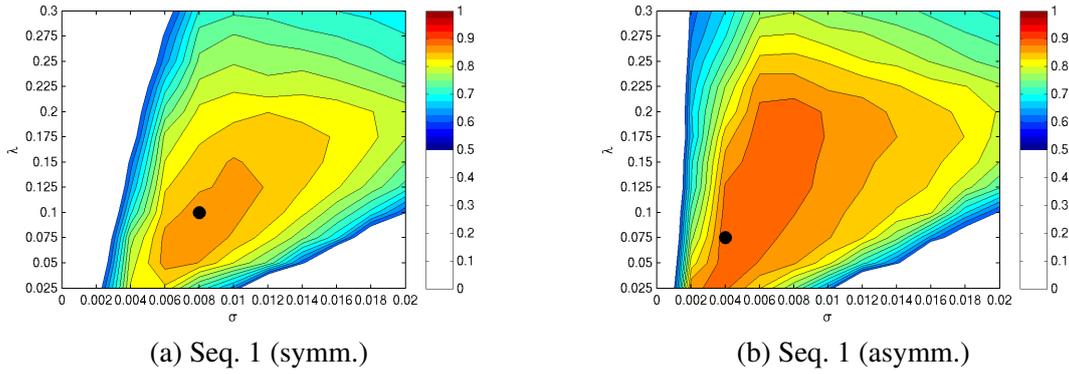


Figure 4.8: Stability of results. F-measure of the boundary detection results as function of the parameters λ and σ . F-measure *isolines* are plotted in the range $[0.5, 1.0]$ in 0.025 intervals. *Black dots* indicate best performing parameters. **(a)** Results for the standard graph cut with symmetric boundary costs. **(b)** Results for our approach using asymmetric boundary costs. Our approach yields better results and is less sensitive to the selected parameters. Compare to table 4.1.

Qualitative results In figure 4.9 qualitative segmentation results are given. They show improvements for detecting very weak phase contrast boundaries, *e.g.* figure 4.9a right column. Also the characteristic halo artifacts in phase contrast microscopy are handled well, due to segmentation at the correct dark-bright intensity transition. In contrast, symmetric boundary costs are strongly affected by the halo effect and predominantly show leaking segmentation at these borders.

Comparison to reported results We further evaluated our method using asymmetric costs in the measures of the ISBI Cell Tracking Challenge (CTC) (ISB). The “average segmentation performance” (Av. SEG) measures the average intersection over union of all reference objects to their matching segmented objects. The “average tracking performance” (Av. TRA) measures how difficult it is to change the computed tracking graph to the ground truth graph. For more details, we refer to ISB. Table 4.2a summarizes our results for segmentation and tracking on the *training* dataset. For comparison, table 4.2b shows results of the top ranked methods reported at the 2nd CTC (ISB).

Our results at the 3rd ISBI CTC (2015) We participated in the 3rd ISBI CTC (2015). The results in table 4.2c show that we were able to win against the previously 1st ranked method (KTH-SE) and also confirm the competitive performance already indicated by the results on the *training* dataset, see table 4.2a. The only approach that was able to win against ours was the U-net (Ronneberger et al., 2015). It was ranked 1st and corresponds to another method we submitted to the 3rd ISBI CTC. It won a Bitplane Attendance Awards 2015⁵. Compared to our method the U-net relies on a sufficient amount of training data and additional data augmentation techniques, while our method can cope with very few training data. In our experiments we used only one frame of each training sequence. The algorithm description of our method submitted to the challenge is available on the challenge website⁶. Our results were published in Nature Methods (Ulman et al., 2017) under the method acronym FR-Be-GE.

Table 4.3 shows the gain achieved by adding segmentation propagation components for temporal consistency to our pure single-frame segmentation approach.

⁵http://www.codesolorzano.com/celltrackingchallenge/Cell_Tracking_Challenge/Attendance_Awards.html

⁶http://www.codesolorzano.com/Challenges/CTC/FR-Be-GE_2015.html

Sequence	Av. SEG	Av. TRA
Seq. 1	0.865	0.983
Seq. 2	0.756	0.915
Seq. 1+2	0.811	0.949

(a) Our results (*training* dataset)

Rank	Group	Av. SEG	Av. TRA
1st	KTH-SE	0.795	0.982
2nd	HOUS-US	0.532	0.921
3rd	IMCB-SG	0.267	0.960

(b) Top ranked methods 2nd ISBI CTC (2014)

Rank	Method/Group (acronym)	Av. SEG	Av. TRA	Av. SEG+TRA
1st	U-net (FR-Ro-GE)	0.920	0.981	1.901
2nd	Ours (FR-Be-GE)	0.826	0.965	1.792
3rd	KTH-SE (KTH-SE)	0.795	0.977	1.773
4th	:	:	:	:
5th	:	:	:	:
:	:	:	:	:

(c) Our results at the 3rd ISBI CTC (2015)

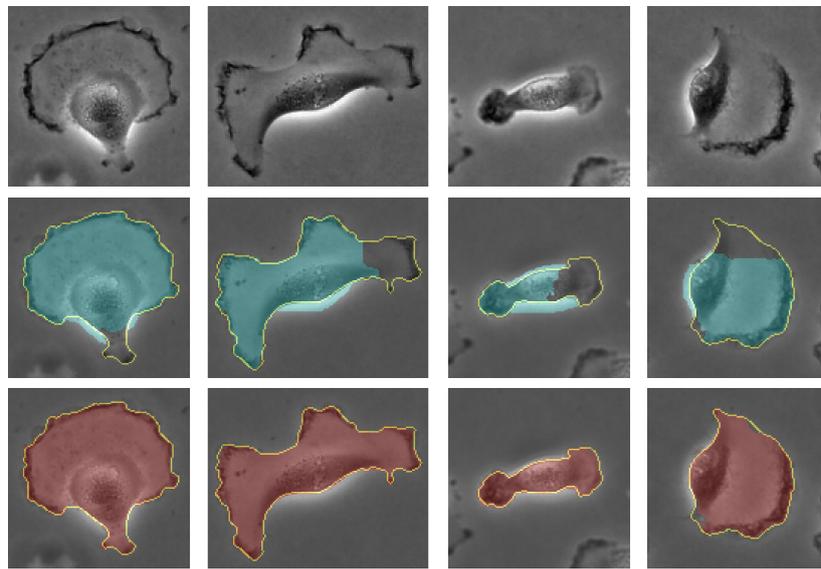
Table 4.2: Our results at the ISBI Cell Tracking Challenge (CTC) (ISB) on the PhC-C2DH-U373 dataset. Results are reported in terms of the average segmentation (Av. SEG) and the average tracking (Av. TRA) performance. **(a)** Our results on the *training* dataset obtained using best parameters from grid-search ($\lambda = 0.2$, $\sigma = 0.006$), for best average segmentation performance on Seq. 1+2. **(b)** Results on the *challenge* dataset of the top ranked methods reported at the 2nd ISBI CTC^a, where the method by Magnusson et al. (2014) (KTH-SE) received 1st rank. **(c)** Results on the *challenge* dataset from the 3rd ISBI CTC^b. Our method (Bensch and Ronneberger, 2015) achieved **2nd rank** in terms of the Av. SEG measure, as well as the overall measure (Av. SEG+TRA), among five international participants with a valid submission for this dataset. We improved over the previously top ranked method (KTH-SE, 2nd CTC), which received 3rd rank. 1st rank results achieved another method we submitted (U-net, described in Ronneberger et al. (2015)).

^aFrom the old challenge website: http://www.codesolorzano.com/celltrackingchallenge/Cell_Tracking_Challenge/Results_Second_CTC.html, which is currently not available anymore (June 9th, 2017).

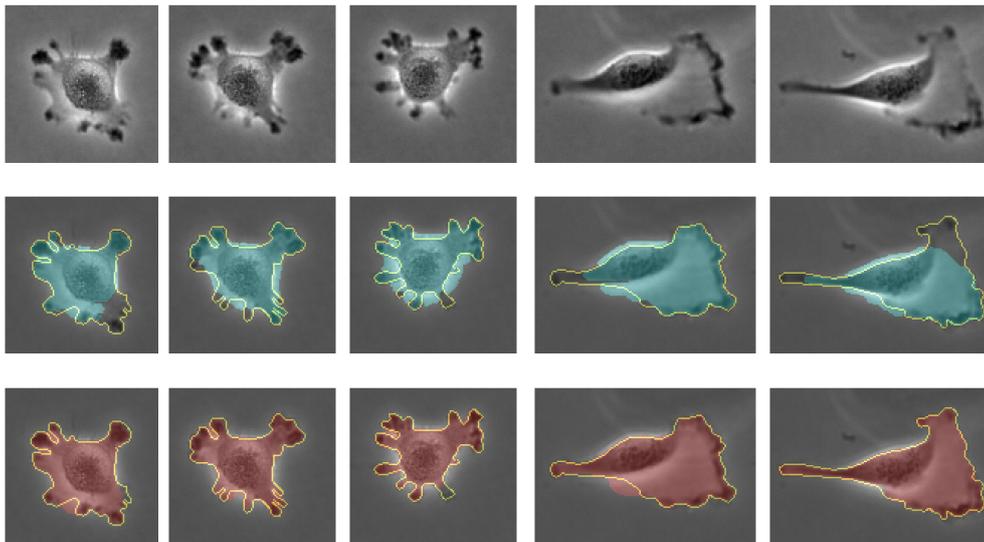
^bReported in Nature Methods (Ulman et al., 2017, figure 6) among the top-three performing methods of the three challenge editions and presented on the challenge website: http://www.codesolorzano.com/Challenges/CTC/Latest_Results.html. Since 2017 the challenge is open for online submissions and up-to-date results are reported on the new challenge website: <http://www.celltrackingchallenge.net>

Setting	Av. SEG	Av. TRA
	(Seq. 1+2)	
Asymm. only	0.738	0.896
Asymm. + FP	0.803	0.937
Asymm. + FP + NM	0.811	0.949

Table 4.3: Evaluation of different settings of our approach when incrementally adding segmentation propagation components: *Foreground propagation* (FP) yields significant improvements, since it supports the data costs in case of insufficient foreground evidence. The *non-merging constraint* (NM) improves results in case of false merging segments.

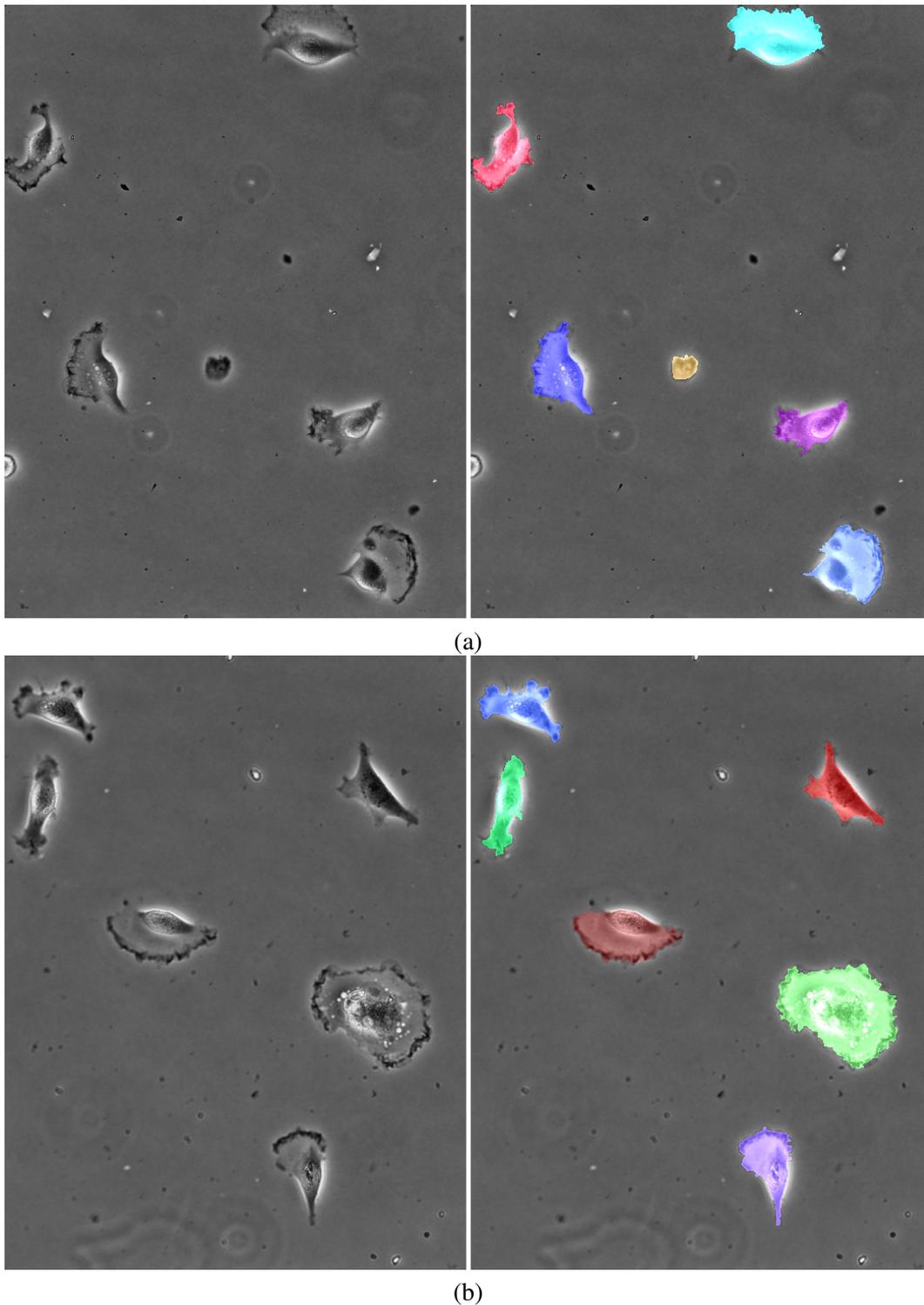


(a)



