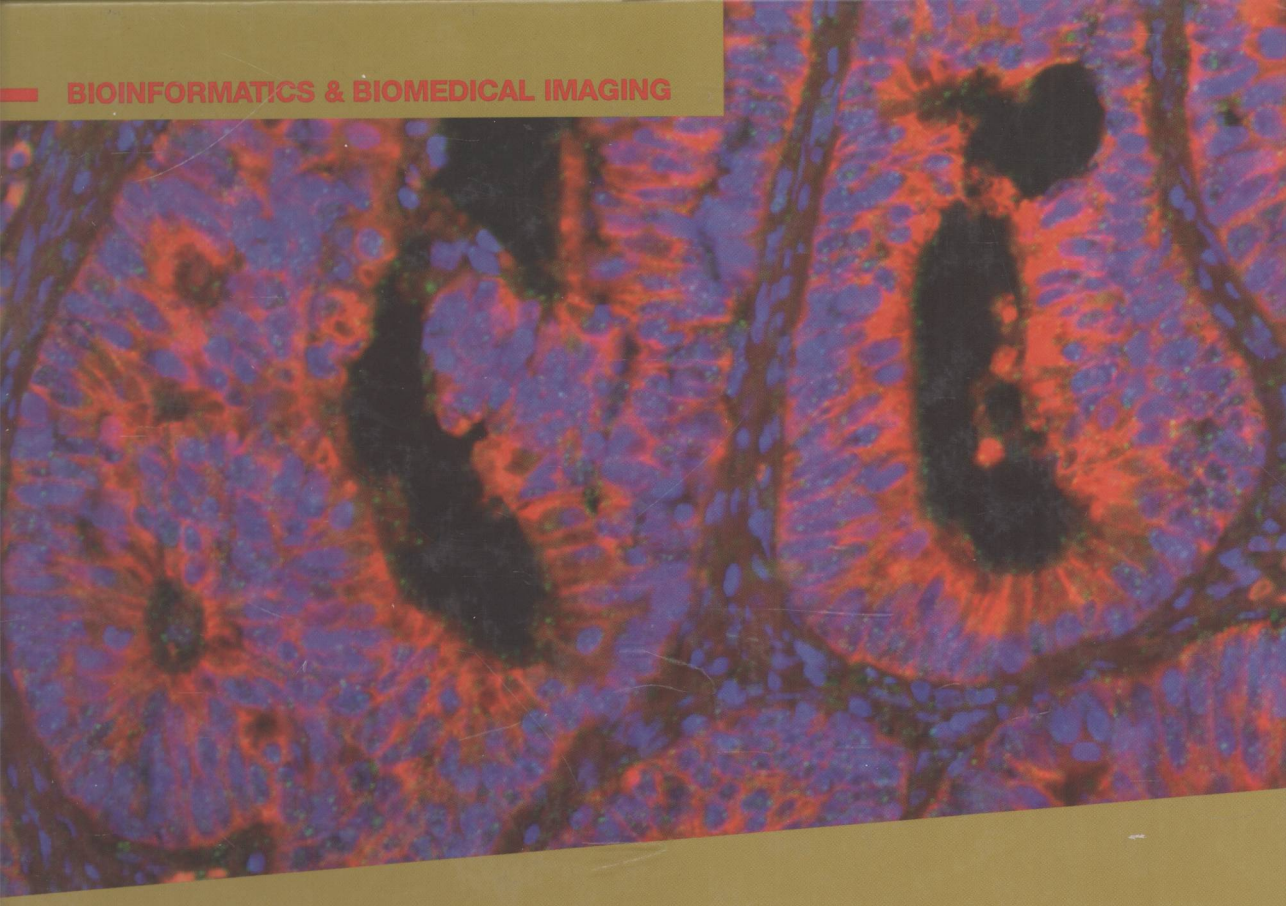
BIOINFORMATICS & BIOMEDICAL IMAGING



# Microscopic Image Analysis

## *for Life Science Applications*

Jens Rittscher
Raghu Machiraju
Stephen T. C. Wong

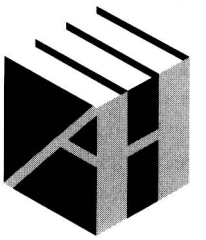
*editors*

*CD-ROM Included*

# Microscopic Image Analysis for Life Science Applications

Jens Rittscher
Raghu Machiraju
Stephen T. C. Wong

*Editors*

ARTECH
HOUSE
BOSTON | LONDON
artechhouse.com

**Library of Congress Cataloging-in-Publication Data**
A catalog record for this book is available from the U.S. Library of Congress.

**British Library Cataloguing in Publication Data**
A catalogue record for this book is available from the British Library.

ISBN-13: 978-1-59693-236-4

**Cover design by Yekaterina Ratner**
**Cover image courtesy of Ali Can, Musodiq Bello, Xiaodong Tao, Michael Gerdes, Anup Sood, Michael C. Montalto, and Fiona Ginty, Authors of Chapter 8, GE Research**


**685 Canton Street**
**Norwood, MA 02062**



DISCLAIMER OF WARRANTY

The technical descriptions, procedures, and computer programs in this book have been developed with the greatest of care and they have been useful to the author in a broad range of applications; however, they are provided as is, without warranty of any kind. Artech House, Inc. and the authors and editors of the book titled *Microscopic Image Analysis for Life Science Applications* make no warranties, expressed or implied, that the equations, programs, and procedures in this book or its associated software are free of error, or are consistent with any particular standard of merchantability, or will meet your requirements for any particular application. They should not be relied upon for solving a problem whose incorrect solution could result in injury to a person or loss of property. Any use of the programs or procedures in such a manner is at the user's own risk. The editors, authors, and publisher disclaim all liability for direct, incidental, or consequent damages resulting from use of the programs or procedures in this book or the associated software.

10 9 8 7 6 5 4 3 2 1

# Microscopic Image Analysis for Life Science Applications

## Artech House Series
## Bioinformatics & Biomedical Imaging

Series Editors
