

Advances in Pattern Recognition

Springer-Verlag London Ltd.

Advances in Pattern Recognition is a series of books which brings together current developments in all areas of this multi-disciplinary topic. It covers both theoretical and applied aspects of pattern recognition, and provides texts for students and senior researchers.

Springer also publishes a related journal, **Pattern Analysis and Applications**. For more details see: <http://link.springer.de>

The book series and journal are both edited by Professor Sameer Singh of Exeter University, UK.

Also in this series:

Principles of Visual Information Retrieval

Michael S. Lew (Ed.)

1-85233-381-2

Statistical and Neural Classifiers: An Integrated Approach to Design

Šarūnas Raudys

1-85233-297-2

Advanced Algorithmic Approaches to Medical Image Segmentation

Jasjit Suri, Kamaledin Setarehdan and Sameer Singh (Eds)

1-85233-389-8

NETLAB: Algorithms for Pattern Recognition

Ian T. Nabney

1-85233-440-1

M. Bennamoun and G.J. Mamic

Object Recognition

Fundamentals and Case Studies

With 133 Figures



M. Bennamoun Ing, MSc, PhD
G.J. Mamic, BEng, PhD (pending)
Space Centre for Satellite Navigation,
School of Electrical and Electronic Systems Engineering,
Queensland University of Technology, Brisbane, Queensland 4001, Australia

Series editor

Professor Sameer Singh, PhD
Department of Computer Science, University of Exeter, Exeter, EX4 4PT, UK

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

Library of Congress Cataloging-in-Publication Data
Bennamoun, M. (Mohammed)

Object recognition : fundamentals and case studies / Mohammed Bennamoun
and George Mamic.

p. cm. – (Advances in pattern recognition, ISSN 1617-7916)

Includes bibliographical references and index.

ISBN 978-1-4471-3724-5

ISBN 978-1-4471-3722-1 (eBook)

DOI 10.1007/978-1-4471-3722-1

1. Image processing. 2. Pattern recognition systems. I. Mamic, George, 1977
III. Title. IV. Series.

TA1632 .B45 2002

621.367—dc21

2001049625

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

Advances in Pattern Recognition ISSN 1617-7916

ISBN 978-1-4471-3724-5

© Springer-Verlag London 2002

Originally published by Springer-Verlag London Berlin Heidelberg in 2002

Softcover reprint of the hardcover 1st edition 2002

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant laws and regulations and therefore free for general use.

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Typesetting: Camera ready by authors

34/3830-543210 Printed on acid-free paper SPIN 10785377

To my father whose image is always before my eyes and whose values run my life. To my mother, my wife and my kids whose presence is as precious as life.

M. Bennamoun

To my parents whose unconditional love, intelligence, compassion and tenacity are a source of awe and inspiration.

G.J. Mamic

Preface

Automatic object recognition is a multidisciplinary research area using concepts and tools from mathematics, computing, optics, psychology, pattern recognition, artificial intelligence and various other disciplines. The purpose of this research is to provide a set of coherent paradigms and algorithms for the purpose of designing systems that will ultimately emulate the functions performed by the Human Visual System (HVS). Hence, such systems should have the ability to recognise objects in two or three dimensions independently of their positions, orientations or scales in the image.

The HVS is employed for tens of thousands of recognition events each day, ranging from navigation (through the recognition of landmarks or signs), right through to communication (through the recognition of characters or people themselves). Hence, the motivations behind the construction of recognition systems, which have the ability to function in the real world, is unquestionable and would serve industrial (e.g. quality control), military (e.g. automatic target recognition) and community needs (e.g. aiding the visually impaired).

Scope, Content and Organisation of this Book

This book provides a comprehensive, yet readable foundation to the field of object recognition from which research may be initiated or guided. It represents the culmination of research topics that I have either covered personally or in conjunction with my PhD students. These areas include image acquisition, 3-D object reconstruction, object modelling, and the matching of objects, all of which are essential in the construction of an object recognition system.

The book is divided into three parts. Part A provides the reader with an introduction into object recognition and covers the acquisition of images. Part B details the 3-D object reconstruction, modelling and matching, and forms the fulcrum of the manuscript. Finally, Part C describes typical recognition systems using case studies, which illustrates the application of the fundamentals covered in the previous parts. The code used to generate the results that are presented in the case studies may be found at the following FTP site: ftp://ftp.springer.de/pub/cs/object_recognition/. The reader is given a road map through all of this literature in the form of a taxonomy that is

constructed in the introduction. In addition, each chapter in Parts A and B of the manuscript contains an extensive literature survey of state-of-the-art systems available in specific areas. These surveys are summarised in detailed tables and taxonomies at the beginning of each of the chapters.

Who Should Read this Book?

The book should appeal to both engineers and computer scientists, in particular advanced undergraduate students, postgraduate students and researchers. Hopefully, researchers in other disciplines, such as psychology, will also find the manuscript useful.