(b)

Figure 4.9: Qualitative segmentation results. *Top rows* show the raw data. *Cyan masks* show results of the standard graph cut with symmetric boundary costs. *Red masks* show results of our approach with asymmetric costs. The ground truth contour is shown in *yellow*. **(a)** Single frame results for cells of Seq. 1. **(b)** Time-lapse results of Seq. 2 for one cell on frames 5, 9, 15, and another cell on frames 62, 68.



Movie 4.10: Qualitative segmentation results on the *challenge* sequences of dataset PhC-C2DH-U373 from the ISBI Cell Tracking Challenge (ISB; Ullman et al., 2017). (a) Sequence 1. (b) Sequence 2. Raw images (*left*) are shown side by side with an overlay of the segmentation result (*right*). Each segment is shown in a different *color*. (Full movie results are available online: (a) http://lmb.informatik.uni-freiburg.de/people/bensch/phdthesis/PhC-C2DH-U373_01-combined.mkv, (b) http://lmb.informatik.uni-freiburg.de/people/bensch/phdthesis/PhC-C2DH-U373_02-combined.mkv)

4.4 Conclusion and Outlook

The segmentation of cells in phase contrast images is significantly improved by using direction dependent boundary costs. Our approach outperforms the standard min-cut segmentation with symmetric boundary costs. On the PhC-C2DH-U373 phase contrast dataset it beats the top-ranked methods from the ISBI Cell Tracking Challenge 2014. In the following Challenge 2015, the U-net, a deep learning method that we also submitted won, but the proposed method still ranked 2nd. Our results were reported in Nature Methods (Ulman et al., 2017) among the top-three performing methods from all the three challenge editions together.

We assume that cell segmentation in other modalities (transmitted light, dark field, fluorescence, etc.) also profits from asymmetric boundary costs. Our open-source MATLAB implementation will enable a large audience to try it on their data sets.

The accurate segmentation of cells is an important basis for further analysis and quantification in biomedical applications. The next chapter (chapter 5) gives an example, it presents an application where the motion pattern of migrating cells is analyzed from the extracted cell contours.

5 Contour-based Motion Pattern Analysis of Migrating Cells

In the previous chapter we introduced our approach to cell segmentation and tracking in phase contrast images. Based on this, in this chapter we present a direct application in a field of biomedical research investigating migrating cells.

In the following, we first give an introduction including the biological background and related work. Then, in section 5.2 we describe how we detect so-called symmetry-breaking events. These are special cellular events, that are in the focus of the analysis. Finally, we represent cell contour motion patterns using protrusion/retractions maps. These visualize the (signed) contour velocities in a 2D map with polar contour coordinates on the y-axis and time on the x-axis. Section 5.3 details how they are computed. The experimental section 5.4, on the one hand, gives qualitative results of our methods on real biological data in terms of cell detection, segmentation, the detection of symmetry-breaking events and cell contour protrusion/retraction analysis. On the other hand, in section 5.4.4 we summarize the biological results that were extracted using the presented methods.

The methods were developed and the biological results were obtained as a part of a joint project and are partially contained in the PhD thesis by Deniz Saltukoglu (Saltukoglu, 2015) and published in *Molecular Biology of the Cell* (Saltukoglu et al., 2015).

5.1 Introduction

Vicente-Manzanares et al. (2005) introduces cell migration as “... a fundamental process, from simple, uni-cellular organisms such as amoeba, to complex multi-cellular organisms such as mammals. Whereas its main functions comprise mating and the search for food in simple organisms (Manahan et al., 2004), complexity brings a requirement for specialization, which necessitates cell migration-mediated tissue organization, organogenesis and homeostasis (Ridley et al., 2003).” Furthermore, Horwitz and Parsons (1999) emphasize that “cell migration is crucial for embryonic development, the inflammatory immune response, wound repair, and tumor formation and metastasis (Lauffenburger and Horwitz, 1996).” In conclusion, Vicente-Manzanares et al. (2005) states that “... a thorough understanding of the mechanisms underlying cell migration will facilitate development of therapies for the treatment of migration-related disorders.”

In this context, consider figure 5.1a. It illustrates the particular biological application this chapter focuses on. A single cell (a human keratinocyte) is shown that starts migrating from the left to the right side. It can be observed that the cell polarizes while migration is initiated and establishes a distinct front and rear cell edge. This process is referred to as *symmetry-breaking* by biologists.

In Saltukoglu et al. (2015) we studied molecular requirements for random and electrical field (EF)-controlled symmetry-breaking. We tested the role of a cascade of molecules in polarization, ranging from cytoskeleton force generators, cytoskeleton regulators (Rac1, Cdc42 GTPases), inner cell surface molecules to cell surface receptors. For several questions we addressed the “readout” is the migration behavior, and more specifically, the way cells change their shape and cell contour during migration, that is observed in the recorded image data (see figure 5.1). When cells break symmetry and polarize, the *lamellipodium*, which is formed at the front, provides the protrusive force, and the lagging edge, at the back, contracts to detach the cell from the substratum (Saltukoglu et al., 2015).

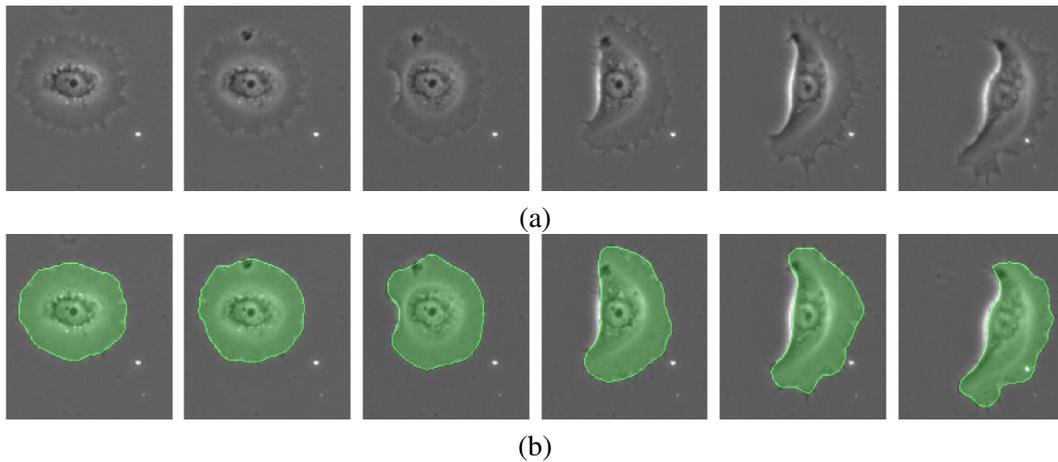


Figure 5.1: Cell shape and contour changes of a single migrating cell and segmentation results. **(a)** Raw data time-lapse of a migrating human keratinocyte acquired by phase contrast microscopy. The cell starts migrating from the *left* to the *right*. **(b)** Segmentation results obtained when applying our approach (Bensch and Ronneberger, 2015), presented in chapter 4. Segmentation masks and detected contours are highlighted (in *green*). Note that the contours are smoothed intentionally due to the analysis requested in this application.

The observed behavior and details of cell contour changes are linked with the underlying biological mechanisms that are in the focus of our study. Thus, the analysis of cell contour changes plays an important role.

The basis of the desired analysis is a segmentation of the cells in each frame, *i.e.* the extraction of the cell contours. The analysis of a large amount of time-lapse image data is not imaginable without automated image analysis techniques. Even if it was possible, manual analysis would be biased, prone to errors and not repeatable. Reconsidering the raw image data shown in figure 5.1a, there are several challenges image segmentation algorithms are faced with: Very weak edges at the cell boundaries, quickly changing morphology and even changing edge appearance, especially at the rear edge after migration is initiated.

For detecting, segmenting and tracking cells we apply our approach (Bensch and Ronneberger, 2015) presented in chapter 4. It uses graph cut with asymmetric boundary costs and was developed for the segmentation of cells in phase contrast images. In the experiments conducted in Bensch and Ronneberger (2015) it outperformed other approaches. Figure 5.1b exemplarily shows the segmentation results we rely on in this chapter and that we obtain with our segmentation approach. Implementation details deviating from Bensch and Ronneberger (2015) are given in section 5.4.2. Further cell detection and segmentation results are shown in section 5.4.4.

Based on the results of cell segmentation and tracking the desired analysis of cell contour changes can be performed. In the following, we first introduce the related work before we detail how we detect symmetry-breaking events and perform cell contour protrusion/retraction analysis.

5.1.1 Related Work

A few works exist in the literature that deal with the analysis of cell contour changes of migrating cells from phase contrast time-lapse microscopy. More specifically, cell contour dynamics of migrating keratocytes (cells that look and behave similar to keratinocytes shown in figure 5.1) are investigated in both Ambühl et al. (2012) and Yam et al. (2007). Both first perform cell segmentation and then analyze cell contour protrusions and retractions.

Ambühl et al. (2012) propose a series of morphological image processing steps combined with level set approaches for segmentation. Yam et al. (2007) use “a variation of the active contours algorithm

(Kass et al., 1988) derived from the gradient vector flow method of Xu and Prince (1998)” for segmenting cells, for which the contour was initialized manually in the first frame. For more information on the related work of cell segmentation in phase contrast images we refer to the related work section 4.1.1.

For the description of contour changes and measuring protrusion and retraction three questions have to be addressed: 1) Which contour parameterization to choose? 2) How to define temporal correspondences? 3) How to transform vectorial contour displacements into a scalar measure that is positive for contour protrusion and negative for contour retraction?

In Ambühl et al. (2012) curvilinear coordinates are used and each contour point in frame t is assigned to the closest contour point in the next frame $t + \Delta t$. This results in a displacement field $\mathbf{d}(s, t)$, with curvilinear coordinates s and time t . Protrusion and retraction is measured by multiplying the norm of the field of displacement with $\text{sign}(\mathbf{d} \cdot \mathbf{n})$, where \mathbf{n} is the outer normal vector onto the cell outline (Ambühl et al., 2012). In Yam et al. (2007) boundary movements are computed by a mechanical model proposed in Machacek and Danuser (2005), it tries to find a mapping that is consistent with and continuous along the contour normal direction. Cell contour positions are finally represented in polar coordinates and continuous space-time plots of protrusion and retraction are generated by the scalar map function (see Yam et al. (2007) for details). Unfortunately, it is not explicitly stated how the scalar protrusion/retraction measure is computed.

Similar to Yam et al. (2007) we use a contour representation in polar coordinates and measure protrusion and retraction in direction of the contour normal, since the tangential component does not contribute to the movement of the cell boundary, see section 5.3.

5.2 Detection of Symmetry-Breaking Events

In this work automated detection of symmetry-breaking events was necessary to facilitate and assist browsing the large amount of data and to extract the relevant time intervals which are in the focus of the analysis. When cells break symmetry their shape changes from a roundish to an elliptic shape, which is illustrated in figure 5.2. We identify symmetry-breaking events by detecting these characteristic shape transitions. We use the eccentricity to measure the “roundness” of the shape as a function of time. To detect symmetry-breaking we define a simple heuristic on the eccentricity function.

