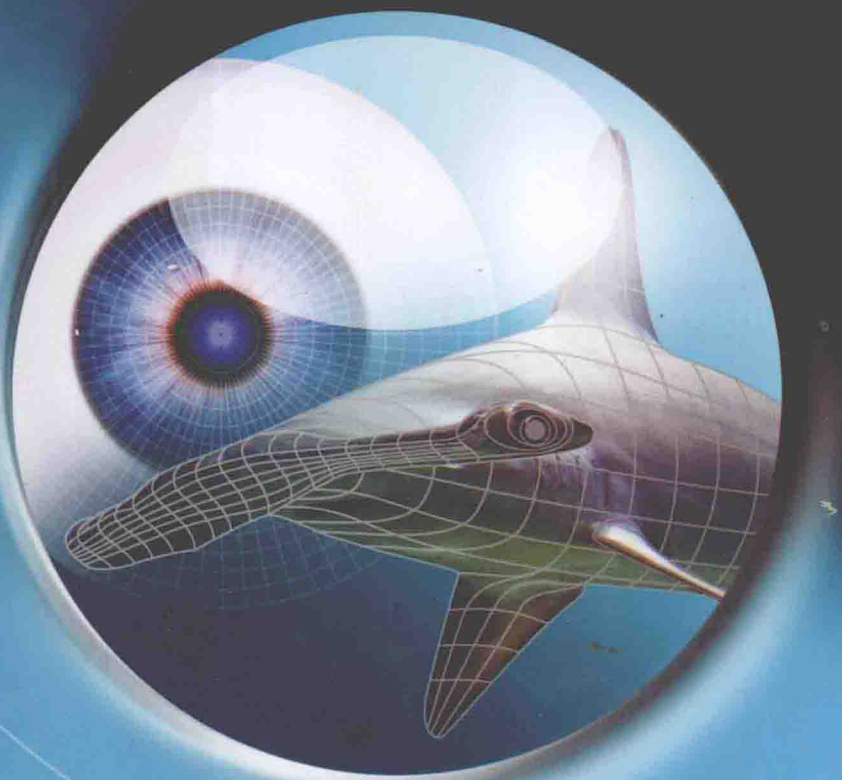
An Introduction to 3D Computer Vision Techniques and Algorithms



Bogusław Cyganek
J. Paul Siebert

WILEY

AN INTRODUCTION TO 3D COMPUTER VISION TECHNIQUES AND ALGORITHMS

Bogusław Cyganek
Department of Electronics, AGH University of Science and Technology, Poland

J. Paul Siebert
Department of Computing Science, University of Glasgow, Scotland, UK

WILEY
A John Wiley and Sons, Ltd., Publication

This edition first published 2009

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

Cyganek, Boguslaw.
An introduction to 3D computer vision techniques and algorithms / by Boguslaw Cyganek and J. Paul Siebert.
p. cm.
Includes index.
ISBN 978-0-470-01704-3 (cloth)
1. Computer vision. 2. Three-dimensional imaging. 3. Computer algorithms. I. Siebert, J. Paul. II. Title
TA1634.C94 2008
006.3′7–dc22

2008032205

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

ISBN 978-0-470-01704-3 (H/B)

Set in 10/12pt Times by Aptara Inc., New Delhi, India.
Printed in Great Britain by CPI Antony Rowe, Chippenham, Wiltshire

AN INTRODUCTION TO 3D COMPUTER VISION TECHNIQUES AND ALGORITHMS

To Magda, Nadia and Kamil
From Bogusław

To Sabina, Konrad and Gustav
From Paul

Preface

Recent decades have seen rapidly growing research in many areas of computer science, including computer vision. This comes from the natural interest of researchers as well as demands from industry and society for qualitatively new features to be afforded by computers. One especially desirable capability would be automatic reconstruction and analysis of the surrounding 3D environment and recognition of objects in that space. Effective 3D computer vision methods and implementations would open new possibilities such as automatic navigation of robots and vehicles, scene surveillance and monitoring (which allows automatic recognition of unexpected behaviour of people or other objects, such as cars in everyday traffic), medical reasoning, remote surgery and many, many more.

This book is a result of our long fascination with computers and vision algorithms. It started many years ago as a set of short notes with the only purpose 'to remember this or that' or to have a kind of 'short reference' just for ourselves. However, as this diary grew with the years we decided to make it available to other people. We hope that it was a good decision! It is our hope that this book facilitates access to this enthralling area, especially for students and young researchers. Our intention is to provide a very concise, though as far as possible complete, overview of the basic concepts of 2D and 3D computer vision. However, the best way to get into the field is to try it oneself! Therefore, in parallel with explaining basic concepts, we provide also a basic programming framework with the hope of making this process easier. We greatly encourage the reader to take the next step and try the techniques in practice.

Bogusław Cyganek, Kraków, Poland
J. Paul Siebert, Glasgow, UK

Acknowledgements

We would like to express our gratitude to all the people who helped in the preparation of this book!

In particular, we are indebted to the whole Wiley team who helped in the preparation of the manuscript. In this group special thanks go to Simone Taylor who believed in this project and made it happen. We would also like to express our gratitude to Sian Andrews, Laura Bell, Liz Benson, Emily Bone, Lucy Bryan, Kate Griffiths, Wendy Hunter, Alex King, Erica Peters, Kathryn Sharples, and Nicky Skinner.

We are also very grateful to the individuals and organizations who agreed to the use of their figures in the book. These are Professor Yuichi Ohta from Tsukuba University, as well as Professor Ryszard Szeliski from Microsoft Research. Likewise we would like to thank Dimensional Imaging Ltd. and Precision 3D Ltd. for use of their images. In this respect we would also like to express our gratitude to Springer Science and Business Media, IEEE Computer Society Press, the IET, Emerald Publishing, the ACM, Maney Publishing and Elsevier Science.