Stephen T. C. Wong, The Methodist Hospital and Weill Cornell Medical College
Guang-Zhong Yang, Imperial College

*Advances in Diagnostic and Therapeutic Ultrasound Imaging*, Jasjit S. Suri, Chirinjeev Kathuria, Ruey-Feng Chang, Filippo Molinari, and Aaron Fenster, editors

*Biological Database Modeling*, Jake Chen and Amandeep S. Sidhu, editors

*Biomedical Informatics in Translational Research*, Hai Hu, Michael Liebman, and Richard Mural

*Genome Sequencing Technology and Algorithms*, Sun Kim, Haixu Tang, and Elaine R. Mardis, editors

*Image Modeling of the Human Eye*, Rajendra Acharya U, Eddie Y. K. Ng, and Jasjit S. Suri, editors

*Life Science Automation Fundamentals and Applications*, Mingjun Zhang, Bradley Nelson, and Robin Felder, editors

*Microscopic Image Analysis for Life Science Applications*, Jens Rittscher, Raghu Machiraju, and Stephen T. C. Wong, editors

*Next Generation Artificial Vision Systems: Reverse Engineering the Human Visual System*, Maria Petrou and Anil Bharath, editors

*Systems Bioinformatics: An Engineering Case-Based Approach*, Gil Alterovitz and Marco F. Ramoni, editors

# Foreword

Microscopy is one of the oldest biological research techniques, and yet it has lagged behind other methods in adopting the statistical, quantitative, and automated approaches that characterize biology in the age of systems biology. However, work performed over the past dozen years has begun to make biological imaging, especially fluorescence microscopy, a suitable tool for large-scale, systematic, and automated studies of biological processes. The need for computational analysis of biological images is driven in part by the advent of automated, high throughput microscopes, but the significant limitations of visual analysis of biological images is an equally important factor. The goal of the field of bioimaging informatics is to enable the direct extraction of assertions about biological events from images or sets of images so that these assertions can be integrated with those from other biological methods into comprehensive models of biological systems.