5.2.1 Calculation of Eccentricity

For each cell segment the eccentricity is computed as a function over time. The eccentricity

$$e = \frac{f}{a} = \sqrt{1 - \frac{b^2}{a^2}} \quad (5.1)$$

is the ratio of the distance f between the center of the ellipse (with the same second-moments as the segment) and each focus to the length of the semi-major axis a . The semi-minor axis of the ellipse is denoted by b . The eccentricity of an ellipse is in the interval $[0, 1)$, it is 0 for a circle and 1 for an ellipse that is degenerated to a line segment. The function is smoothed using an average filter of width w_{ecc} . Figure 5.2b gives the eccentricity values computed for the depicted segments and figure 5.3a shows an example of the smoothed eccentricity function.

5.2.2 Detection

We detect events as symmetry-breaking events in frame intervals where the eccentricity traverses the range $[0.6, 0.7]$. A reference orientation is computed at the frame where eccentricity is closest to 0.7. It is the orientation of the major axis of the ellipse that has the same second-moments as the cell segment.

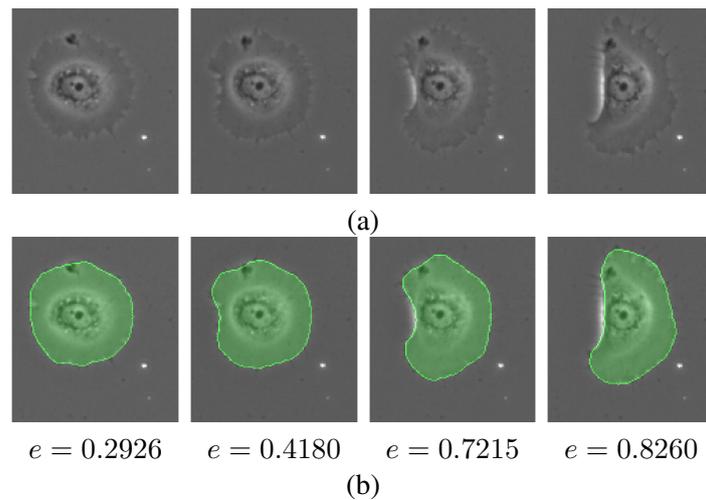


Figure 5.2: Symmetry breaking and initiation of cell migration. **(a)** A subset of frames from within the sequence shown in figure 5.1 depicts the moment when symmetry-breaking takes place. The initially roundish and unpolarized cell turns into an elliptically shaped and polarized cell. The polarized cell exhibits a distinct front/leading edge (*right* side) and rear/lagging edge (*left* side). **(b)** Segmentation results. Segmentation masks and detected contours are highlighted (in *green*). Calculated eccentricity values are given below.

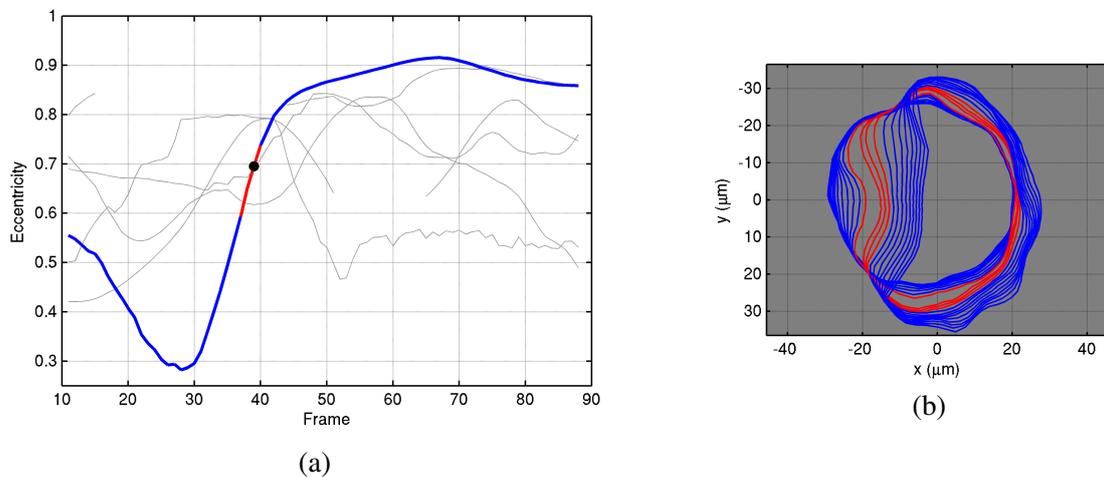


Figure 5.3: Detection of symmetry-breaking events. **(a)** The eccentricity as a function over time (frames) is shown (in *blue*) computed for the cell segment depicted in figure 5.1b. The range for which a symmetry-breaking event is detected is plotted in *red*. The point for which eccentricity is closest to 0.7 and that is used to define the reference orientation is indicated by a *black dot*. Eccentricity functions of other cells in the same field of view are indicated in *light gray*. **(b)** Cell contours falling into a larger time window around the detection are plotted in *blue*, contours within the detection range are plotted in *red*.

In the example in figure 5.3 the detected range is shown in red and the reference point is indicated by a black dot in figure 5.3a.

5.3 Cell Contour Protrusion/Retraction Analysis

The initiation of cell migration goes along with symmetry-breaking and entails contour protrusion at the leading edge and contour retraction at the rear cell edge. Figure 5.4 indicates cell contour protrusion and retraction at an example. In this study, we wish to investigate the underlying biological mechanism that are linked to these contour movements. Thus, it is important to precisely describe and visualize these motion patterns.

We compute protrusion/retraction velocity maps that describe cell contour movements in a polar coordinate system and distinguish outwards (protrusion) and inwards pointing movements (retraction). Figure 5.5 gives an example. In the following we detail how we compute protrusion/retraction maps (denoted PR maps in the following).

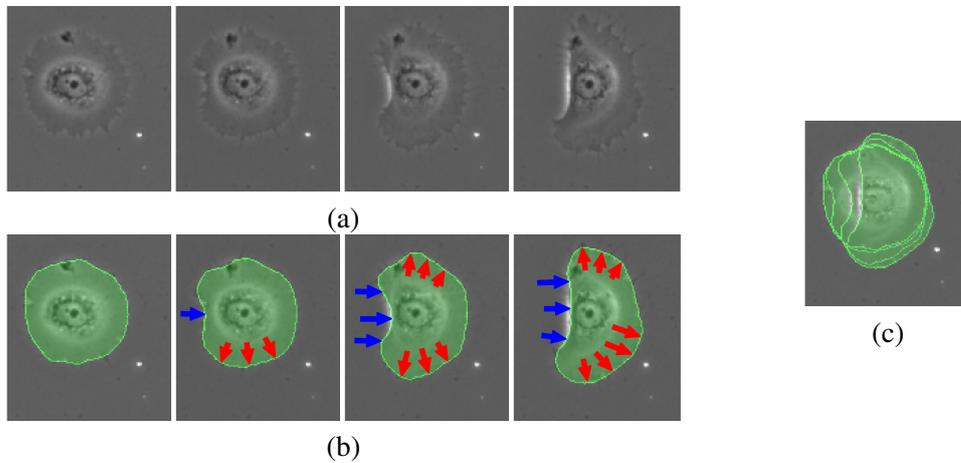


Figure 5.4: Cell contour protrusion and retraction. **(a)** Raw data time-lapse of a migrating cell breaking symmetry. **(b)** Cell contour protrusion is indicated by *red arrows* and retraction is indicated by *blue arrows*. Segmentation results are highlighted in *green*. **(c)** An overlay of all segmentation masks from (b) shows the retracting and protruding cell contour.

5.3.1 Contour Extraction and Normalization

Contours are extracted as the boundary pixel positions of the segmented cells. For cells breaking symmetry the orientation is normalized to the reference orientation defined in section 5.2.2. *I.e.* the set of contours is rotated such that the orientation of the major axis of the segment in the reference frame is upright after normalization.

5.3.2 Calculation of Contour Velocities

For calculating cell contour movements, we establish temporal point correspondences by uniform angular sampling (N_{samples} angular samples) of the contours centered to their mean position. Contour positions are smoothed in temporal direction to compensate for inconsistent fluctuations (using an average filter of width w_{pos}). Finally, contour velocity vectors are computed as the displacement vectors between corresponding contour points.

5.3.3 Protrusion/Retraction Measure

Cell edge protrusion/retraction is measured by the scalar projection

$$\langle \mathbf{v}_i, \mathbf{n}_i \rangle \quad (5.2)$$

of the velocity vectors \mathbf{v}_i onto the local normal direction \mathbf{n}_i of the sampled cell contour, with samples $i \in \{1, \dots, N_{\text{samples}}\}$ (normal direction is defined as pointing outwards). The measure is positive for protrusion and negative for retraction. Figure 5.5 shows an example of extracted contours and the corresponding PR map. The PR map in figure 5.5c clearly shows a cone-shaped blue region corresponding to the retraction of the rear cell edge and red regions corresponding to the protrusion of the front cell edge shown in figure 5.4b and figure 5.5a.

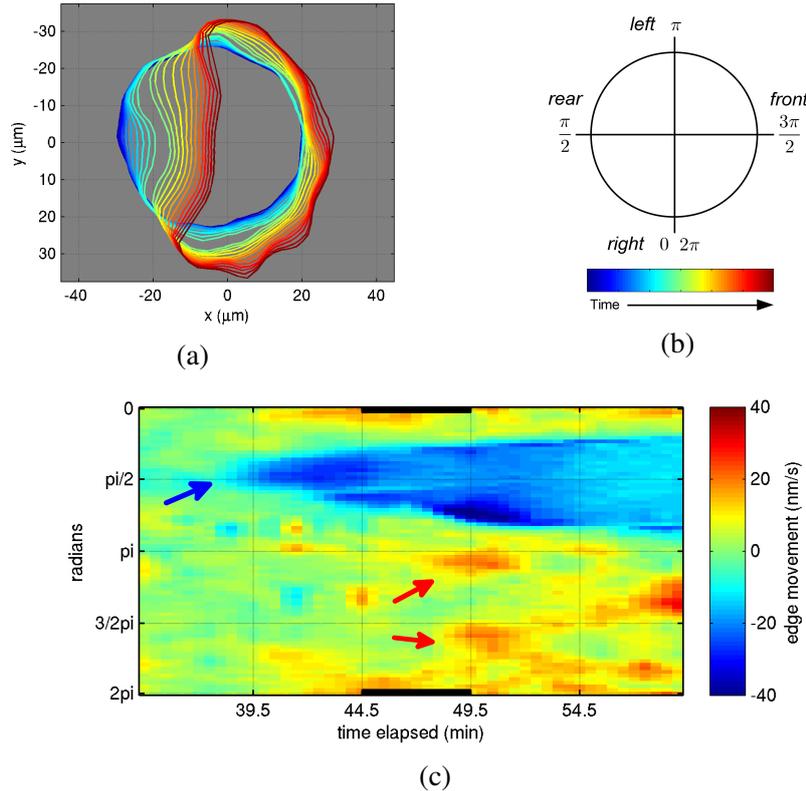


Figure 5.5: Cell contour protrusion/retraction analysis illustrated at the example given in figure 5.4. **(a)** Sequence of detected contours with time color-coded in the depicted time range (from *blue* to *red*). **(b)** Polar coordinate system used for PR maps. Time color code used in contour sequence plots. **(c)** The PR map visualizes movements of contour points (in polar coordinates) over time (on the x-axis). Protrusion is positive signed and color-coded in *red*, retraction is negative signed and color-coded in *blue*. Regions of protrusion and retraction are indicated by *red* and *blue* arrows. These regions correspond to the protrusions and retractions shown in figure 5.4b. *Black* horizontal bars indicate the detection of a symmetry-breaking event.

5.4 Experiments

In our experiments we evaluated cell contour protrusion and retraction of migrating human keratinocytes from several time-lapse recordings acquired by phase contrast microscopy. Migrating keratinocytes were recording under different conditions, such as with and without inhibition of cytoskeletal regulators (Cdc42 and Rac1), low and normal pH concentrations, or with and without an external

electric field (EF). The goal of the experiments was to quantify and compare cell contour protrusion and retraction patterns under these different conditions. In our experiments, we detected, segmented and tracked cells using graph cut with asymmetric boundary costs as described in chapter 4. Furthermore, we detected symmetry-breaking events and performed cell contour protrusion/retraction analysis as described in sections 5.2 and 5.3.

