Robust Computer Vision

Computational Imaging and Vision

Volume 26

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

To Dafina and my parents
Nicu

To Hyowon
Michael

Contents

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 runlengths 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 Loupias, 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, support, 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.