Advances in Pattern Recognition

For futher volumes:
http://www.springer.com/series/4205

Nanning Zheng · Jianru Xue

Statistical Learning and Pattern Analysis for Image and Video Processing

Springer

Prof. Nanning Zheng
Xi'an Jiaotong University
Inst. Artificial Intelligence &
Robotics
28 Xianningxilu, Xi'an, 710049
China, People's Republic
nnzheng@mail.xjtu.edu.cn

Prof. Jianru Xue
Xi'an Jiaotong University
Inst. Artificial Intelligence &
Robotics
28 Xianningxilu, Xi'an, 710049
China, People's Republic
jrxue@mail.xjtu.edu.cn

Series editor
Professor Sameer Singh, PhD
Research School of Informatics
Loughborough University
Loughborough, UK

ISSN 1617-7916
ISBN 978-1-84882-311-2 e-ISBN 978-1-84882-312-9
DOI 10.1007/978-1-84882-312-9
Springer Dordrecht Heidelberg London New York

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

Library of Congress Control Number: 2009932139

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Why are We Writing This Book?

Visual data (graphical, image, video, and visualized data) affect every aspect of modern society. The cheap collection, storage, and transmission of vast amounts of visual data have revolutionized the practice of science, technology, and business. Innovations from various disciplines have been developed and applied to the task of designing intelligent machines that can automatically detect and exploit useful regularities (patterns) in visual data. One such approach to machine intelligence is statistical learning and pattern analysis for visual data.

Over the past two decades, rapid advances have been made throughout the field of visual pattern analysis. Some fundamental problems, including perceptual grouping, image segmentation, stereo matching, object detection and recognition, and motion analysis and visual tracking, have become hot research topics and test beds in multiple areas of specialization, including mathematics, neuron-biometry, and cognition. A great diversity of models and algorithms stemming from these disciplines has been proposed. To address the issues of ill-posed problems and uncertainties in visual pattern modeling and computing, researchers have developed rich toolkits based on pattern analysis theory, harmonic analysis and partial differential equations, geometry and group theory, graph matching, and graph grammars.

Among these technologies involved in intelligent visual information processing, statistical learning and pattern analysis is undoubtedly the most popular and important approach, and it is also one of the most rapidly developing fields, with many achievements in recent years. Above all, it provides a unifying theoretical framework for intelligent visual information processing applications.

The main topics of this book are the modeling and computing of visual patterns in image sequences and the methods required to construct intelligent video analysis systems. For visual pattern modeling, we apply statistical learning and statistical pattern analysis to extract semantic visual objects. In our view, such models can be learned efficiently to emulate complex scenes in the image sequence. For video analysis system building, the methods presented here are based on techniques of statistical computing, such as motion analysis, inferring underlying states of objects

of interest, and reducing the dimensionality of video data while preserving useful information as much as possible.

How the Book is Organized

This book provides a comprehensive overview of theories, methodologies, and recent developments in the field of statistical learning and statistical analysis for visual pattern modeling and computing. We had three objectives in selecting topics to cover. We wish to 1) describe a solid theoretical foundation, 2) provide a comprehensive summary of the latest advances of recent years, and 3) present typical issues to be considered in making a real system for visual information processing. We have tried to achieve a balance between these three objectives. The rest of this book is organized as follows:

Chapter 1 is devoted to constructing the theoretic basis for pattern analysis and statistical learning. The fundamentals of statistical pattern recognition and statistical learning are presented via introducing the general framework of a statistical pattern recognition system. We also discuss pattern representation and classification, two important components of such a system, as well as concepts involved in three main approaches to statistical learning: supervised learning, semistatistical learning, and unsupervised learning. This introduction leads to the development of three parts of the whole book.

In the first part, we focus on the unsupervised learning of visual pattern representational models for objects in images, which covers through Chapters 2 to 5. Usually, what a vision algorithm can accomplish depends crucially on how much it knows about content of the visual scenes. This knowledge can be mathematically represented by simple but general models that can realistically characterize visual patterns in the ensemble of visual data. Representation and computation are thus two principal problems in visual computing. We provide a comprehensive survey of recent advances in statistical learning and pattern analysis with respect to these two problems. Chapter 2 discusses cluster analysis and perceptual grouping algorithms used in unsupervised visual pattern analysis. The systematic approaches for deriving these models are also illustrated step by step. Chapters 3 through 5 focus on representing and learning visual patterns in both spatial and temporal domains. Chapter 3 describes component analysis approaches, which are used to find hidden components via visual data analysis techniques. Chapter 4 discusses the manifold learning perspective on visual pattern representation, dimensionality reduction, and classification problems. Chapter 5 presents a review of recent advances in the adaptive wavelet transform for image and video coding.

In the second part, we introduce the supervised learning of visual patterns in images, which is covered in Chapter 6. We focus on supervised statistical pattern analysis and introduce concepts and major techniques in feature extraction and selection as well as classifier design. Especially, we introduce statistical machine learning techniques by examining the support vector machine and AdaBoost classifier.

In the third part, we focus on the visual pattern analysis in video, which covers through Chapters 7 to 11. In this part, we discuss methodologies for building intelligent video analysis systems such as object detection, tracking, and recognition in video. Chapter 7 focuses on the critical aspects of motion analysis, including statistical optical flow, model-based motion analysis, and joint motion estimation and segmentation. For the object-level motion analysis, we first introduce the sequential Bayesian estimation framework in Chapter 8, which acts as the theoretic basis for visual tracking, and then present approaches to constructing a representation model of specific objects. Then, in Chapter 9, we present a probabilistic fusion framework for robust tracking. Chapters 10 and 11 offer a multitarget tracking in video (MTTV) formulation that exploits a Markov network whose solution is arrived at using Monte Carlo-based belief propagation. Using this approach, problems including occlusion and various number of objects in MTTV are addressed.

Finally, in Chapter 12, we make an in-depth discussion of visual data processing in the cognitive process. A new scheme of association memory and new architecture of artificial intelligent system with attractors of chaos are also addressed. We argue that to make a breakthrough in current research on intelligent visual data processing, people should pay great attention to the mechanism of cognition and selective attention.

Acknowledgments

First, we thank the National Science Foundation of China (Grant Nos. 608750081 and 60635050) for their continuing support of our research in the image and video processing field. Our interest in statistical learning and pattern analysis approaches to image and video processing could not have developed to the point of writing this book without their funding.

We also thank Xi'an Jiaotong University, for their support of our efforts in writing this book, and Wayne Wheeler, Senior editor of Springer-Verlag, for his suggestion in writing this book. Another major source of stimulation for this effort comes from our former graduate students, who have provided a continuing stream of new ideas and insights. Those who have made contributions directly reflected in this book include Mr. Xiaoping Zhong, Dr. Weixiang Liu, and Dr. Zaide Liu.

A special thanks goes to Dr. Clayton McMillan, who read the first draft very carefully and made numerous corrections for improvements. We also thank Dr. Shaoyi Du and Dr. Gaofeng Meng for their dedication and patience in preparing drafts and writing Chapter 6. The second author would like to thank Prof. Songchun Zhu for the hosting of his visiting in the Center for Image and Vision Science in University of California, Los Angeles, when he was writing this book.

Last but not least, we are greatly indebted to our families: our parents and our wives and children for their unconditional support, encouragement, and continuing love throughout our research work.

Contents