In our experimental results in section 5.4.4, on the one hand, we give qualitative results of cell detection and segmentation, the detection of symmetry-breaking events and cell contour protrusion/retraction analysis. On the other hand, we show which biologically relevant results were extracted using the presented methods. Prior to this, in the following we give details on the datasets and implementation.

5.4.1 Datasets

The data covers about 20 different experimental conditions, for example composed from the application of molecular switches, such as Cdc42 and Rac1, the application of an external electrical field (EF), and the application of an extracellular potential of hydrogen (pH). The effect of different conditions on the migratory behavior of the cells was in the focus of the biological experiments.

We processed 58 datasets of time-lapse recordings of migrating human keratinocytes consisting of 64-190 frames and containing 1-12 cells (mostly 5-8 cells) in the field of view. The image resolution was 692×520 pixels, with $0.65 \mu\text{m}^2$ pixel size. The temporal resolution was 1 min per frame. Figure 5.6 shows exemplary frames from the recorded data.

5.4.2 Implementation Details

We first normalized the image intensities to the interval $[0, 1]$. Compared to our implementation in [Bensch and Ronneberger \(2015\)](#), we do not apply background correction.

Furthermore, no region prior is required, and data costs C_{obj} are initialized to zero. However, we apply hard foreground constraints using heuristically initialized foreground seeds in the first frame. These are initialized reliably from characteristic dark spots inside the cells by intensity thresholding. For the following frames, foreground propagation is applied. In contrast to [Bensch and Ronneberger \(2015\)](#), foreground regions are propagated using both, a hard foreground constraint from the eroded previous segmentation mask, and a hard background constraint from the inverse of the dilated previous segmentation mask. Hence, min-cut optimization is performed only in a narrow band around the cell contour determined in the previous frame. The data costs for hard-constraint foreground and background pixels are set to $C_{\mathcal{O}} = 10000/N_{\mathcal{O}}$ and $C_{\mathcal{B}} = 10000/N_{\mathcal{B}}$ respectively. $N_{\mathcal{O}}$ is the number of foreground and $N_{\mathcal{B}}$ is the number of background pixels. As touching cells are excluded from the analysis before anyway, it is not necessary to apply a non-merging constraint (NM).

Asymmetric edge costs for the intensity derivative d are defined by an ordinary step-function and computed as $C_{\text{edge}}(d) = 0$, if $d > d_{\text{thr}}$ and $C_{\text{edge}}(d) = C_{\text{max}}$ else. For graph construction we used a 4-neighborhood.

The min-cut is computed using the maxflow algorithm MATLAB interface ([Boykov and Kolmogorov, 2004](#)). Small segments below pixel area a_{min} are discarded.

After segmenting the first frame the method uses segmentation propagation for segmenting subsequent frames. We computed erosion and dilation for foreground propagation using a disk-shaped structuring element with radius s_{erosion} and s_{dilation} respectively. For the analysis requested in this application smooth contours are preferred over contours that include all the fine details along the cell boundary, such as the thin “cell hairs” termed *filopodia*. To this end, cell contours are smoothed by applying a Gaussian filter with sigma σ_{smooth} to the binary segmentation mask and then thresholding at 0.5.

These parameters are used: $C_{\max} = 0.005$, $d_{\text{thr}} = 0.005$, $\lambda = 1$, $a_{\min} = 400\text{px}$, $s_{\text{erosion}} = 11\text{px}$, $s_{\text{dilation}} = 17\text{px}$ and $\sigma_{\text{smooth}} = 5.0\text{px}$. The parameters were set manually by inspecting several sample results.

For smoothing the eccentricity function and the contour positions temporally we used an average filter of width $w_{\text{ecc}} = 10$ frames and $w_{\text{pos}} \approx 3$ frames respectively. For the uniform angular sampling of contours we used $N_{\text{samples}} = 80$ angular samples.

The methods were implemented in MATLAB (The MathWorks Inc.). Eccentricity and orientation was computed by MATLAB function `regionprops`.

5.4.3 Quality Control

For the biological evaluations it was essential to ensure high quality segmentation results, since even small contour deviations can lead to strong artifacts in the PR maps. Although automated cell segmentation overall yielded precise results, the extracted cell contours, relevant for the biological evaluations (section 5.4.4), were manually inspected and corrected, which was necessary in multiple cases.

5.4.4 Results

Cell Detection, Segmentation and Tracking

In figure 5.6 we show qualitative results for the segmentation and detection of cells in the recorded images. Figure 5.7 shows qualitative results of tracked single cells. The segmentation results for the full field of view demonstrate that our approach allows to simultaneously detect and segment multiple cells. We obtained precise cell contour segmentations, despite various challenges these phase contrast microscopic images pose. Cell contours are very weak at the leading edge, and at the rear edge the “halo effect” results in a very bright edge with misleading intensity gradients. Even more challenging, the edge appearance is changing when cells start to migrate, which can be observed in figure 5.7. The segmentation is challenging not only because of the complex edge appearance, but also because of the “shade-off effect” in phase contrast microscopy. It results in similar intensities inside and outside the cells. In addition, there is a strong variation in the cell shape.

In figure 5.7 we show the evolving appearance and shape of the cells in the phase contrast microscopy time-lapse. Three cells are given as an example, including two cells that start to migrate and break symmetry, and one cell that, in contrast, does not migrate and does not undergo symmetry-breaking. Apart from the discussed artifacts in phase contrast microscopy, artifacts originating from the data acquisition occur that hamper the segmentation task. As the field of view is recorded in a single slice only and cell positions may slightly vary in z-direction, single cells can get out-of-focus and appear unsharp. From the examples in figure 5.7 it can be observed that the segmentation results remain robust under the described artifacts. Finally, the obtained segmentation results precisely capture the movement of the cell boundary and the evolution of the cell shape.

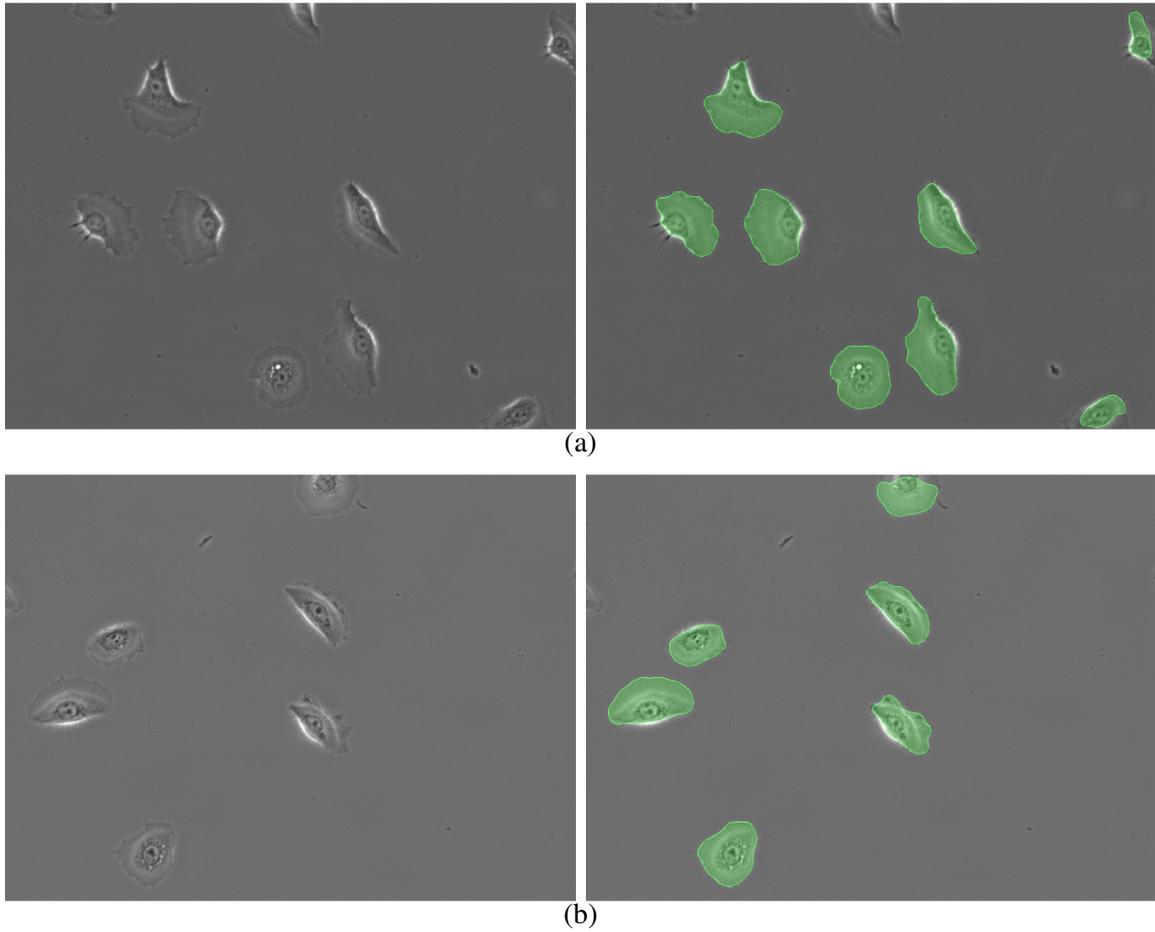


Figure 5.6: Cell detection and segmentation results in the recorded images. *Left column* shows the raw data from phase contrast microscopy. *Right column* shows the segmentation masks and detected contours highlighted in *green*. **(a-b)** Single frame results of two sequences from different experiments are shown. (Cell segments from only partially visible cells at the borders are discarded for the evaluation.)

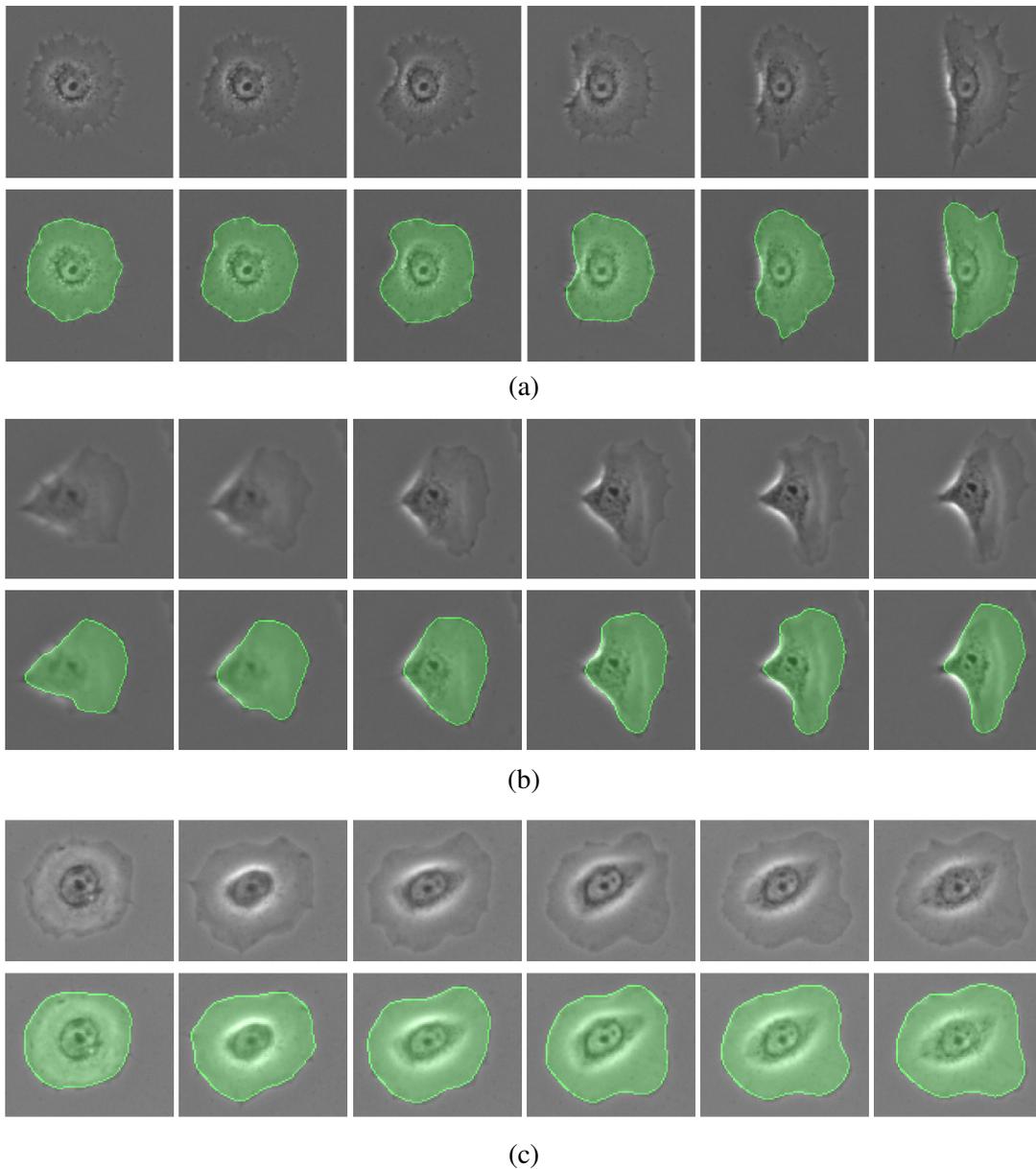


Figure 5.7: Single cell segmentation results. *Top rows* show the raw data from phase contrast microscopy. *Bottom rows* show the segmentation masks and detected contours highlighted in *green*. **(a-c)** Three exemplary cells from different sequences and experiments are shown. **(a,b)** The first two cells migrate and break symmetry. **(c)** The third cell in contrast does not migrate and does not break symmetry, but only shows protrusions along the contour.