The chapters that Jens Rittscher, Raghu Machiraju, and Stephen T. C. Wong have assembled in this book describe the range of methods used in bioimaging informatics and provide the reader with both an introduction to the challenges of the field and the current state of the art in addressing those challenges.

The book begins with four introductory chapters on microscopy, fluorescent probes, and basic image processing. The next 12 chapters review work from the authors' groups in the context of particular goals but from which two brad themes emerge. The first theme is recognizing patterns using both supervised and unsupervised methods from images from sources as varied as high-throughput screen, histochemistry, and whole embryo imaging. The second theme is analyzing spatiotemporal dynamics for processes with time scales as varied as organelle movement and cell cycle progression. The need for approaches to combine analysis results from different methods and channels is also addressed.

The last four chapters of this book focus primarily on a different task: the reconstruction from an image set (which may include images from different modalities) of an adequate model describing the original sample. The first two of these chapters focus on neuronal images, and the last two chapters focus on bridging microscopy with larger-scale imaging modalities such as magnetic resonance imaging.

This book should be valuable both to readers familiar with biology and microscopy, but lacking knowledge of current image informatics methods, and to those familiar with image analysis, but seeking challenging new applications.

The field is in its infancy, but the challenges and opportunities make it one of the most exciting in all of biomedical research. It is hoped that this book

will stimulate the reader to contribute to realizing the potential of bioimage informatics for understanding complex biological processes and behaviors.

*Robert F. Murphy*
*Ray and Stephanie Lane Professor of Computational Biology*
*Carnegie Mellon University*
*Pittsburgh, Pennsylvania*
*July 2008*

# Preface

Advances in the microscopy of living cells, small organisms, and tissues have changed life science research. Modern microscopy has developed from a convergence of advances in cell manipulation, probe chemistry, solid state physics, electronics, optics, and image analysis. Today, fluorescence markers suitable for live cell imaging are readily available. Not only is it possible to image biological structures in three dimensions by tagging particular proteins, but it is also possible to monitor a variety of biological processes occurring in cells and tissues. State-of-the-art microscopes enable us to record such biological processes at extremely high spatial and temporal resolutions.

Optical microscopy enables the determination of the structure within a cell and the location of specific molecules within a cell. More importantly, these insights are obtained with relatively less sophisticated instrumentation than other forms of microscopy, especially those relying on electron tunneling methods. More recently, certain protocols of optical microscopy provide more than structural information. Various subcellular properties and events, as manifested through the presence and interationc of certain molecules, can also be determined through the use of chemical and biological probes.

Important application areas include cancer research, toxicity screening, digital pathology, and neuroscience research. All of these applications require an in-depth understanding of the biology, the probe chemistry, the underlying imaging physics, and the associated analysis of images. It is clear that new breakthroughs in this rapidly emerging field will only be accomplished through effective interdisciplinary research. Automated state-of-the-art microscopes enable researchers to acquire data that can no longer be analyzed manually. Such data sets pose a number of challenges that are very distinct from conventional clinical imagery in their size and abundance and the detail of relevant features and their statistics. Sophisticated algorithms are necessary to process such imagery and extract biologically relevant measurements.

This book presents an in-depth exploration of issues related to automated image analysis applied to life science applications. The primary aim of this book is to provide a bridge between the biomedical image analysis and bioscience communities. Striking a balance between an entirely self-contained presentation and a review of the state of the art in the filed is always challenging. This book is divided into five major parts: Introduction, Subcellular Structures and Events, Structure and Dynamics of Cell Populations, Automated Tissue Analysis, and In Vivo Microscopy. The introduction consist of five chapters that are tutorial chapters. The remaining chapters cover a wide spectrum of examples illustrating the potential of biomedical vision algorithms.

While each of the chapters can be read independently, some of the image analysis concepts that are applied require an understanding of basic concepts such as

image segmentation and visual tracking. Rather than presenting a review of relevant algorithm and image analysis methods, this book presents techniques in a concrete application context. The biological background necessary to understand the specific technical challenges associated to each problem is included in the chapter. The content of each section is briefly described next.