Acknowledgements

First I would like to thank my co-author George Mamic for his patience, hard work, perseverance, contribution and support. Particular thanks to my past students whose research has contributed enormously to this manuscript, namely, Jasmine Banks, John Williams, and Anthony Cheung. Thanks to the series editor S. Singh for encouraging us to write the manuscript and to our reviewers for their valuable comments and constructive inputs. Also, special thanks to my past and current students, namely (in alphabetical order): Jason Baker, Adriana Bodnarova, PiChi Chou, Clinton Fookes, Greg Hislop, Haris Pandzo, Birgit Planitz, Alex Pongpech and Jeremy Tay. I would also like to acknowledge the support of the Queensland University of Technology for all the research support that I have received over the years. Thanks to Professor Kurt Kubik for his numerous technical discussions and continual support. Thanks also to Professors M. Bayoumi, N. Bergmann and A. Zoubir for their collaboration. Acknowledgements also go to the Australian Research Council (ARC) for their grant support of some projects directly related to the area of object recognition. Finally, I would like to thank the “Laboratoire de Traitement et Transport de l’Information”, L2TI in Paris 13, France; particularly Professor A. Beghdadi for hosting me whilst the finishing touches were being applied to this manuscript.

Mohammed Bennamoun
July 2001

Contents

Part A – Introduction and Acquisition Systems	1
1. Introduction	3
1.1 What Is Computer Vision?	3
1.2 Background and History	4
1.3 Classification of Existing Vision Systems	5
1.3.1 Marr’s Theory	6
1.3.2 Model-based Object Recognition	8
1.4 Problem Formulation	9
1.4.1 Mathematical Formulation	9
1.5 Why Is Automatic Object Recognition a Difficult Problem?	12
1.6 Motivations and Significance	13
1.6.1 Industry	13
1.6.2 Community	14
1.7 A 2-D System or a 3-D System?	14
1.8 Specifications/Themes of Interest in Object Recognition	15
1.9 Acquisition Systems	17
1.9.1 Intensity Images	18
1.9.2 Range Imaging Technologies	19
1.9.3 Miscellaneous Modalities	27
1.10 Taxonomy	27
2. Stereo Matching and Reconstruction of a Depth Map	29
2.1 Fundamentals of Stereo Vision	29
2.1.1 Stereo Vision Paradigm	30
2.1.2 Image Matching	31
2.1.3 Matching Problems	31
2.2 Review of Existing Techniques	33
2.3 Area-based Techniques	33
2.3.1 Simple Matching Measures	34
2.3.2 Validation Techniques	35
2.3.3 Hierarchical Methods	38
2.3.4 Adaptive Window Techniques	38

2.3.5	Sparse Point Matching	41
2.3.6	Dense Matching	42
2.3.7	Symmetric Multi-Window Technique	45
2.3.8	Unmanned Ground Vehicle Implementation	46
2.3.9	Multiple Baseline Techniques	47
2.3.10	Least Squares Matching	50
2.4	Transform-based Techniques	53
2.4.1	Sign Representation	54
2.4.2	Non-parametric Techniques	58
2.5	Symbolic Feature-based Techniques	61
2.5.1	Zero Crossing Matching	62
2.5.2	Edge Matching	62
2.5.3	Patch Matching	73
2.5.4	Relational Matching	74
2.6	Hybrid Techniques	74
2.6.1	Cross Correlation Combined with Edge Information ..	74
2.7	Phase-based Techniques	78
2.8	Combining Independent Measurements	79
2.9	Relaxation Techniques	80
2.9.1	Cooperative Algorithm	80
2.9.2	Relaxation Labelling	81
2.10	Dynamic Programming	81
2.10.1	Viterbi Algorithm	82
2.10.2	Intra- and Inter-Scanline Search	83
2.10.3	Disparity Space Image	86
2.11	Object Space Techniques	87
2.11.1	Combining Matching and Surface Reconstruction	87
2.11.2	Object Space Models	88
2.12	Existing Matching Constraints and Diagnostics	89
2.12.1	Matching Constraints	89
2.12.2	Matching Diagnostics	92
2.12.3	Discussion	92
2.13	Conclusions	95
Part A – Summary		97
Part B – Database Creation and Modelling for 3-D Object Recognition		99
3.	3-D Object Creation for Recognition	101
3.1	Preliminaries of 3-D Registration	102
3.2	Registration Paradigm	103
3.2.1	General Specifications	104
3.3	Chronological Literature Review	104
3.4	Fundamental Techniques	107