Detection of Symmetry-Breaking Events

In figure 5.8 we show symmetry-breaking detection results corresponding to the single cell segmentation results in figure 5.7. For each cell the function of the evolving shape eccentricity is plotted together with the corresponding sequence of extracted contours.

When cells start migrating their shape changes from a roundish to an elliptic shape. This behavior is reflected in the results by the fact that the eccentricity is monotonously increasing in the corresponding time window. Symmetry-breaking is defined as the time windows in which eccentricity traverses the range $[0.6, 0.7]$. Results corresponding to these detected time windows are plotted in red in figure 5.8. The examples suggest that based on the defined, simple criterion for the detection of symmetry-breaking, these events can be detected robustly when they occur (figure 5.8a,b). Accordingly, when no clear symmetry-breaking is present, no event is detected (figure 5.8c).

In the results shown in figure 5.3 and figure 5.8 eccentricity first decreases before it increases again and cells perform symmetry-breaking. This matches the observation that cells flatten and spread before they start to migrate.

The two examples in figure 5.8a,b show that the duration and the speed of symmetry-breaking varies across cells. We measure the duration from the eccentricity function as the time span between the first local minimum before and the first local maximum after the detected time window. The speed can be measured by the slope of the function within the detected time window, for example at the reference point for which eccentricity is closest to the value 0.7 (indicated by a black dot).

Cell Contour Protrusion/Retraction Analysis

In figure 5.9 we show the time color-coded contours and PR maps in line with figures 5.7 and 5.8. The color-code used for the contours indicates the temporal progression of the shape sequences (time is color-coded from blue to red). The PR maps, on the other hand, depict contour protrusions (color-coded in red) and contour retractions (color-code in blue), as described in section 5.3. Cells that undergo symmetry-breaking show strong protrusions and retractions, such as the two cells shown in figure 5.9a,b. In contrast, cells that do not break symmetry only show minor protrusions and retractions along the contour, see figure 5.9c for example. In the examples in figure 5.9a,b we observe retraction of the rear cell edge (see the upper half of the PR maps) and protrusion of the leading edge (see the lower half of the PR maps). The plot allows to extract additional information about the contour deformations. For instance, the pattern of rear cell edge retraction is cone-shaped starting with a peak of retraction which broadens with the progression of rear edge retraction. This pattern in the PR maps geometrically corresponds to a contour retraction that starts with a strong retraction in the middle of the rear cell edge and slowly proceeds to continuously pull the rest of the rear edge towards cell-inside direction. The pattern of leading edge protrusion encodes details about location, strength and temporal progression of protrusions. Consider figure 5.9a for example, there the protrusion pattern indicates first stronger protrusions in south/south-east direction only after the initiation of retraction and first stronger protrusion of the cell front in east direction later, when the complete rear cell edge is already undergoing stronger retraction.

5 Contour-based Motion Pattern Analysis of Migrating Cells

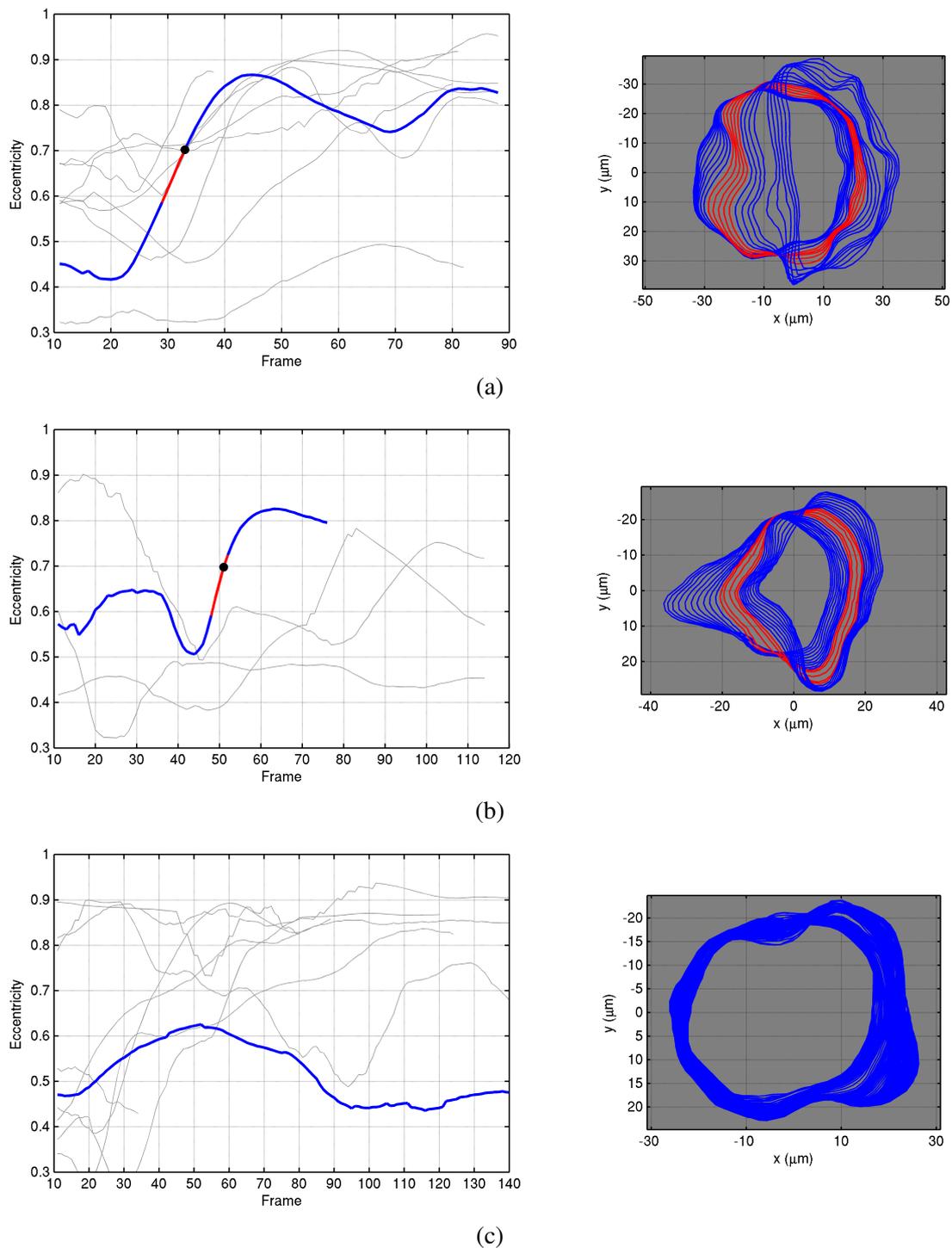


Figure 5.8: Detection of symmetry-breaking events corresponding to the cells shown in figure 5.7a-c. *Left column* shows the eccentricity as a function over time (frames) in *blue*. The range in which symmetry-breaking is detected is plotted in *red*. Note that only intervals are detected in which eccentricity traverses the range $[0.6, 0.7]$. The point for which eccentricity is closest to 0.7 is indicated by a *black dot*. It is used to define the reference orientation. The eccentricity functions of other cells in the same field of view are indicated in *light gray*. *Right column* shows cell contour plots corresponding to a larger time window around the detection, contours within the detection range are plotted in *red*. **(a,b)** Symmetry-breaking is detected for the first two cells, which is in line with the observations in figure 5.7. **(c)** For the third cell however no symmetry-breaking is detected.

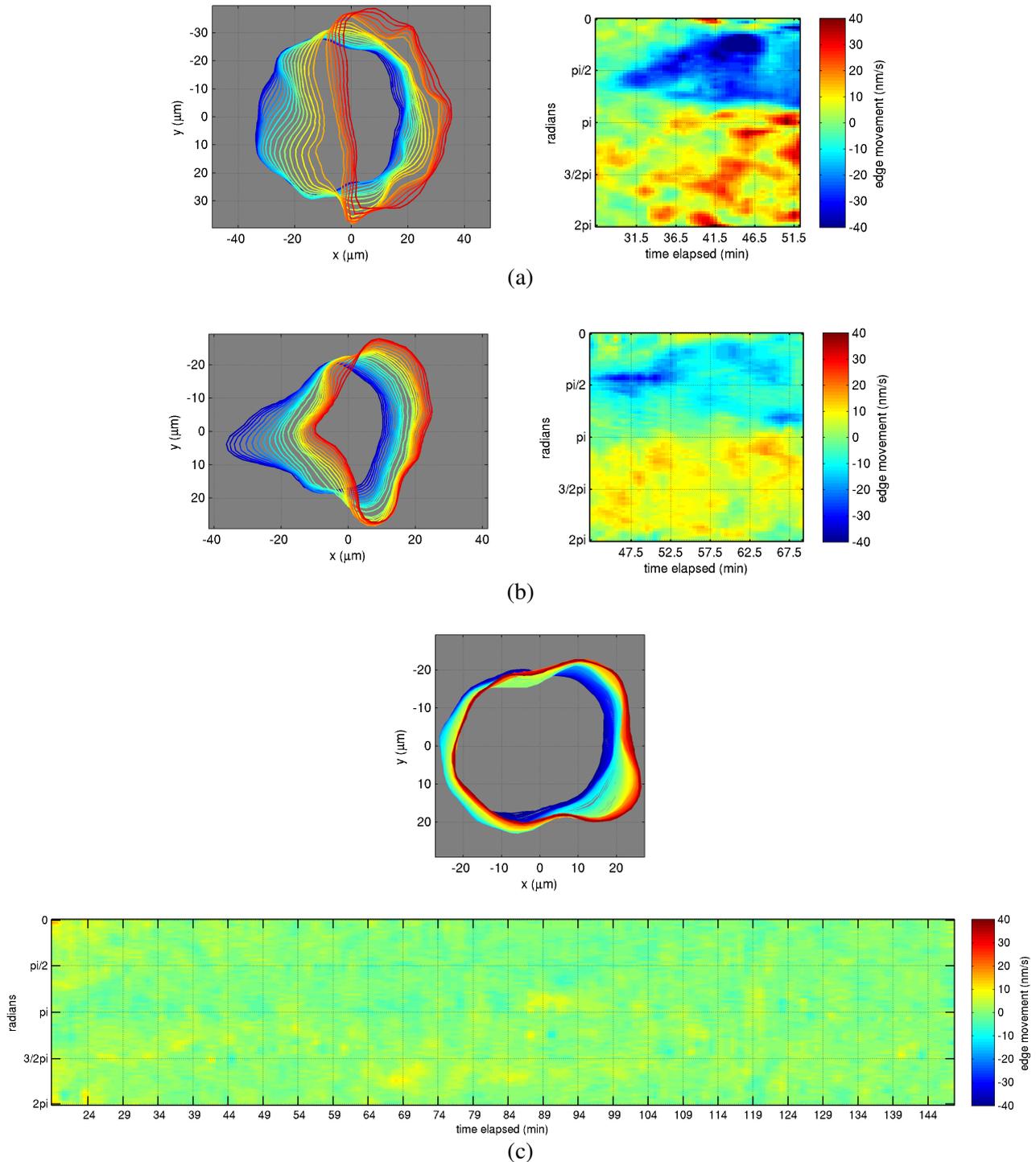


Figure 5.9: Cell contour protrusion/retraction analysis corresponding to the cells shown in figure 5.7a-c. *Left column (a,b) / top row (c)* shows contour plots with time color-coded from *blue* to *red* in the depicted time range. *Right column (a,b) / bottom row (c)* shows the corresponding PR maps with protrusion color-coded in *red* and retraction color-coded in *blue*. **(a,b)** The first two cells show detailed patterns of rear cell edge retraction and front edge protrusion. **(c)** The third cell in contrast only shows minor protrusions and retractions occurring along the contour (*yellow* and *cyan* spots in the PR map).