### Part I: Introduction

In Chapter 1, Gunjan Agarwal presents a brief but comprehensive overview of light microscopy. The various forms of microscopy, contrast mechanisms, and aspects of sample preparation techniquies that are commonly employed for biological light microscopy are introduced.

A key aspect of fluorescence microscopy is molecular probes that enable the targeting of structural components and dynamic processes in living celss and tissues. In Chapter 2 Michael Davidson and his collaborators present a detailed overview of fluorescent probes. Throughout the chapter, the underlying chemical structure of such probes and associated fluorescent mechanisms are discussed.

The remaining three chapters of the introduction provide an overview of the necessary signal processing and image analysis tools that are necessary to examine and quantify acquired fluorescent microscopy images. Jelena Kovacevic and Gustavo K. Rohde present a summary of computer-aided-image analysis tools and tasks for microscopy in Chapter 3. This chapter is structured around mathematical tools and concepts that hold great potential to advance the state of the art for automated artifact removal and information extraction.

Michael Unser and his collaborators present an introduction to fluorescence microscopy in Chapter 4. This chapter puts the current developments into a historical perspective, clearly formulating the main challenges that need to be addressed while highlighting possible future perspectives in this field. In addition, it describes a broad and comprehensive framework for capturing the multitude of geometric structures and their dynamic behavior at both cellular and subcellular levels. It also emphasizes the need to build scalable algorithms, given the ever-increasing size of datasets. These datasets pose a number of hallenges that are very distinct from conventional clinical imagery in their size and abundance, the detail of relevant features, and their statistics.

The goal of the FARSIGHT concept, developed by Badri Roysam, is the formulation of a unified approach that captures the richness of such data sets. Chapter 5 presents an overview of this approach and illustrates how computational models can be applied to establish associations among cellular compartments, surfaces, and functional markers.

### Part II: Subcellular Structures and Events

The first group of chapters higlighting the application of advance image analysis tools focuses on the analysis and measurement of subcellular structures and events. In Chapter 6, Hanchuan Peng addresses the analysis of in situ gene expression pattern images obtained from the use of specialized mRNA probes. The goal

is to utilize pattern recognition as an important source of information towards understanding the functions of genes required for the development of various organisms. Automated image analysis methods are applied to build image informatics tools that analyze the spatial-temporal patterns of gene expressions during embryogenesis of Drosophila melangaster (fruit fly). The end results of the analysis are appropriate ontological annotations of expression patterns.

Techniques for probing intracellular dynamics in living cells using optical methods are still very crude. The large number of small moving particles complexity of particle interactions pose significant challenges. In Chapter 7 Jérôme Boulanger and his associates present a general estimation and simulation framework that allows the modeling of image sequences depicting the motion of small regions as complex motion patterns that, in turn, are manifestations of intracellular dynamics and trafficking in biology.

Immunofluorescent methods that allow the staining of a single tissue section with an entire set of different protein biomarkers are beginning to emerge. In Chapter 8, Ali Can and his associates review how such a new staining technique enables multiplexing a large number of markers in a sigle tissue. They illustrate how specialized image analysis methods enable the quantification of the expression levels of different markers. By extractin the expressions in different subcellular compartments, these techniques allow one to gain insight into unexamined relationships between spatial locations an various protein-protein interactions. Preliminary data certainly demonstrates the value of automated fluorescence-based image analysis when used on clinical studies of breast and lung cancers.

Supporting large screenings of RNA interference experiments and small molecule screenings is the focus of Chapter 9. Thouis R. Jones and his associates present methods for automatic image cytometry. Developing methods for illumination normalization, foreground/background separation, and cell segmentation that work robustly on thousands of images acquired with a high-throughput robotic microscope is challenging. The authors describe how the CellProfiler system, which is available as open source software, can be applied for making discoveries about cellular processes, genetic patways, and drug candidates.

