Machine Interpretation of Patterns

Image Analysis and Data Mining

Statistical Science and Interdisciplinary Research

Series Editor: Sankar K. Pal (*Indian Statistical Institute*)

Description:

In conjunction with the Platinum Jubilee celebrations of the Indian Statistical Institute, a series of books will be produced to cover various topics, such as Statistics and Mathematics, Computer Science, Machine Intelligence, Econometrics, other Physical Sciences, and Social and Natural Sciences. This series of edited volumes in the mentioned disciplines culminate mostly out of significant events — conferences, workshops and lectures — held at the ten branches and centers of ISI to commemorate the long history of the institute.

- Vol. 3 Algorithms, Architectures and Information Systems Security
*edited by Bhargab B. Bhattacharya, Susmita Sur-Kolay,
Subhas C. Nandy & Aditya Bagchi
(Indian Statistical Institute, India)*
- Vol. 4 Advances in Multivariate Statistical Methods
edited by A. SenGupta (Indian Statistical Institute, India)
- Vol. 5 New and Enduring Themes in Development Economics
*edited by B. Dutta, T. Ray & E. Somanathan
(Indian Statistical Institute, India)*
- Vol. 6 Modeling, Computation and Optimization
*edited by S. K. Neogy, A. K. Das and R. B. Bapat
(Indian Statistical Institute, India)*
- Vol. 7 Perspectives in Mathematical Sciences I: Probability and Statistics
*edited by N. S. N. Sastry, T. S. S. R. K. Rao, M. Delampady and
B. Rajeev (Indian Statistical Institute, India)*
- Vol. 8 Perspectives in Mathematical Sciences II: Pure Mathematics
*edited by N. S. N. Sastry, T. S. S. R. K. Rao, M. Delampady and
B. Rajeev (Indian Statistical Institute, India)*
- Vol. 9 Recent Developments in Theoretical Physics
edited by S. Ghosh and G. Kar (Indian Statistical Institute, India)
- Vol. 11 Machine Interpretation of Patterns: Image Analysis and Data Mining
*edited by R. K. De, D. P. Mandal and A. Ghosh
(Indian Statistical Institute, India)*
- Vol. 12 Recent Trends in Surface and Colloid Science
edited by Bidyut K. Paul (Indian Statistical Institute, India)

Platinum Jubilee Series

Statistical Science and
Interdisciplinary Research — Vol. 11

Machine Interpretation of Patterns

Image Analysis and Data Mining

Editors

Rajat K. De
Deba Prasad Mandal
Ashish Ghosh

Indian Statistical Institute, India

Series Editor: Sankar K. Pal

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

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

Statistical Science and Interdisciplinary Research — Vol. 11

MACHINE INTERPRETATION OF PATTERNS

Image Analysis and Data Mining

Copyright © 2010 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN-13 978-981-4299-18-3

ISBN-10 981-4299-18-9

Printed in Singapore.

Foreword

The Indian Statistical Institute (ISI) was established on 17th December, 1931, by a great visionary Prof. Prasanta Chandra Mahalanobis to promote research in the theory and applications of statistics as a new scientific discipline in India. In 1959, Pandit Jawaharlal Nehru, the then Prime Minister of India introduced the ISI Act in the parliament and designated it as an Institution of National Importance because of its remarkable achievements in statistical work as well as its contribution to economic planning.

Today, the Indian Statistical Institute occupies a prestigious position in the academic firmament. It has been a haven for bright and talented academics working in a number of disciplines. Its research faculty has done India proud in the arenas of Statistics, Mathematics, Economics, Computer Science, among others. Over seventy five years, it has grown into a massive banyan tree, like the institute emblem. The Institute now serves the nation as a unified and monolithic organization from different places, namely Kolkata, the Headquarters, Delhi, Bangalore and Chennai, three centers, a network of five SQC-OR Units located at Mumbai, Pune, Baroda, Hyderabad and Coimbatore, and a branch (field station) at Giridih.

The platinum jubilee celebrations of ISI have been launched by Honorable Dr. Manmohan Singh on December 24, 2006, and the Government of India has declared 29th June as the “Statistics Day” to commemorate the birthday of Prof. Mahalanobis nationally.