Biological Results

In this section we give an overview of the results the biologists could extract using the presented methods. We show the results in figures 5.10 and 5.11 and describe them in the following paragraphs. These are summarized versions from Saltukoglu et al. (2015) as indicated below.

The role of cytoskeletal regulators in spontaneous and directional polarization

“Cdc42 and Rac1 are molecular switches that orchestrate the organization of the actomyosin cytoskeleton and regulate the activities of cytoskeletal force generators (Jaffe and Hall, 2005). While Cdc42 is a key molecule in polarity establishment, Rac1 controls the formation of a lamellipodium (Heasman and Ridley, 2008). Both Cdc42 and Rac1 abolished symmetry-breaking in keratinocytes, but produced different morphologies [...]. Cdc42 inhibition led to the apparent loss of protrusive activity, which was confirmed by automated cell outlining of the cellular boundary over time during the polarization assay (figure 5.10). By contrast, Rac1 inhibition led to small but detectable protrusions occurring randomly at the cell periphery (figure 5.10). [...] Whereas Cdc42 decreases the overall ability of cells to make protrusions, Rac1 was required to organized the stability and size of protrusions to form a stable leading edge.”, (Saltukoglu et al., 2015, p. 5).

The electric field increases polarization in cells with compromised cytoskeletal signaling

We next tested the requirements for cytoskeletal force generators and small GTPases Cdc42 and Rac1 in EF-controlled polarization. [...] The inhibition of the small GTPases Cdc42 and Rac1 did not affect directionality determination. However, [...] EF stimulation raised the percentage of symmetry-breaking in Rac1-inhibited cells from 30% (without EF) to 54% (with EF). By extracting the cellular outlines over time, we could further demonstrate how EF could increase symmetry-breaking in Rac1-inhibited cells. This may occur via the clustering of the uncoordinated protrusive activity, and therefore also of the protrusive force essential for symmetry-breaking, towards the cathodal side (figure 5.10). By contrast, the EF was unable to enhance symmetry-breaking in Cdc42-inhibited cells, where all peripheral protrusions were suppressed (figure 5.10). Taken together, these findings show that the cytoskeleton force generators and cytoskeleton regulators are not responsible in directionality determination and that in some conditions, EF application can partially restore the capacity to polarize.”, (Saltukoglu et al., 2015, p. 5).

Morphological differences in anodal and cathodal polarization

To understand the reversal of polarization with low pH_e further, we used protrusion/retraction maps to compare the morphology of symmetry-breaking cells in normal and low pH_e . In normal pH_e , before any movement in the position of the future lagging edge, cells protruded their plasma membrane toward the cathode (in the protrusion/retraction map, this can be seen as red spots before the event of actual symmetry-breaking). This type of leading-edge protrusion was lacking in the low-pH medium condition. By contrast, an invagination of the back preceded the leading-edge protrusion. Initially, this invagination started in the midpoint of the future lagging edge and extended laterally with time, forming the characteristic conical shape of the blue areas representing retraction in the PR map (figure 5.11). Leading-edge protrusion and cell translocation during polarization was generally slower in this condition. Of importance, this type of symmetry-breaking was not a general feature of cells in low-pH medium, as it was not observed in low pH_e without EF stimulation (figure 5.11). The morphological difference in the mode of symmetry-breaking for cathodal and anodal polarization suggested that there may be a distinct set of cytoskeletal regulators responsible for either type of polarization. [...] These results further underscore that cytoskeleton elements are not determining the direction of polarization. Moreover, the observed changes

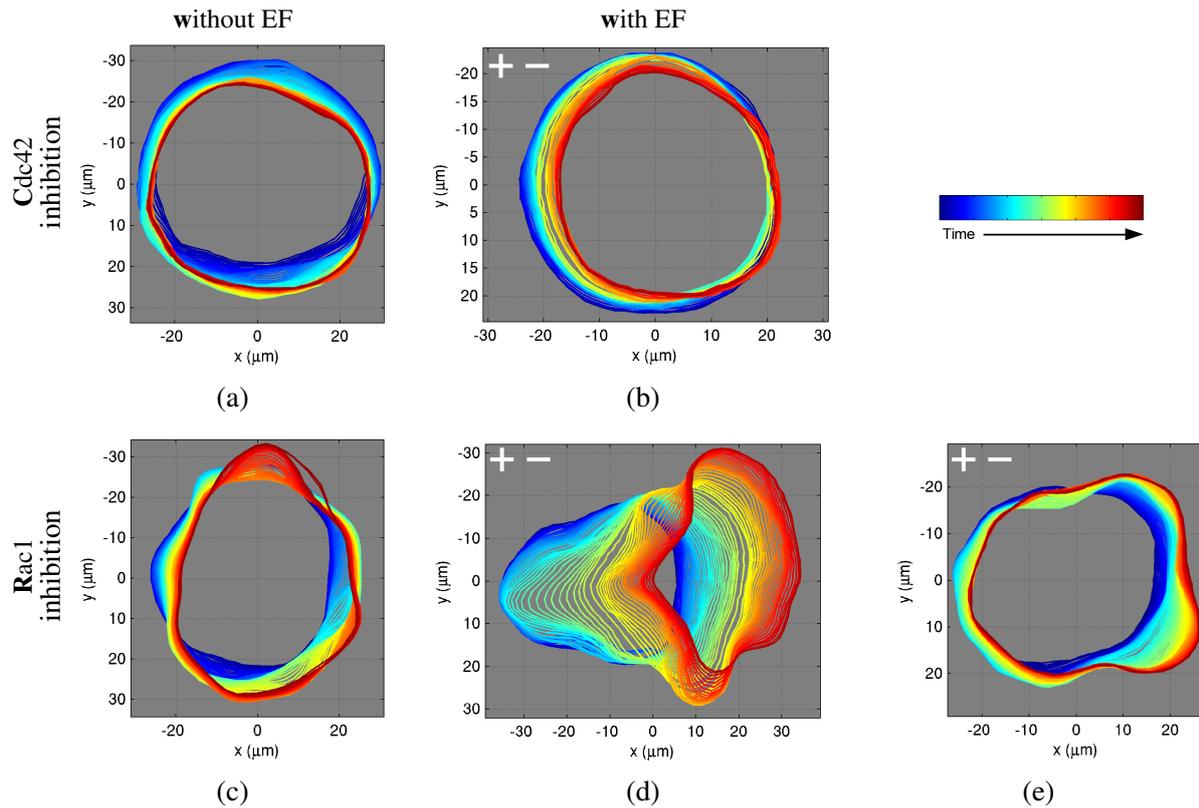


Figure 5.10: “The role of cytoskeletal regulators in spontaneous and directional polarization. [...] Automated boundary detection from phase contrast videos for the representative cells with the indicated treatments. Time progression is represented from *blue* to *red*. **(a-b)** In Cdc42-inhibited cells, EF does not induce any protrusions towards the cathode [...]. **(c-e)** Rac1-inhibited cells show small, unsustained and random boundary protrusions without EF. Two examples are shown for Rac1-inhibited cells with EF. **(d)** In the left example, the cell clearly polarizes to the cathode. **(e)** In the right, small protrusions are driven towards the cathode and suppressed at the anodal side.” Figure and caption adopted from Saltukoglu et al. (2015, figure 2C, p. 4-5).

in pH_i upon pH_e reduction do not seem to affect the cytoskeletal organization in polarizing cells.”, (Saltukoglu et al., 2015, p. 9).

These constitute our contributions to the overall biological findings, which are expected “to have a number of direct implications for epidermal wound healing”, according to Saltukoglu et al. (2015, p. 11).

5 Contour-based Motion Pattern Analysis of Migrating Cells

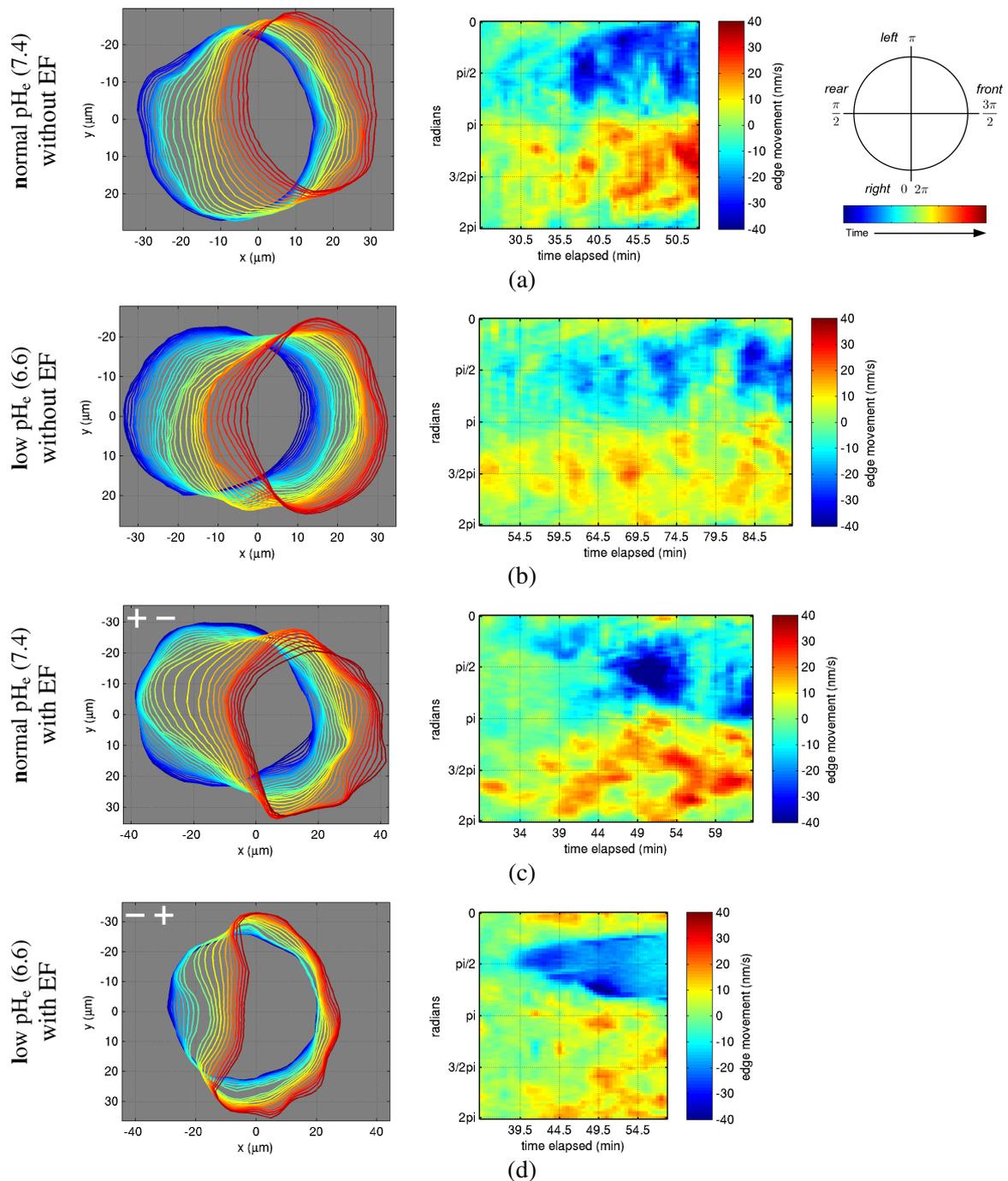


Figure 5.11: “Lowering pH_e reverses the direction of symmetry-breaking. Automated boundary detection from phase contrast videos and the associated PR maps constructed from boundary movements in time for the conditions of low- and normal-pH medium with and without EF. In the detected boundaries, time progression is represented from *blue* to *red*. In the PR maps, the x-axis represents the time, and the y-axis is the position of the cellular boundary from 0 to 2π. Representative cells have been turned [...] so that their leading edges always face the right side. **(a-c)** Symmetry breaking in the normal-pH condition (with and without EF), as well as low-pH condition (without EF), is featured by front protrusions preceding the retraction of the back. **(d)** By contrast, in low pH (with EF), there is an invagination of the cellular boundary at the lagging edge. The invagination event is represented by the unique conical shape of the blue retraction area of the corresponding PR map. Moreover, there is also a general reduction of leading-edge protrusion [...] in this condition.” Figure and caption adopted from [Saltukoglu et al. \(2015, Figure 6C, p. 8-9\)](#).