3.4.1	Registration with Point Correspondences	108
3.4.2	Registration Without Correspondences	112
3.5	Uncertainty in 3-D Registration	117
3.5.1	Weighted Correspondences	118
3.5.2	A Better Approach	119
3.6	Simultaneous Multiple View Registration	122
3.6.1	Simple Approaches	122
3.6.2	Rigid Body Modelling	122
3.6.3	Multiple View Chen and Medioni	123
3.7	View Integration and Surface Reconstruction	125
3.7.1	Integration versus Reconstruction	126
3.7.2	Volumetric Integration Methods	127
3.7.3	Volumetric Reconstruction	130
3.7.4	Geometric Integration Methods	135
3.7.5	Geometric Reconstruction	140
3.8	Registration – Case Study	143
3.8.1	Notation and Terminology	144
3.8.2	Problem Reformulation	147
3.8.3	Iterative Algorithm to Solve for $\mathbf{R}$	149
3.8.4	Results	155
3.8.5	Conclusions	157
3.9	Surface Reconstruction Summary	158
4.	Object Representation and Feature Matching	161
4.1	Preliminaries	161
4.2	Object-centred Representations	163
4.2.1	Boundary and Curve-based Representations	163
4.2.2	Axial Descriptions	169
4.2.3	Surface Descriptions	170
4.2.4	Volumetric Descriptions	177
4.3	Viewer-centred Representations	180
4.3.1	Aspect Graphs	180
4.3.2	Silhouettes	180
4.3.3	Principal Component Analysis	183
4.3.4	Miscellaneous Techniques	184
4.4	Representation Conclusions	186
4.5	Matching	186
4.5.1	Hypothesise and Test	187
4.5.2	Relational Structures	188
4.5.3	Pose Clustering	190
4.5.4	Geometric Hashing	190
4.5.5	Interpretation Trees	191
4.5.6	Registration and Distance Transforms	194
4.6	Matching Conclusions	194

Part B – Summary	195
Part C – Vision Systems – Case Studies	197
5. Optical Character Recognition	199
5.1 Examples of Existing Systems	200
5.1.1 Prototype Extraction and Adaptive OCR	201
5.1.2 Direct Grayscale Extraction of Features for Character Recognition	201
5.2 Optical Character Recognition System for Cursive Scripts – A Case Study	201
5.2.1 Background	202
5.2.2 An Overview of the Case Study System	205
5.2.3 The Document Image Analysis Step	206
5.2.4 The Recognition-based Segmentation Step	213
5.2.5 The Feature Extraction Stage	214
5.2.6 Results	218
5.2.7 Conclusions	220
6. Recognition by Parts and Part Segmentation Techniques	221
6.1 Examples of Existing Vision Systems	221
6.1.1 HYPER	222
6.1.2 The Logarithmic Complexity Matching Technique	223
6.2 Recognition by Parts and Part Segmentation – A Case Study	224
6.2.1 The Edge Detection Stage	225
6.2.2 The Part Segmentation Stage	229
6.2.3 Part Isolation	233
6.2.4 The Part Identification Stage	234
6.2.5 The Structural Description and Recognition Stage	237
6.2.6 Results	240
6.2.7 Discussion	249
6.2.8 Conclusions	259
7. 3-D Object Recognition Systems	261
7.1 Examples of Existing Systems	261
7.1.1 ACRONYM	262
7.1.2 SCERPO	263
7.1.3 3DPO	263
7.1.4 PREMIO	263
7.1.5 Recognition of MSTAR Targets	264
7.1.6 Bayesian Recognition by Parts in FLIR	265
7.2 3-D Free-form Object Recognition Using Bayesian Splines – A Case Study	265
7.2.1 Preliminaries	267

7.2.2	Bayesian Formulation	269
7.2.3	The RJMCMC Algorithm for Splines	272
7.2.4	Simulated Annealing RJMCMC	275
7.2.5	Matching Splines	276
7.2.6	Results	279
7.2.7	Conclusions	292
Part C – Summary		295
Appendices		297
A.	Vector and Matrix Analysis	299
A.1	Preliminaries	299
A.1.1	Determinant	300
A.1.2	Inversion	301
A.2	Derivatives and Integrals of Matrices	302
A.3	Vectors and Vector Analysis	302
A.4	Eigenvalues and Eigenvectors	304
A.5	Quadratic Forms	304
B.	Principal Component Analysis	305
C.	Optimisation Fundamentals	307
C.1	Fundamental Concepts	307
C.2	Linear Least Squares	308
C.3	Non-linear Optimisation	309
C.4	Direct Search Techniques	310
C.4.1	Simplex Method	310
C.5	Gradient Methods	311
C.5.1	Newton-Raphson Technique	311
C.5.2	Davidon-Fletcher-Powell	312
C.6	Simulated Annealing	313
D.	Differential Geometry – Basic Principles	315
E.	Spline Theory	319
E.1	Spline Definitions	319
F.	Detailed Derivation of Registration Equations	323
F.1	Simplification of $\Phi_{R,T}$	323
F.2	Simplification of Φ_R	327
References		329
Index		347