AUTOMATIC DIATOM IDENTIFICATION

SERIES IN MACHINE PERCEPTION AND ARTIFICIAL INTELLIGENCE*

Editors: **H. Bunke** (Univ. Bern, Switzerland)
P. S. P. Wang (Northeastern Univ., USA)

- Vol. 34: *Advances in Handwriting Recognition*
(Ed. *S.-W. Lee*)
- Vol. 35: *Vision Interface — Real World Applications of Computer Vision*
(Eds. *M. Cheriet and Y.-H. Yang*)
- Vol. 36: *Wavelet Theory and Its Application to Pattern Recognition*
(*Y. Y. Tang, L. H. Yang, J. Liu and H. Ma*)
- Vol. 37: *Image Processing for the Food Industry*
(*E. R. Davies*)
- Vol. 38: *New Approaches to Fuzzy Modeling and Control — Design and Analysis*
(*M. Margaliot and G. Langholz*)
- Vol. 39: *Artificial Intelligence Techniques in Breast Cancer Diagnosis and Prognosis*
(Eds. *A. Jain, A. Jain, S. Jain and L. Jain*)
- Vol. 40: *Texture Analysis in Machine Vision*
(Ed. *M. K. Pietikäinen*)
- Vol. 41: *Neuro-Fuzzy Pattern Recognition*
(Eds. *H. Bunke and A. Kandel*)
- Vol. 42: *Invariants for Pattern Recognition and Classification*
(Ed. *M. A. Rodrigues*)
- Vol. 43: *Agent Engineering*
(Eds. *Jiming Liu, Ning Zhong, Yuan Y. Tang and Patrick S. P. Wang*)
- Vol. 44: *Multispectral Image Processing and Pattern Recognition*
(Eds. *J. Shen, P. S. P. Wang and T. Zhang*)
- Vol. 45: *Hidden Markov Models: Applications in Computer Vision*
(Eds. *H. Bunke and T. Caelli*)
- Vol. 46: *Syntactic Pattern Recognition for Seismic Oil Exploration*
(*K. Y. Huang*)
- Vol. 47: *Hybrid Methods in Pattern Recognition*
(Eds. *H. Bunke and A. Kandel*)
- Vol. 48: *Multimodal Interface for Human-Machine Communications*
(Eds. *P. C. Yuen, Y. Y. Tang and P. S. P. Wang*)
- Vol. 49: *Neural Networks and Systolic Array Design*
(Eds. *D. Zhang and S. K. Pal*)

*For the complete list of titles in this series, please write to the Publisher.

AUTOMATIC DIATOM IDENTIFICATION

Editors

Hans du Buf

University of Algarve, Portugal

Micha M. Bayer

Royal Botanic Garden Edinburgh, UK

World Scientific

New Jersey • London • Singapore • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

AUTOMATIC DIATOM IDENTIFICATION

Copyright © 2002 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-02-4886-5

Printed in Singapore by Uto-Print

Preface

Computer processing of diatom images used to be an area reserved for a very few specialists. With the advent of the ADIAC project (Automatic Diatom Identification And Classification), funded by the European Marine Science And Technology (MAST) programme, we may say that diatom research has entered a new era. In the future, computers could assist diatom researchers in many different ways: on-line electronic atlases could largely replace printed ones, automatic slide scanning could free diatomists from routine microscopy tasks, and the identification of taxa could be carried out automatically or semi-automatically.

This book is actually the final scientific report of ADIAC, the first project ever to attempt diatom identification on a large scale by means of image analysis, using both contour and valve-face features. It also allowed the creation of large image databases, and explored the automatic scanning of microscope slides.

ADIAC was designed as a pilot study, that is, it was intended to explore the subject of automatic diatom identification and to pinpoint problems. Success cannot be bought with research grants, especially when a relatively new area is being explored, when there are only three years available to do the work, and when much of it is done by fresh PhD students. Now, at the end of the project, it is fair to say that our results have exceeded all expectations, and that the project's participants have done a magnificent job. However, much remains to be done before our techniques can be applied in routine practice. We hope that investigators and practitioners in both diatom research and pattern recognition will find this book useful, as it points the way to future research.

Hans du Buf and Micha Bayer

For public databases and online demo see:

<http://www.ualg.pt/adiac/>

<http://www.rbge.org.uk/ADIAC/>

<http://www.rbge.org.uk/ADIAC/db/adiacdb.htm>

<http://www.cs.rug.nl/hpci/demos/adiac/>

This page is intentionally left blank

Acknowledgments

The ADIAC project was funded by the European MAST (MARine Science and Technology) programme, contract MAS3-CT97-0122. Additional funding came from the Swiss Federal Office for Education and Science (BBW 98.00.48), the CSIC in Spain and the Spanish programmes III Pricit CAM as well as MAR98-1537CE.

We would also like to thank Annemarie Clarke, Richie Head, Mary Lewis and Kaarina Weckström (Univ. of Newcastle) for providing diatom slides and for their expertise in selecting, identifying and scanning a substantial number diatom specimens for the image databases.