5.5 Conclusions

This chapter demonstrates an example in which we successfully applied our image segmentation approach (Bensch and Ronneberger, 2015) to the analysis of cell migration behavior in human keratinocytes. Our approach resulted in accurately extracted cell contours. Based on these, the detection of symmetry-breaking events allowed to automatically browse large amounts of data and focus the analysis on these relevant cellular events. Finally, we visualized and precisely measured cell contour protrusions and retractions in PR maps. Through the analysis of the protrusion/retraction behavior of cells under different conditions we answered several biologically relevant questions.

6 Conclusion

In this thesis, we presented novel methods for motion pattern analysis in biomedical applications that enable quantitative comparisons in 2D and 3D time-lapse microscopy. Motion patterns are very relevant in developmental biology, for example, as they describe processes, such as cell migration, that lead to the formation of tissue, the creation of organs and even whole organisms - generally speaking, the creation of life. Automated quantitative comparisons between wild-type and genetically manipulated specimen are important in biomedical research to identify significant differences. The results can help to decode the function of specific genes.

We developed a general method to detect motion anomalies in 3D+time data using a new, efficient and robust “supertrajectory” representation. Anomalies are detected by reconstructing a whole test pattern via placing spatiotemporally deformed instances of a prototype pattern. A modified hashing approach and a new method to elastically register trajectory patterns allows efficient and robust reconstruction. A prototype model is learned from training sequences to define accepted variations. Our approach performed well in detecting subtle anomalies on a new motion anomaly dataset of juggling patterns on which it outperformed chaotic invariants for anomaly detection (Wu et al., 2010). We successfully applied the method on the quantitative comparison of motion patterns from the early development of zebrafish embryos and obtained a detailed spatiotemporal analysis of differences between wild-type and morphant embryos.

We presented a method to detect specific motion patterns in 3D+time data using spatiotemporal geometric models. In particular, we developed a model to detect *cell intercalations*, which is an essential motion pattern in developmental biology. The method builds on motion trajectories of single cells, while a spatiotemporal geometric model defines the motion pattern to be detected. Our approach is robust to noisy and interrupted measurements and handles the variability within the class of 3D intercalations. We successfully applied the method to obtain quantitative comparisons of cell intercalations and their motion statistics between wild-type and mutant embryos in the early development of zebrafish.

To segment cells in phase contrast microscopy we presented a new robust, effective and surprisingly simple approach. The key feature of our method is that it strongly favors dark-to-bright transitions at the boundary of (arbitrarily shaped) segmentation masks. The segmentation mask is effectively found via a fast min-cut approach. Our graph contains directed edges with asymmetric edge weights, which has never been applied to phase contrast microscopy, to the best of our knowledge. Our evaluation shows that asymmetric edge weights yield better results while being less sensitive to the graph cut parameters. Our method outperformed the top ranked methods from the ISBI Cell Tracking Challenge (CTC) (ISB) 2014 on the phase contrast dataset and reached second place in the ISBI CTC 2015.

A successful application of our approach for cell segmentation in phase contrast images was the analysis of cell migration and cell contour motion patterns in human keratinocytes. We developed methods to detect *symmetry-breaking* events and to compute protrusion/retraction maps. These enabled automatic browsing of large amounts of data and an analysis of cell contour protrusion/retraction patterns focused on symmetry-breaking events. Our approach provided the biologists with a quantification of several experimental conditions (Saltukoglu et al., 2015; Saltukoglu, 2015).

Outlook

As deep learning (LeCun et al., 2015; Rusk, 2016) is currently the most exciting and promising field for advances, it seems reasonable to focus future research into that direction. Therefore, in this outlook we discuss some ideas about using deep learning in the context of the presented work.

Currently, solving rather “basic” tasks, such as classification, detection, or segmentation, works well with deep learning. However, realizing an end-to-end deep learning approach that solves complex tasks, such as those considered in this thesis, is very challenging to date. The main challenges have been discussed in the introduction in section 1.3. A mixture of conventional and deep learning methods might be required on the way of approaching a pure deep learning solution.

In the following, we list several ideas about using deep learning. At first, two general basically different approaches are discussed.

- **Discriminative approach (hypothesis testing):** In a discriminative setting a classifier, in this case a deep neural network, could be trained to distinguish between datasets of different classes. If this succeeds, then differences exist. This setting is comparable to hypothesis testing. This approach, however, has several drawbacks. First of all, in case training fails, it is not clear, whether there are no differences, or whether the learning algorithm fails to train the classifier. Secondly, the fact that a classifier can distinguish the classes does not imply that the differences are relevant. They might just be artifacts, for example, introduced by different imaging settings. Finally, this approach only addresses the existence not the localization of differences.
- **Generative approach (anomaly detection):** In a generative approach, a generative model could be learned from the normal class data. Differences in the test data would be detected as deviations from the learned model. This setting is comparable to anomaly detection. In deep learning this approach could be realized using auto-encoders. Auto-encoders have already been used in video anomaly detection, *e.g.* in Hasan et al. (2016); Chong and Tay (2017). An auto-encoder consists of an encoder, that is a contracting network for extracting a compact feature representation, and a decoder, that is an expanding network for reconstructing the original image. When trained with normal class data, the auto-encoder learns a representation for the normal class. When applied to test data, the auto-encoder tries to reconstruct the data using the learned representation, which will succeed only up to the degree of similarity to the normal class. The difference between the reconstructed and the input image highlights locations that are not reconstructed well, *i.e.* that deviate from the normal model. This approach sounds promising, however several drawbacks remain. First of all, it is not guaranteed whether the detected differences are relevant, or just device-dependent variations or other artifacts. Secondly, although the localization of deviations is available, a meaningful representation of the differences is beyond the scope of this approach.
- **Unsupervised domain adaptation:** An interesting approach for unsupervised domain adaptation has been presented in Kamnitsas et al. (2016). It uses an adversarial network architecture to learn a segmentation task, while at the same time forcing the network to learn features that are robust across different domains of the input data, specifically different imaging protocols. The advantage of this approach is that domain adaptation only requires the domain labels, but no annotations for the primary segmentation task. This approach could be integrated to obtain device-independent properties. It could be used in the discriminative setting described above, for example. A drawback of adversarial networks is that they are very hard to train.
- **Recurrent neural networks (RNNs):** RNNs allow processing of sequential data and have proven to be very powerful in numerous scenarios. RNNs could be used in our context in multiple ways, for example, to process temporal sequences, or a sequence of datasets. Potentially, RNNs could also be used to sequentially iterate the data to identify differences. Furthermore,

recent encoder-recurrent-decoder (ERD) architectures (Fragkiadaki et al., 2015; Donahue et al., 2017) might be helpful. The separate encoder and decoder networks allow to integrate different input and output formats, while the recurrent part of the network allows to learn temporal dynamics.

- **Introducing prior knowledge:** Introducing prior knowledge helps to simplify the problem and steers the solution into the desired direction. Instead of building on the raw input data only, one could augment the input data by precomputed segmentation, detection, or optical flow information. One could also directly build on a precomputed trajectory representation and possibly use graph convolutional networks (Defferrard et al., 2016; Kipf and Welling, 2016) for processing this type of data. However, including specific prior knowledge is very difficult in deep learning and moreover deviates from the concept of end-to-end learning.
 - **Extraction and localization of differences:** With regard to the discriminative approach described above, for a successfully trained discriminator network, differences could be localized in the image domain by using backpropagation techniques or methods that invert representations, such as the method presented by Dosovitskiy and Brox (2016), or other methods discussed in the related work section in Dosovitskiy and Brox (2016). The auto-encoder approach discussed above, among generative approaches, directly provides a localization of differences in the image space. Extracting a meaningful representation of differences will however be very difficult.
 - **Statistical significance:** As mentioned in Angermueller et al. (2016) in the section “Pitfalls”, assessing statistical significance “is currently difficult using deep learning methods”, and will probably remain a challenge for future research.
-

Bibliography

- The Berkeley Segmentation Dataset and Benchmark*. Available at: <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>.
- ISBI Cell Tracking Challenge*. Available at: <http://www.celltrackingchallenge.net>. Previous URL: <http://www.codesolorzano.com/Challenges/CTC>.
- A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz. Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):555–560, March 2008.
- M. Ambühl, C. Brepsant, J.-J. Meister, A. Verkhovsky, and I. Sbalzarini. High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *Journal of Microscopy*, 245(2): 161–170, 2012.
- C. Angermueller, T. Pärnamaa, L. Parts, and O. Stegle. Deep learning for computational biology. *Molecular Systems Biology*, 12(7):878–n/a, 2016.
- B. Antic and B. Ommer. Video parsing for abnormality detection. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2415–2422, 2011.
- Y. Benezeth, P. M. Jodoin, V. Saligrama, and C. Rosenberger. Abnormal events detection based on spatio-temporal co-occurrences. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2458–2465, June 2009.
- R. Bensch and O. Ronneberger. Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1220–1223, April 2015. (oral presentation).
- R. Bensch, O. Ronneberger, B. Greese, C. Fleck, K. Wester, M. Hülskamp, and H. Burkhardt. Image analysis of arabidopsis trichome patterning in 4d confocal datasets. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 742–745, June 2009.
- R. Bensch, S. Song, O. Ronneberger, and W. Driever. Non-directional radial intercalation dominates deep cell behavior during zebrafish epiboly. *Biology Open*, 2(8):845–854, 2013.
- R. Bensch, T. Brox, and O. Ronneberger. Spatiotemporal deformable prototypes for motion anomaly detection. In *British Machine Vision Conference (BMVC)*, pages 189.1–189.12, September 2015. (oral presentation).
- R. Bensch, N. Scherf, J. Huisken, T. Brox, and O. Ronneberger. Spatiotemporal deformable prototypes for motion anomaly detection. *International Journal of Computer Vision*, 122(3):502–523, May 2017.
- P. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

Bibliography

- G. B. Blanchard, A. J. Kabla, N. L. Schultz, L. C. Butler, B. Sanson, N. Gorfinkiel, L. Mahadevan, and R. J. Adams. Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation. *Nature Methods*, 6(6):458–464, June 2009.
- O. Boiman and M. Irani. Similarity by composition. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems (NIPS)*, pages 177–184. MIT Press, 2007a.
- O. Boiman and M. Irani. Detecting irregularities in images and in video. *International Journal of Computer Vision*, 74(1):17–31, 2007b.
- Y. Boykov and G. Funka-Lea. Graph cuts and efficient n-d image segmentation. *International Journal of Computer Vision*, 70(2):109–131, 2006.
- Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, September 2004.
- T. Brox and J. Malik. Large displacement optical flow: descriptor matching in variational motion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):500–513, 2011.
- R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, September 1995.
- V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3):15:1–15:58, July 2009.
- Y. S. Chong and Y. H. Tay. Abnormal event detection in videos using spatiotemporal autoencoder. *CoRR*, abs/1701.01546, 2017.
- Y. Cong, J. Yuan, and J. Liu. Sparse reconstruction cost for abnormal event detection. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3449–3456, 2011.
- Y. Cong, J. Yuan, and Y. Tang. Video anomaly search in crowded scenes via spatio-temporal motion context. *IEEE Transactions on Information Forensics and Security*, 8(10):1590–1599, 2013.
- H. M. Dee and D. C. Hogg. Detecting inexplicable behaviour. In *British Machine Vision Conference (BMVC)*, pages 477–486, 2004.
- M. Defferrard, X. Bresson, and P. Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *CoRR*, abs/1606.09375, 2016.
- S. Dimopoulos, C. E. Mayer, F. Rudolf, and J. Stelling. Accurate cell segmentation in microscopy images using membrane patterns. *Bioinformatics*, 30(18):2644–2651, 2014.
- J. Donahue, L. A. Hendricks, M. Rohrbach, S. Venugopalan, S. Guadarrama, K. Saenko, and T. Darrell. Long-term recurrent convolutional networks for visual recognition and description. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):677–691, April 2017.
- A. Dosovitskiy and T. Brox. Inverting visual representations with convolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- B. Everitt, S. Landau, M. Leese, and D. Stahl. *Cluster Analysis*. Wiley series in probability and statistics. Wiley, 2011.