Prof. Mahalanobis, was a great believer in interdisciplinary research because he thought that this will promote the development of not only Statistics, but also the other natural and social sciences. To promote interdisciplinary research, major strides were made in the areas of computer science, statistical quality control, economics, biological and social sciences, physical and earth sciences.

The Institute’s motto of “unity in diversity” has been the guiding principle of all its activities since its inception. It highlights the unifying

role of statistics in relation to various scientific activities.

In tune with this hallowed tradition, a comprehensive academic programme, involving Nobel Laureates, Fellows of the Royal Society, Abel prize winner and other dignitaries, has been implemented throughout the Platinum Jubilee year, highlighting the emerging areas of ongoing frontline research in its various scientific divisions, centers, and outlying units. It includes international and national-level seminars, symposia, conferences and workshops, as well as series of special lectures. As an outcome of these events, the Institute is bringing out a series of comprehensive volumes in different subjects under the title Statistical Science and Interdisciplinary Research, published by the World Scientific Press, Singapore.

The present volume titled “Machine Interpretation of Patterns — Image Analysis and Data Mining” is the eleventh one in the series. The volume consists of fifteen chapters, written by eminent scientists from different parts of the world. These chapters provide a current perspective of different areas of research and development, both from theoretical and application points of view, emphasizing the major challenging issues primarily in pattern recognition, computer vision, machine learning and data mining. Interesting application areas covered include mapping earthquake damage in urban areas, personalization and context modeling in web mining, human fall detection from videos of multiple persons in computer vision, and prediction of soluble proteins in bioinformatics. Tools used are both classical and modern including case based reasoning, granular support vector machine, decision tree, parallel genetic algorithms and bipolar fuzzy sets. I believe the state-of-the art studies presented in this book will be very useful to students, researchers as well as practitioners.

Thanks to the contributors for their excellent research contributions and to the volume editors Dr. Rajat K. De, Dr. D. P. Mandal and Prof. Ashish Ghosh for their sincere effort in bringing out the volume nicely. Initial design of the cover by Mr. Indranil Dutta is acknowledged. Sincere efforts by Prof. Dilip Saha and Dr. Barun Mukhopadhyay for editorial assistance are appreciated. Thanks are also due to World Scientific for their initiative in publishing the series and being a part of the Platinum Jubilee endeavor of the Institute.

January 2010
Kolkata

Sankar K. Pal
Series Editor and Director

Preface

This book provides some recent developments and state of the art review in various areas of pattern recognition, image processing, machine learning, data mining and web intelligence, both from theoretical and application points of view. It is a peer-reviewed research monograph under the ISI Platinum Jubilee Series on Statistical Science and Interdisciplinary Research. The book may be considered as an integrated volume to the researchers interested in doing interdisciplinary research where computer science is a component. Since a wide variety of current topics are covered in the book, the readers are expected to have some basic idea of pattern recognition, machine learning, image processing, data mining and web intelligence; but will be able to follow each of the chapters independent of the others. The book contains a total of 15 chapters. The first 5 chapters deal with development of a few methodologies for various tasks of pattern recognition and image processing, along with some theoretical aspects. Chapters 6, 7 and 8 consider text processing and other related tasks. The next four chapters (Chapters 9–12) consider various issues of computer vision from application points of view. The last 3 chapters (Chapters 13–15) describe the classification task of pattern recognition, involving signal processing, case-based reasoning and support vector machines. These chapters are on applications of pattern recognition to speech processing, broker system designing and protein classification.

Chapter 1 deals with the design of a probabilistic multi-class multi-kernel machine based on a kernel combination methodology. The classifier is able to combine diverse sources of information and multiple feature spaces. The methodology follows a well-founded hierarchical Bayesian paradigm that models uncertainty in the parameters via a hierarchy of prior and hyper-prior distributions. Chapter 2 presents an image quality metric which integrates the notions of structural similarity measure mimicking the overall functionality of the human visual system and perceptually important regions based on the characteristics of intermediate

and higher level visual processes. In Chapter 3, a new approach for image entropy reduction is described, which is based on two-dimensional fractional differentiation. Chapter 4 deals with the development of two novel methods for object-background classification using the notion of parallel genetic algorithms. Chapter 5 addresses the issue of bipolarity in spatial domains. It introduces mathematical morphological operations on bipolar fuzzy sets.