We would like to thank Stephen Droop, formerly Royal Botanic Garden Edinburgh, for his help in planning and developing the ADIAC project in the early stages, and for his participation in the first half of the project. We are also grateful to Dr. Martin Pullan, Royal Botanic Garden Edinburgh, for help in developing the ADIAC diatom database and the web-based search engine, and for continued advice and support in data management issues.

David Mann is grateful to the Royal Society for an equipment grant enabling the purchase of a photomicroscope.

The authors of Chapter 5 would like to thank all the participants that took part in the ring-test and the Internet-based experiments, and so generously gave up much of their precious time; without their contribution these valuable data could not have been generated. Many thanks also to AWEL (Amt für Abfall, Wasser, Energie und Luft), Kanton Zürich (Switzerland), for supporting the ring-test, to Dr. Pius Niederhauser (AWEL) who performed field and laboratory work, and to Kathryn Lyttle (Univ. of Newcastle) who kindly helped analyze the data from the Internet-based experiments.

Gabriel Cristóbal would like to thank J. Chamorro, J. Fernandez-Valdivia, B. Prados and D. Alvarez of Univ. Granada (Spain), and J. I. Uzquiano of CSIC (Spain) for their contributions and support of the work described in Chapter 13.

The editors also gratefully acknowledge the series editor Horst Bunke, who encouraged us to prepare this book, the editors of World Scientific for their valuable help, as well as Luís Santos for L^AT_EXing Chapters 2 to 5. Last but not least, to all authors: thanks for a timely production.

This page is intentionally left blank

Authors' affiliations

Authors for correspondence are marked with a †.

Micha M. Bayer, David G. Mann[†]

Royal Botanic Garden Edinburgh (RBGE)
20A Inverleith Row, Edinburgh EH3 5LR, Scotland, UK

Richard J. Telford, Steve Juggins[†]

University of Newcastle-upon-Tyne, Dept. of Geography
Daysh Building, Claremont Road, Newcastle-upon-Tyne
Tyne and Wear NE1 7RU, UK

Bertrand Ludes

Université Louis Pasteur, Inst. de Médecine Légale
11 rue Humann, F-67085, Strasbourg Cedex, France

Michael H.F. Wilkinson, Andrei C. Jalba, Erik R. Urbach,

Michel A. Westenberg, Jos B.T.M. Roerdink[†]

Inst. for Mathematics and Computing Science
P.O. Box 800, 9700 AV Groningen, The Netherlands

Stefan Fischer, Horst Bunke[†]

Universität Bern, Inst. für Informatik und angewandte Mathematik
Neubrückstrasse 10, CH-3012 Bern, Switzerland

Jose L. Pech-Pacheco, Gabriel Cristóbal[†]

Instituto de Optica – CSIC
Serrano 121, 28006 Madrid, Spain

Robert E. Loke, Adrian Ciobanu, Luís M. Santos,

Hamid R. Shahbazkia, Hans du Buf[†]

Vision Laboratory, Faculty of Sciences and Technology
University of Algarve, Campus de Gambelas, 8000-810 Faro, Portugal

Martyn G. Kelly

Bowburn Consultancy
11 Monteigne Drive, Bowburn, Durham DH6 5QB, UK

Joachim Hürlimann

AquaPlus
Bundesstrasse 6, CH-6300 Zug, Switzerland

This page is intentionally left blank

Contents

Preface	v
Acknowledgments	vii
Authors' Affiliations	ix
1 Introduction to ADIAC and This Book	1
<i>Hans du Buf and Micha M. Bayer</i>	
2 Diatoms: Organism and Image	9
<i>David G. Mann</i>	
3 Diatom Applications	41
<i>Richard J. Telford, Steve Juggins, Martyn G. Kelly, and Bertrand Ludes</i>	
4 ADIAC Imaging Techniques and Databases	55
<i>Micha M. Bayer and Steve Juggins</i>	
5 Human Error and Quality Assurance in Diatom Analysis	75
<i>Martyn G. Kelly, Micha M. Bayer, Joachim Hürlimann, and Richard J. Telford</i>	
6 Contour Extraction	93
<i>Stefan Fischer, Hamid R. Shahbazkia, and Horst Bunke</i>	
7 Identification Using Classical and New Features in Combination with Decision Tree Ensembles	109
<i>Stefan Fischer and Horst Bunke</i>	
8 Identification by Curvature of Convex and Concave Segments	141
<i>Robert E. Loke and Hans du Buf</i>	
9 Identification by Contour Profiling and Legendre Polynomials	167
<i>Adrian Ciobanu and Hans du Buf</i>	
10 Identification by Gabor Features	187
<i>Luís M. Santos and Hans du Buf</i>	
11 Identification by Mathematical Morphology	221
<i>Michael H. F. Wilkinson, Andrei C. Jalba, Erik R. Urbach, and Jos B. T. M. Roerdink</i>	
12 Mixed-Method Identifications	245
<i>Michel A. Westenberg and Jos B. T. M. Roerdink</i>	
13 Automatic Slide Scanning	259
<i>José L. Pech-Pacheco and Gabriel Cristóbal</i>	
14 ADIAC Achievements and Future Work	289
<i>Hans du Buf and Micha M. Bayer</i>	
Appendix: The Mixed Genera Data Set	299