- K. Fragkiadaki, S. Levine, P. Felsen, and J. Malik. Recurrent network models for human dynamics. In *IEEE International Conference on Computer Vision (ICCV)*, pages 4346–4354, December 2015.
- J. D. Gibbons and S. Chakraborti. *Nonparametric Statistical Inference*. Chapman & Hall/CRC Press, Taylor & Francis Group, Boca Raton, FL, 5th edition, 2011.
- B. Greese, K. Wester, R. Bensch, O. Ronneberger, J. Timmer, M. Hülkamp, and C. Fleck. Influence of cell-to-cell variability on spatial pattern formation. *IET Systems Biology*, 6(4):143–153, 2012.
- M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury, and L. S. Davis. Learning temporal regularity in video sequences. *CoRR*, abs/1604.04574, 2016.
- E. Heller, K. Kumar, S. Grill, and E. Fuchs. Forces generated by cell intercalation tow epidermal sheets in mammalian tissue morphogenesis. *Developmental Cell*, 28(6):617–632, March 2014.
- M. Hollander and D. A. Wolfe. *Nonparametric Statistical Methods*. John Wiley & Sons, Inc., Hoboken, NJ, 1999.
- A. R. Horwitz and J. T. Parsons. Cell migration—movin’ on. *Science*, 286(5442):1102–1103, 1999.
- W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank. A system for learning statistical motion patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1450–1464, 2006.
- K. Irvine and E. Wieschaus. Cell intercalation during drosophila germband extension and its regulation by pair-rule segmentation genes. *Development*, 120(4):827–841, 1994.
- K. Kamnitsas, C. F. Baumgartner, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, A. Nori, A. Criminisi, D. Rueckert, and B. Glocker. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. *CoRR*, abs/1612.08894, 2016.
- D. A. Kane, K. N. McFarland, and R. M. Warga. Mutations in half baked/e-cadherin block cell behaviors that are necessary for teleost epiboly. *Development*, 132(5):1105–1116, 2005.
- M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.
- P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H. Stelzer. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science*, 322(5904):1065–1069, 2008.
- P. J. Keller, A. D. Schmidt, A. Santella, K. Khairy, Z. Bao, J. Wittbrodt, and E. H. K. Stelzer. Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. *Nature Methods*, 7(8):637–642, August 2010.
- R. Keller, L. Davidson, A. Edlund, T. Elul, M. Ezin, D. Shook, and P. Skoglund. Mechanisms of convergence and extension by cell intercalation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1399):897–922, July 2000.
- R. Keller, L. A. Davidson, and D. R. Shook. How we are shaped: The biomechanics of gastrulation. *Differentiation*, 71(3):171 – 205, 2003.
- M. Keuper, R. Bensch, K. Voigt, A. Dovzhenko, K. Palme, H. Burkhardt, and O. Ronneberger. Semi-supervised learning of edge filters for volumetric image segmentation. In *Pattern Recognition (Proc. DAGM)*, LNCS, pages 462–471. Springer, 2010.
- J. Kim and K. Grauman. Observe locally, infer globally: A space-time mrf for detecting abnormal activities with incremental updates. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2921–2928, June 2009.

Bibliography

- C. B. Kimmel, W. W. Ballard, S. R. Kimmel, B. Ullmann, and T. F. Schilling. Stages of embryonic development of the zebrafish. *Developmental Dynamics*, 203(3):253–310, 1995.
- T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. *CoRR*, abs/1609.02907, 2016.
- L. Kratz and K. Nishino. Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1446–1453, June 2009.
- M. Lachnit, E. Kur, and W. Driever. Alterations of the cytoskeleton in all three embryonic lineages contribute to the epiboly defect of pou5f1/oct4 deficient {MZspg} zebrafish embryos. *Developmental Biology*, 315(1):1 – 17, 2008.
- Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.
- C. Li, Z. Han, Q. Ye, and J. Jiao. Visual abnormal behavior detection based on trajectory sparse reconstruction analysis. *Neurocomputing*, 119:94 – 100, 2013.
- K. Li and T. Kanade. Nonnegative mixed-norm preconditioning for microscopy image segmentation. In *International Conference on Information Processing in Medical Imaging (IPMI)*, LNCS, pages 362–373, 2009.
- S. S. Lienkamp, K. Liu, C. M. Karner, T. J. Carroll, O. Ronneberger, J. B. Wallingford, and G. Walz. Vertebrate kidney tubules elongate using a planar cell polarity-dependent, rosette-based mechanism of convergent extension. *Nature Genetics*, 44(12):1382–1387, December 2012.
- K. Lunde, H.-G. Belting, and W. Driever. Zebrafish pou5f1/pou2, homolog of mammalian oct4, functions in the endoderm specification cascade. *Current Biology*, 14(1):48 – 55, 2004.
- M. Machacek and G. Danuser. Morphodynamic profiling of protrusion phenotypes. *Biophysical Journal*, 90(4):1439–1452, November 2005.
- K. Magnusson, J. Jaldén, and H. M. Blau. *Cell tracking using bandpass filtering and the viterbi algorithm*, 2014. Description of the algorithm available at: http://www.codesolorzano.com/Challenges/CTC/KTH-SE_2014_files/isbi170.pdf.
- K. Magnusson, J. Jaldén, and H. M. Blau. *Cell tracking using bandpass filtering and the viterbi algorithm*, 2015. Description of the algorithm available at: http://www.codesolorzano.com/Challenges/CTC/KTH-SE_2015_files/KTH-SE.pdf.
- V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos. Anomaly detection in crowded scenes. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1975–1981, June 2010.
- M. Maška, V. Ulman, D. Svoboda, P. Matula, and P. Matula, *et al.* A benchmark for comparison of cell tracking algorithms. *Bioinformatics*, 30(11):1609–1617, 2014.
- R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 935–942, June 2009.
- A. M. Middleton, C. D. Bosco, P. Chlap, R. Bensch, H. Harz, F. Ren, S. Bergmann, S. Wend, W. Weber, K. Hayashi, M. D. Zurbriggen, R. Uhl, O. Ronneberger, K. Palme, C. Fleck, and A. Dovzhenko. Data-driven modeling of intracellular auxin fluxes indicates a dominant role of the er in controlling nuclear auxin uptake. *Cell Reports*, 22(11):3044–3057, March 2018.

- D. Murphy. Phase contrast microscopy. *Fundamentals of Light Microscopy and Electronic Imaging*, pages 97–112, 2001.
- E. Málaga-Trillo, G. P. Solis, Y. Schrock, C. Geiss, L. Luncz, V. Thomanetz, and C. A. O. Stuermer. Regulation of embryonic cell adhesion by the prion protein. *PLoS Biology*, 7(3):1–15, March 2009.
- H. Nait-Charif and S. J. McKenna. Activity summarisation and fall detection in a supportive home environment. In *IEEE International Conference on Pattern Recognition (ICPR)*, volume 4, pages 323–326, August 2004.
- C. Piciarelli, C. Micheloni, and G. L. Foresti. Trajectory-based anomalous event detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(11):1544–1554, November 2008.
- O. P. Popoola and K. Wang. Video-based abnormal human behavior recognition - a review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):865–878, November 2012.
- G. Reim and M. Brand. Maternal control of vertebrate dorsoventral axis formation and epiboly by the pou domain protein spg/pou2/oct4. *Development*, 133(14):2757–2770, 2006.
- O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of *LNCS*, pages 234–241. Springer, 2015.
- M. Rose. *Elementary Theory of Angular Momentum*. Dover books on physics and chemistry. Dover, 1995.
- M. J. Roshtkhari and M. D. Levine. An on-line, real-time learning method for detecting anomalies in videos using spatio-temporal compositions. *Computer Vision and Image Understanding*, 117(10):1436 – 1452, 2013.
- N. Rusk. Deep learning. *Nature Methods*, 13(1):35–35, January 2016.
- V. Saligrama and Z. Chen. Video anomaly detection based on local statistical aggregates. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2112–2119, June 2012.
- V. Saligrama, J. Konrad, and P. m. Jodoin. Video anomaly identification. *IEEE Signal Processing Magazine*, 27(5):18–33, September 2010.
- D. Saltukoglu. *Spontaneous and electric field-induced directional symmetry breaking in keratinocytes*. PhD thesis, Albert-Ludwigs-Universität Freiburg, 2015.
- D. Saltukoglu, J. Grünewald, N. Strohmeyer, R. Bensch, M. Ulbrich, O. Ronneberger, and M. Simons. Spontaneous and electric field-controlled front-rear polarization of human keratinocytes. *Molecular Biology of the Cell*, 26(24):4373–4386, December 2015.
- B. Schmid, G. Shah, N. Scherf, M. Weber, K. Thierbach, C. P. Campos, I. Roeder, P. Aanstad, and J. Huisken. High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics. *Nature Communications*, 4, July 2013.
- T. Shimizu, T. Yabe, O. Muraoka, S. Yonemura, S. Aramaki, K. Hatta, Y.-K. Bae, H. Nojima, and M. Hibi. E-cadherin is required for gastrulation cell movements in zebrafish. *Mechanisms of Development*, 122(6):747 – 763, 2005.
- R. R. Sillito and R. B. Fisher. Semi-supervised learning for anomalous trajectory detection. In *British Machine Vision Conference (BMVC)*, pages 1035–1044, 2008.

Bibliography

- J. Slack. *Essential Developmental Biology*. Wiley, 2nd edition, 2005.
- S. Song, S. Eckerle, D. Onichtchouk, J. A. Marrs, R. Nitschke, and W. Driever. Pou5f1-dependent egf expression controls e-cad endocytosis, cell adhesion, and zebrafish epiboly movements. *Developmental cell*, 24(5):486–501, March 2013.
- N. Sundaram, T. Brox, and K. Keutzer. Dense point trajectories by gpu-accelerated large displacement optical flow. In *European Conference on Computer Vision (ECCV)*, 2010.
- R. Tomer, K. Khairy, F. Amat, and P. J. Keller. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods*, 9(7):755–763, July 2012.
- V. Ulman, M. Maška, K. E. G. Magnusson, O. Ronneberger, C. Haubold, N. Harder, P. Matula, P. Matula, D. Svoboda, M. Radojevic, I. Smal, K. Rohr, J. Jaldén, H. M. Blau, O. Dzyubachyk, B. Lelieveldt, P. Xiao, Y. Li, S.-Y. Cho, A. Dufour, J. C. Olivo-Marin, C. C. Reyes-Aldasoro, J. A. Solis-Lemus, R. Bensch, T. Brox, J. Stegmaier, R. Mikut, S. Wolf, F. A. Hamprecht, T. Esteves, P. Quelhas, Ö. Demirel, L. Malmström, F. Jug, P. Tomancák, E. Meijering, A. Muñoz-Barrutia, M. Kozubek, and C. O. de Solor. An objective comparison of cell-tracking algorithms. *Nature Methods*, 14:1141–1152, 2017.
- S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991.
- M. Vicente-Manzanares, D. J. Webb, and A. R. Horwitz. Cell migration at a glance. *Journal of Cell Science*, 118(21):4917–4919, 2005.
- S. E. Webb and A. L. Miller. Ca²⁺ signalling and early embryonic patterning during zebrafish development. *Clinical and Experimental Pharmacology and Physiology*, 34(9):897–904, 2007.
- S. Winkelbach, S. Molkenstruck, and F. M. Wahl. Low-cost laser range scanner and fast surface registration approach. In *Pattern Recognition (Proc. DAGM)*, pages 718–728. Springer, LNCS, 2006.
- S. Wu, B. E. Moore, and M. Shah. Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2054–2060, 2010.
- C. Xu and J. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, March 1998.
- P. T. Yam, C. A. Wilson, L. Ji, B. Hebert, E. L. Barnhart, N. A. Dye, P. W. Wiseman, G. Danuser, and J. A. Theriot. Actin-myosin network reorganization breaks symmetry at the cell rear to spontaneously initiate polarized cell motility. *The Journal of Cell Biology*, 178(7):1207–1221, 2007.
- T. Yasunaga, S. Hoff, C. Schell, M. Helmstädter, O. Kretz, S. Kuechlin, T. Yakulov, C. Engel, B. Müller, R. Bensch, O. Ronneberger, T. Huber, S. Lienkamp, and G. Walz. The polarity protein inturnd links nphp4 to daam1 to control the subapical actin network in multiciliated cells. *The Journal of Cell Biology*, 211(5):963–973, 2015.
- F. Zernike. Phase contrast, a new method for the microscopic observation of transparent objects part {II}. *Physica*, 9(10):974 – 986, 1942.