Chapter 6 describes the present day scenario of the area of information retrieval. The main focus is on the problem of personalization and context modelling, and on the definition of new paradigms for content representation. An online archiving and retrieving system is developed in Chapter 7. This is followed by the development of a two stage recognition scheme, in Chapter 8, for recognition of handwritten Devanagari words.

Chapter 9 is concerned with a problem of computer vision. It deals with the development of a model based approach for human fall detection from videos in the presence of multiple persons. Chapter 10 shows how multi-temporal SAR data allows mapping earthquake damage in urban areas with an acceptable accuracy, provided the ancillary information defining urban blocks are available. Chapters 11 and 12 describe the designs of safety related systems based on the theories of machine learning.

In Chapter 13, a methodology is developed for signal reconstruction using the notion of non-linear quantization based on logarithmic and Incomplete Beta Function (IBF) which dynamically assign the number of quantization levels exploiting the auditory motivation. Chapter 14 addresses some discovery tasks as correlated with supervised learning. Textual case-based and decision tree induction approaches are considered to these tasks, and the use of multiple representations is explored. The performance evaluation is made on a broker system that discovers and mediates requests, and responses for meteorological and oceanographic data. Finally in Chapter 15, an important bioinformatics problem is considered where a Granular Support Vector Machine (GSVM) is used for prediction of soluble proteins based on over expression in *Escherichia coli*.

Rajat K. De
Deba Prasad Mandal
Ashish Ghosh
Editors

Contents

<i>Foreword</i>	v
<i>Preface</i>	vii
1. Combining Information with a Bayesian Multi-class Multi-kernel Pattern Recognition Machine <i>T. Damoulas and M. A. Girolami</i>	1
2. Image Quality Assessment Based on Weighted Perceptual Features <i>D. V. Rao and L. P. Reddy</i>	29
3. Quasi-reversible Two-dimension Fractional Differentiation for Image Entropy Reduction <i>A. Nakib, A. Nait-Ali, H. Oulhadj and P. Siarry</i>	57
4. Parallel Genetic Algorithm Based Clustering for Object and Background Classification <i>P. Kanungo, P. K. Nanda and A. Ghosh</i>	69
5. Bipolar Fuzzy Spatial Information: First Operations in the Mathematical Morphology Setting <i>I. Bloch</i>	91
6. Approaches to Intelligent Information Retrieval <i>G. Pasi</i>	113
7. Retrieval of On-line Signatures <i>H. N. Prakash and D. S. Guru</i>	131

8.	A Two Stage Recognition Scheme for Offline Handwritten Devanagari Words <i>B. Shaw and S. K. Parui</i>	145
9.	Fall Detection from a Video in the Presence of Multiple Persons <i>V. Vishwakarma, S. Sural and C. Mandal</i>	167
10.	Fusion of GIS and SAR Statistical Features for Earthquake Damage Mapping at the Block Scale <i>G. Trianni, G. Lisini, P. Gamba and F. Dell'Acqua</i>	195
11.	Intelligent Surveillance and Pose-invariant 2D Face Classification <i>B. C. Lovell, C. Sanderson and T. Shan</i>	207
12.	Simple Machine Learning Approaches to Safety-related Systems <i>C. Moewes, C. Otte and R. Kruse</i>	231
13.	Nonuniform Multi Level Crossings for Signal Reconstruction <i>N. Poojary, H. Kumar and A. Rao</i>	251
14.	Adaptive Web Services Brokering <i>K. M. Gupta and D. W. Aha</i>	269
15.	Granular Support Vector Machine Based Method for Prediction of Solubility of Proteins on Over Expression in Escherichia Coli and Breast Cancer Classification <i>P. Kumar, B. D. Kulkarni and V. K. Jayaraman</i>	289