We would also like to thank numerous colleagues from the AGH University of Science and Technology in Kraków. We owe a special debt of gratitude to Professor Ryszard Tadeusiewicz and Professor Kazimierz Wiatr, as well as to Lidia Krawentek for their encouragement and continuous support.

We would also like to thank members of the former Turing Institute in Glasgow (Dr Tim Niblett, Joseph Jin, Dr Peter Mowforth, Dr Colin Urquhart and also Arthur van Hoff) as well as members of the Computer Vision and Graphics Group in the Department of Computing Science, University of Glasgow, for access to and use of their research material (Dr John Patterson, Dr Paul Cockshott, Dr Xiangyang Ju, Dr Yijun Xiao, Dr Zhili Mao, Dr Zhifang Mao (posthumously), Dr J.C Nebel, Dr Tim Boyling, Janet Bowman, Susanne Oehler, Stephen Marshall, Don Whiteford and Colin McLaren). Similarly we would like to thank our collaborators in the Glasgow Dental Hospital and School (Professor Khursheed Moos, Professor Ashraf Ayoub and Dr Balvinder Khambay), Canniesburn Plastic Surgery Unit (Mr Arup Ray), Glasgow, the Department of Statistics (Professor Adrian Bowman and Dr Mitchum Bock), Glasgow University, Professor Donald Hadley, Institute of Neurological Sciences, Southern General Hospital, Glasgow, and also those colleagues formerly at the Silsoe Research Institute (Dr Robin Tillett, Dr Nigel McFarlane and Dr Jerry Wu), Silsoe, UK.

Special thanks are due to Dr Sumitha Balasuriya for use of his Matlab codes and graphs. Particular thanks are due to Professor "Keith" van Rijsbergen and Professor Ray Welland without whose support much of the applied research we report would not have been possible.

We wish to express our special thanks and gratitude to Steve Brett from Pandora Inc. for granting rights to access their software platform.

Some parts of the research for which results are provided in this book were possible due to financial support of the European Commission under RACINE-S (IST-2001-37117) and IP-RACINE (IST-2-511316-IP) as well as Polish funds for scientific research in 2007–2008. Research described in these pages has also been funded by the UK DTI and the EPSRC & BBSRC funding councils, the Chief Scientist Office (Scotland), Wellcome Trust, Smith's Charity, the Cleft Lip and Palate Association, the National Lottery (UK) and the Scottish Office. Their support is greatly appreciated.

Finally, we would like to thank Magda and Sabina for their encouragement, patience and understanding over the three-year period it took to write this book.

Notation and Abbreviations

$I_k(x, y)$	Intensity value of a k-th image at a point with local image coordinates (x, y)
$\overline{I_k(x, y)}$	Average intensity value of a k-th image at a point with local image coordinates (x, y)
$\mathbf{I}$	Identity matrix; image treated as a matrix
$\mathbf{P}$	A vector (a point), matrix, tensor, etc.
$T[\mathbf{I}, \mathbf{P}]$	The Census transformation T for a pixel $\mathbf{P}$ in the image $\mathbf{I}$
i, j	Free coordinates
d_x, d_y	Displacements (offset) in the x and y directions
$D(\mathbf{p}_l, \mathbf{p}_r)$	Disparity between points $\mathbf{p}_l$ and $\mathbf{p}_r$
$\mathbf{D}$	Disparity map (a matrix)
$U(x, y)$	Local neighbourhood of pixels around a point (x, y)
$\mathbf{O}_c$	Optical centre point
$\mathbf{P}_c = [X_c, Y_c, Z_c]^T$	Coordinates of a 3D point in the camera coordinate system
$\mathbf{\Pi}$	Camera plane; a projective plane
$\mathbf{o} = (o_x, o_y)$	Central point of a camera plane
f	Focus length of a camera
b	Base line in a stereo system (a distance between cameras)
h_x, h_y	Physical horizontal and vertical dimensions of a pixel
$\mathbf{P} = [X, Y, Z]^T$	3D point and its coordinates
$\wp^n$	N-dimensional projective space
$\mathbf{P} = [X_h, Y_h, Z_h, 1]^T$	Homogenous coordinates of a point
$\mathbf{M}$	Camera matrix
$\mathbf{M}_\mathrm{i}$	Intrinsic parameters of a camera
$\mathbf{M}_\mathrm{e}$	Extrinsic parameters of a camera
$\mathbf{E}$	Essential matrix.
$\mathbf{F}$	Fundamental matrix.
$\mathbf{e}_i$	Epipole in an i-th image
SAD	Sum of absolute differences
SSD	Sum of squared differences
ZSAD	Zero-mean sum of absolute differences
ZSSD	Zero-mean sum of squared differences
ZSSD-N	Zero-mean sum of squared differences, normalized

SCP	Sum of cross products
SCP-N	Sum of cross products, normalized
RMS	Root mean square
RMSE	Root mean square error
<Lxx, Lyy>	Code lines from a line *Lxx* to *Lyy*
HVS	Human Visual System
SDK	Software Development Kit
$\wedge$	logical 'and'
$\vee$	logical 'or'
LRC	Left-right checking (cross-checking)
OCC	Occlusion constraint
ORD	Point ordering constraint
BMD	Bimodality rule
MGJ	Match goodness jumps
NM	Null method
GT RMS	Ground-truth RMS
WTA	Winner-takes-all
*	Convolution operator

Contents