

Robust Computer Vision

Computational Imaging and Vision

Managing Editor

MAX A. VIERGEVER

Utrecht University, Utrecht, The Netherlands

Editorial Board

GUNILLA BORGEFORS, *Centre for Image Analysis, SLU, Uppsala, Sweden*

THOMAS S. HUANG, *University of Illinois, Urbana, USA*

SABURO TSUJI, *Wakayama University, Wakayama, Japan*

Robust Computer Vision

Theory and Applications

by

Nicu Sebe

*LIACS Media Lab,
Leiden University,
Leiden, The Netherlands*

and

Michael S. Lew

*LIACS Media Lab,
Leiden University,
Leiden, The Netherlands*



Springer-Science+Business Media, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-90-481-6290-1 ISBN 978-94-017-0295-9 (eBook)
DOI 10.1007/978-94-017-0295-9

Printed on acid-free paper

All Rights Reserved

© 2003 Springer Science+Business Media Dordrecht
Originally published by Kluwer Academic Publishers in 2003.
Softcover reprint of the hardcover 1st edition 2003

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

To Dafina and my parents
Nicu

To Hyowon
Michael

Contents

Foreword	xi
Preface	xiii
1. INTRODUCTION	1
1 Visual Similarity	2
1.1 Color	4
1.2 Texture	7
1.3 Shape	9
1.4 Stereo	11
1.5 Motion	13
1.6 Facial expression	13
1.7 Summary	15
2 Evaluation of Computer Vision Algorithms	16
3 Overview of the Book	19
2. MAXIMUM LIKELIHOOD FRAMEWORK	25
1 Introduction	25
2 Statistical Distributions	26
2.1 Gaussian Distribution	27
2.2 Exponential Distribution	38
2.3 Cauchy Distribution	41
3 Robust Statistics	43
3.1 Outliers	44
4 Maximum Likelihood Estimators	45
5 Maximum Likelihood in Relation to Other Approaches	47
6 Our Maximum Likelihood Approach	50
6.1 Scale Parameter Estimation in a Cauchy Distribution	54

7	Experimental Setup	57
8	Concluding Remarks	59
3.	COLOR BASED RETRIEVAL	61
1	Introduction	61
2	Colorimetry	64
3	Color Models	64
3.1	<i>RGB</i> Color System	65
3.2	<i>HSV</i> Color System	66
3.3	$l_1l_2l_3$ Color System	67
4	Color Based Retrieval	68
4.1	Color Indexing	69
5	Experiments with the Corel Database	73
5.1	Early Experiments	73
5.2	Usability Issues	74
5.3	Printer-Scanner Noise Experiments	75
5.4	Color Model	76
5.5	Quantization	76
5.6	Distribution Analysis	77
6	Experiments with the Objects Database	79
7	Concluding Remarks	81
4.	ROBUST TEXTURE ANALYSIS	83
1	Introduction	83
2	Human Perception of Texture	86
3	Texture Features	87
3.1	Texture Distribution Models	88
3.1.1	Gray-level differences	89
3.1.2	Laws' texture energy measures	89
3.1.3	Center-symmetric covariance measures	89
3.1.4	Local binary patterns and trigrams	91
3.1.5	Complementary feature pairs	91
3.2	Gabor and Wavelet Models	92
4	Texture Classification Experiments	95
4.2	Distribution Analysis	97
4.3	Misdetection Rates	99
4.3.1	Summary	104
5	Texture Retrieval Experiments	104

5.1	Texture Features	105
5.2	Experiments Setup	106
5.3	Similarity Noise for QMF-Wavelet Transform	106
5.4	Similarity Noise for Gabor Wavelet Transform	108
6	Concluding Remarks	109
5.	SHAPE BASED RETRIEVAL	111
1	Introduction	111
2	Human Perception of Visual Form	113
3	Active Contours	118
3.1	Behavior of Traditional Active Contours	120
3.2	Generalized Force Balance Equations	124
3.3	Gradient Vector Flow	125
4	Invariant Moments	130
5	Experiments	131
6	Conclusions	134
6.	ROBUST STEREO MATCHING AND MOTION TRACKING	135
1	Introduction	135
1.1	Stereoscopic Vision	137
2	Stereo Matching	138
2.1	Related Work	142
3	Stereo Matching Algorithms	144
3.1	Template Based Algorithm	144
3.2	Multiple Windows Algorithm	146
3.3	Cox' Maximum Likelihood Algorithm	147
4	Stereo Matching Experiments	150
4.1	Stereo Sets	151
4.2	Stereo Matching Results	151
4.3	Summary	157
5	Motion Tracking Experiments	157
6	Concluding Remarks	160
7.	FACIAL EXPRESSION RECOGNITION	163
1	Introduction	164
2	Emotion Recognition	166
2.1	Judgment Studies	167
2.2	Review of Facial Expression Recognition	167

3	Face Tracking and Feature Extraction	171
4	The Static Approach: Bayesian Network Classifiers	173
4.1	Continuous Naive-Bayes: Gaussian and Cauchy Naive Bayes Classifiers	175
4.2	Beyond the Naive-Bayes Assumption: Finding Dependencies among Features Using a Gaussian TAN Classifier	176
5	The Dynamic Approach: Expression Recognition Using Multi-level HMMs	179
5.1	Hidden Markov Models	182
5.2	Expression Recognition Using Emotion-Specific HMMs	183
5.3	Automatic Segmentation and Recognition of Emotions Using Multi-level HMM.	184
6	Experiments	187
6.1	Results Using the Chen Database	191
6.1.1	Person-Dependent Tests	191
6.1.2	Person-Independent Tests	193
6.2	Results Using the Cohn-Kanade Database	194
7	Summary and Discussion	195
	References	199
	Index	210

Foreword

The field of computer vision is both intellectually stimulating and full of important applications. As the field approaches maturity and commercial products start to appear, one of the most challenging problems is: How to make algorithms robust? Computer vision algorithms are notoriously brittle. This timely book presents a Maximum Likelihood framework to deal with robustness. To paraphrase Kendall and Buckland: "An algorithm is robust if it is not very sensitive to departure from the assumptions on which it depends."