### Part III: Structure and Dynamics of Cell Populations

Computerized video time-lapse microscopy enables the monitoring of cell populations over extended periods of time. While some of the introductory chapters have already touched on the subject, Chapters 10 through 13 focus on the analyses of the dynamics of cell populations and cellular functions. In Chapter 10, Auguste Genovesio and Jean-Christophe Olivo-Marin present a detailed introduction to various stochastic methods for cell tracking. Specific examples demontstrate how certain algorithms work thus highlighting their advantages. They demonstrate how these visual tracking methods can be used to automatically track biological particulate material as manifest in 3-D+time sequences. Analyzing such datasets is a major challenge and constitutes a major bottleneck for the full exploitation of multidimensional microscopy sequences that documents studies of biological object dynamics.

Chapters 11 and 12 are closely related as both author groups study the cell cycle of a large cell population on a single cell level. Nathalie Harder and her associates introduce in Chapter 11 a computational scheme to automatically segment, track, and classify cell nuclei into different mitotic phases. This work is carried out within the European Union project MitoCheck, which aims to explore the coordination of mitotic processes in human cells at a molecular level and to contribute towards the revealing of mechanisms of cellular development. The results demonstrate that the proposed analysis system reliably classifies cells into the seven mitotic phases.

Dirk Padfield and collaborators expand on the topic in Chapter 12 by applying automatic image analysis algorithms to extract information on the cell cycle progression through the interphase, which is divided into G1, S, and G2 phases, of each cell using a cell cycle phase marker. Among other applications, this is necessary for studying the effect of inhibitor compounds that are designed to block the replication of cancerous cells. Hence, this technique holds the promise of allowing a more detailed study of the distribution of cell cycle checkpoints and perturbations. Label-free imaging of cellular systems is an important topic and highly relevant for many practical experiments. Many molecular markers that are used as nuclei stains have toxic side effects and consequently perturb the living cells. In Chapter 13, Xiaoxu Wang and his colleagues present algorithms that enable the automatic segmentation and cell counting in phase contrast images. Due to the low contrast that exists in the image intensity across membrane boundaries, this is a very challenging task. More importantly, the proposed methods show how computerized image analysis techniques can eliminate the need for fluorescent markers.

### Part IV: Automated Cellular and Tissue Analysis

Going beyond analyzing individual cells, Chapters 14 through 17 present applications that relate to the structure and organization of tissues. Chapter 14 is the presentation of the Digital Fish Project written by Sean G. Megason. Rather than presenting particular algorithmic approach, this chapter highlights how imaging can play a crucial role in this study of systems biology. The goal is to understand how multiple components of a biological system interact to give rise to the function of that system. These components can include metabolites, genese proteins, cells, and even whole organisms. Because imaging can provide anatomical landmarks as well as data that is single cell, longitudinal, and quantitative, it can provide salient and key insights that will expand our current understanding of such systems. The subject of the Digital Fish Project is a tropical fish called zebrafish (Danio rerio). The embryos and larvae of such fish are transparent, allowing te entire embryo to be imaged with in too optical microscopy. The goal of in toto imaging is to image every single cell in a tissue and eventually image the entire embryo as it develops over time. Chapter 14 promotes the concept of open data and interested researchers can access the data and contribute to its analysis.

The main goal of the work presented in Chapter 15 by Kun Huang and his collaborators is the 3-D reconstruction of tissue sections, for example, mouse placenta organ, from a set of histology sections. As it is not possible to acquire such

datasets with confocal microscopy, the challenge of sparse sampling needs to be overcome to reconstruct and analyze such datasets. New algorithms for image segmentation, registration, visualization, and quantization are necessary to address such challenges. This work is part of a mouse model phenotyping, which aims to study the roles of tumor suppressor genes iin the regulation of normal cellular processes. This chapter is another example how the increased use of both histology and tagged confocal images and subsequent analyses provide more insight into fundamental mechanisms of developmental and cancer biology. Chapters 16 and 17 are closely related since both author groups study neural morphology which is broadly affected by age, genetic diseases such as Down's Syndrome, and degenerative diseases such as Alzheimer's disease. In Chapter 16, Ioannis A. Kakadiaris and his collaborators present a fully automatic system for the reconstruction of neuronal morphologies from multiphoton microscopy data. The goal of their work is to enable morphological-guided functional imaging and to create an entire library of such neuron morphologies. The proposed methods are general and make no prior assumptions about the shape of tubular structures.