During the past decade, researchers in computer vision have found that probabilistic machine learning methods are extremely powerful. This book describes some of these methods. In addition to the Maximum Likelihood framework, Bayesian Networks, and Hidden Markov models are also used. Three aspects are stressed: features, similarity metric, and models. Many interesting and important new results, based on research by the authors and their collaborators, are presented. To single out one result: Experiments after experiments have shown that in many applications the empirical noise/error can be fitted better with a Cauchy rather than a Gaussian model. This reminds me of an analytical result I derived many years ago when I was working on the compression of two-tone images: The differences between corresponding run-lengths of two successive scan lines obey a Cauchy distribution if we assume the directions of the boundaries between black and white is uniformly distributed. Why does the Cauchy distribution pop up so often in real-life data? Are there theoretical results for the Cauchy distribution, akin to the Central Limit Theorem for Gaussian?

Although this book contains many new results, it is written in a style that suits both experts and novices in computer vision. To quote one of my more junior graduate students, who carefully read the manuscript of the book, "It is very clear and easy to understand even for a non-computer vision expert like me."

Finally, a personal note. Michael Lew was my Ph.D. student at Illinois. Nicu Sebe was Michael's Ph.D. student at Leiden, and thus my grand-student. Nicu has also spent time at Illinois, collaborating with my students and me. I am extremely proud to see my student and grand-student produce such a wonderful book.

Thomas S. Huang
Urbana, Illinois
U.S.A.
February 9, 2003

Preface

Computer vision is the enterprise of automating and integrating a wide range of processes and representations used for vision perception. It includes many techniques that are useful by themselves, such as image processing (transforming, encoding, and transmitting images) and statistical pattern classification (statistical decision theory applied to general patterns, visual or otherwise). Moreover, it also includes techniques for geometric modeling and cognitive processing. The field of computer vision may be best understood by considering different types of applications. Many of these applications involve tasks that require either work in a hostile environment, a high rate of processing, access and use of large databases of information, or are tedious for people to perform. Computer vision systems are used in many and various types of environments - from manufacturing plants, to hospital surgical suits, and to the surface of Mars. For example, in manufacturing systems, computer vision is often used for quality control. In this application, the computer vision system scans manufactured items for defects and provides control signals to a robotic manipulator to remove defective parts automatically. Current examples of medical systems being developed include: systems to diagnose skin tumors automatically, systems to aid neurosurgeons during brain surgery, systems to perform clinical tests automatically, etc. The field of law enforcement and security is also an active area for computer vision system development with applications ranging from automatic identification of fingerprints to DNA analysis.

In a standard approach, statistical techniques in computer vision applications must estimate accurate model parameters despite small-scale noise in the data, occasional large-scale measurement errors (outliers), and measurements from multiple populations in the same data set. Increasingly, robust estimation techniques from statistics are being used to solve these parameter estimation problems. Ideally, these techniques should effectively ignore the outliers when estimating the parameters of a single population. In our approach, we consider

applications that involve similarity where the ground truth is provided. The goal is to find the probability density function which maximizes the similarity probability. Furthermore, we derive the corresponding metric from the probability density function by using the maximum likelihood paradigm and we use it in the experiments.

The goal of this book is to describe and illuminate some fundamental principles of robust approaches. Consequently, the intention is to introduce basic concepts and techniques of a robust approach and to develop a foundation, which can be used in a wide variety of computer vision algorithms. Chapter 1 introduces the reader to the paradigms, issues, and important applications involving visual similarity, followed by an in-depth chapter (Chapter 2) which discusses the most influential robust framework - maximum likelihood.

In recent years, the vision community has generalized beyond grayscale algorithms toward color techniques which prompts the third chapter on color based retrieval of images and objects. The other primary features which are frequently discussed in the vision literature are texture and shape which are covered in the fourth chapter and in the fifth chapter, respectively.

Beyond classification algorithms, the computer vision area has been interested in finding correspondences between pairs of images which have been taken from different spatial positions (stereo matching) or different moments in time (motion tracking). Our analysis extends to both of these with respect to recent developments in robust techniques in Chapter 5.

Images containing faces are essential to intelligent vision-based human computer interaction. The rapidly expanding research in face processing is based on the premise that information about the user's identity, state, and intent can be extracted from images and that computers can then react accordingly, e.g., by observing a person's facial expression. The area of facial emotion recognition is covered in Chapter 7.

In each of the chapters we show how the literature has introduced robust techniques into the particular topic area, discuss comparative experiments made by us, and conclude with comments and recommendations. Furthermore, we survey the topic area and describe the representative work done.

Acknowledgments

This book would not have existed without the assistance of Nies Huijsmans, Ira Cohen, Ashutosh Garg, Etienne Loupiau, and Larry Chen whose technical contributions are directly reflected within the chapters. We would like to thank Frans Peters, Arnold Smeulders, Erwin Bakker, Harry Wijshoff, Joost Kok, and our colleagues from the Leiden Institute of Advanced Computer Science and the IFP group at University of Illinois at Urbana-Champaign who gave us valuable suggestions, discussions, and critical comments. Beyond technical contributions, we would like to thank our families for years of patience, sup-

port, and encouragement. Furthermore, we are grateful to the Leiden Institute of Advanced Computer Science for providing an excellent scientific environment.

We would also like to acknowledge the profound influence of Thomas Huang on our scientific work and on this book in particular.