Firdaus Janoos and his colleagues describe the 3-D reconstruction and classification of dendritic spines. The emphasis is on the use of computer graphics techniques that are best used on surface meshes. The advantage of such methods is that they are amenable to verification and validation. The data is obtained from a confocal microscope.

## Part V: In Vivo Microscopy

In order to understand the full complexity of biological systems it is desirable to study such systems in vivo and in situ. Additionally, it is necessary to establish and explain the images that are acquired on a microscopic scale in the context of higher-resolution images that are typically acquired using magnetic resonance imaging, ultrasound, or computed tomography (CT). Both of these requirements pose significant challenges for images acquisition and are the subject of ongoing and future research. Chapters 18 and 19 present examples of such efforts.

In Chapter 18, Zheng Zia and associates presents an approach to in vivo molecular small animal imaging. Small animals are ofter used as surrogates for humans in the study of normal and disease states. As it is now possible to introduce genetic mutations identical to those commonly found in human cancer tissue into the endogenous murine gene locus, the physiological relevance of such small animal models has dramatically improved. This chapter introduces optical imaging methods for small animal studies and describes a new multimodality fusion method that combines 3-D fluorescence moleculer tomography and CT images to improve the overall information content.

The principle of fibered confocal microscopy is introduced in Chapter 19. This new imaging technology raises several new image processing and image analysis challenges that cannot readily be addressed by using classical approaches. Tom Vercauteren and colleagues present several tools that increase the relevance and utility of such datasets. One of these tools enables, for example, the analysis of the behavior of blood cells and vessels in situ. By applying image sequence mosaicking,

they effectively widen the field of view of the system and bridge the gap between the microscopic and macroscopic scales.

This field is, as Robert F. Murphy concludes in the foreword to this book, still in its infancy. The work presented here demonstrates that biomedical image analysis can enhance our understanding of complex biological processes and behaviors. This book project was initiated by a series of workshops, Microscopic Image Analysis with Applications in Biology (MIAAB). As this is a new and rapidly developing field, there are naturally a number of forums that focus on various aspects of microscopic image analysis and its applications. However, it has become abundantly clear that strong interdisciplinary collaborations are necessary to exploit the potential of image analysis methods. We are already seeing examples of discoveries that otherwise would not have been possible. While algorithm developers need to learn how certain specific application-related challenges can be addressed, it is imperative that researchers in the life sciences develop a good understanding of what advantages these technologies can offer. The goal of this book is to facilitate this exchange and contribute to an increased collaboration between these fields. The presentation of the applications is certainly not complete, and this volume is a presentation of selected examples as opposed to an encyclopedic review of the state of the art, but we hope that this book contributes to achieving this goal.

## Acknowledgments

We would like to express our gratitude to Dimitri Metaxas (Rutgers University), Ross Whitaker (University of Utah), and Stephen Lockett and Terry Yoo (both of NIH) as they have supported the MIAAB workshops in many ways. We are particularly grateful that Wayne Yuhasz, executive acquisitions editor at Artech House, strongly encouraged this book project. His colleagues, Barbara Lovenvirth and Rebecca Allendorf, were instrumental in guiding us to translate the plan into reality. It should be duly noted that, without the contributions of the many authors, we could not have compiled this volume. We owe thanks to all of them for their patience and timely interaction. Finally, we owe thanks to our individual families for letting us spend the time to complete the task of compiling this volume

*Jens Rittscher*
*Raghu Machiraju*
*Stephen T. C. Wong*
*Editors*
*July 2008*

# Contents