

NIKOLAOS V. BOULGOURIS  
KONSTANTINOS N. PLATANIOTIS  
EVANGELIA MICHELI-TZANAKOU  
EDITORS

# BIOMETRICS

Theory, Methods, and Applications

 WILEY

 **IEEE**  
Celebrating 125 Years  
of Engineering the Future



IEEE Press Series on  
Computational Intelligence  
David B. Fogel, Series Editor



# **Biometrics**

**IEEE Press**  
**445 Hoes Lane**  
**Piscataway, NJ 08854**

**IEEE Press Editorial Board**  
Lajos Hanzo, *Editor in Chief*

R. Abari  
J. Anderson  
S. Basu  
A. Chatterjee

T. Chen  
T. G. Croda  
M. El-Hawary  
S. Farshchi

B. M. Hammerli  
O. Malik  
S. Nahavandi  
W. Reeve

Kenneth Moore, *Director of IEEE Book and Information Services (BIS)*

**Technical Reviewer**

Qinghan Xiao, Defence Research and Development Canada

IEEE-CIS Liaison to IEEE Press, Gary B. Fogel

**Books in the IEEE Press Series on Computational Intelligence**

*Evolving Intelligent Systems: Methodology and Applications*

Edited by Plamen Angelov, Dimitar Filev, and Nik Kasabov

2010 978-0470-28719-4

*Biometrics: Theory, Methods, and Applications*

Edited by Nikolaos V. Boulgouris, Konstantinos N. Plataniotis, and Evangelia Micheli-Tzanakou

2010 978-0470-24782-2

*Clustering*

Rui Wu and Donald Wunsch II

2009 978-0470-27680-0

*Computational Intelligence in Bioinformatics*

Edited by David B. Fogel, David W. Corne, and Yi Pan

2008 978-0470-10526-9

*Introduction to Evolvable Hardware: A Practical Guide for Designing Self-Adaptive Systems*

Garrison W. Greenwood and Andrew M. Tyrrell

2007 978-0471-71977-9

*Evolutionary Computation: Toward a New Philosophy of Machine Intelligence, Third Edition*

David B. Fogel

2006 978-0471-66951-7

*Emergent Information Technologies and Enabling Policies for Counter-Terrorism*

Edited by Robert L. Popp and John Yen

2006 978-0471-77615-4

*Computationally Intelligent Hybrid Systems*

Edited by Seppo J. Ovaska

2005 0-471-47668-4

*Handbook of Learning and Appropriate Dynamic Programming*

Edited by Jennie Si, Andrew G. Barto, Warren B. Powell, Donald Wunsch II

2004 0-471-66054-X

*Computational Intelligence: The Experts Speak*

Edited By David B. Fogel and Charles J. Robinson

2003 0-471-27454-2

# **Biometrics**

## **Theory, Methods, and Applications**

Edited by

**Nikolaos V. Boulgouris**

**Konstantinos N. Plataniotis**

**Evangelia Micheli-Tzanakou**

**IEEE Computational Intelligence Society, *Sponsor***

**IEEE Press Series on Computational Intelligence**

David B. Fogel, *Series Editor*



IEEE PRESS



A John Wiley & Sons, Inc., Publication

Copyright © 2010 by the Institute of Electrical and Electronics Engineers, Inc.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey. All rights reserved.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at [www.copyright.com](http://www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

**Limit of Liability/Disclaimer of Warranty:** While the publisher, the editors, and the authors have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor editor/author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at [www.wiley.com](http://www.wiley.com).

***Library of Congress Cataloging-in-Publication Data is available.***

ISBN 978-0470-24782-2

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

# Contents

---

|   |             |
|---|-------------|
| <b>Preface</b>  | <b>vii</b>  |
| <b>Contributors</b>   | <b>xiii</b> |
| <b>1. Discriminant Analysis for Dimensionality Reduction:<br/>An Overview of Recent Developments</b>              | <b>1</b>    |
| <i>Jieping Ye and Shuiwang Ji</i>   |             |
| <b>2. A Taxonomy of Emerging Multilinear Discriminant Analysis<br/>Solutions for Biometric Signal Recognition</b> | <b>21</b>   |
| <i>Haiping Lu, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos</i>                                 |             |
| <b>3. A Comparative Survey on Biometric Identity Authentication<br/>Techniques Based on Neural Networks</b>       | <b>47</b>   |
| <i>Raghudeep Kannavara and Nikolaos Bourbakis</i>   |             |
| <b>4. Designing Classifiers for Fusion-Based Biometric Verification</b>   | <b>81</b>   |
| <i>Kritha Venkataramani and B. V. K. Vijaya Kumar</i>   |             |
| <b>5. Person-Specific Characteristic Feature Selection<br/>for Face Recognition</b>                               | <b>113</b>  |
| <i>Sreekar Krishna, Vineeth Balasubramanian, John Black,<br/>and Sethuraman Panchanathan</i>                      |             |
| <b>6. Face Verification Based on Elastic Graph Matching</b>   | <b>143</b>  |
| <i>Anastasios Tefas and Ioannis Pitas</i>   |             |
| <b>7. Combining Geometrical and Statistical Models<br/>for Video-Based Face Recognition</b>                       | <b>171</b>  |
| <i>Amit K. Roy-Chowdhury and Yilei Xu</i>   |             |
| <b>8. A Biologically Inspired Model for the Simultaneous Recognition<br/>of Identity and Expression</b>           | <b>195</b>  |
| <i>Donald Neth and Aleix M. Martinez</i>  |             |
| <b>9. Multimodal Biometrics Based on Near-Infrared Face Recognition</b>   | <b>225</b>  |
| <i>Rui Wang, Shengcai Liao, Zhen Lei, and Stan Z. Li</i>  |             |
| <b>10. A Novel Unobtrusive Face and Hand Geometry Authentication<br/>System Based on 2D and 3D Data</b>           | <b>241</b>  |
| <i>Filareti Tsalakanidou and Sotiris Malassiotis</i>  |             |
| <b>11. Learning Facial Aging Models: A Face Recognition Perspective</b>   | <b>271</b>  |
| <i>Narayanan Ramanathan and Rama Chellappa</i>  |             |

|   |            |
|---|------------|
| <b>12. Super-Resolution of Face Images</b>  | <b>295</b> |
| <i>Sung Won Park and Marios Savvides</i>  |            |
| <b>13. Iris Recognition</b>   | <b>315</b> |
| <i>Yung-Hui Li and Marios Savvides</i>  |            |
| <b>14. Learning in Fingerprints</b>   | <b>339</b> |
| <i>Alessandra Lumini, Loris Nanni, and Davide Maltoni</i>   |            |
| <b>15. A Comparison of Classification- and Indexing-Based Approaches for Fingerprint Identification</b>                               | <b>365</b> |
| <i>Xuejun Tan, Bir Bhanu, and Rong Wang</i>   |            |
| <b>16. Electrocardiogram (ECG) Biometric for Robust Identification and Secure Communication</b>                                       | <b>383</b> |
| <i>Francis Minhthang Bui, Foteini Agrafioti, and Dimitrios Hatzinakos</i>   |            |
| <b>17. The Heartbeat: The Living Biometric</b>  | <b>429</b> |
| <i>Steven A. Israel, John M. Irvine, Brenda K. Wiederhold, and Mark D. Wiederhold</i>   |            |
| <b>18. Multimodal Physiological Biometrics Authentication</b>   | <b>461</b> |
| <i>Alessandro Riera, Aureli Soria-Frisch, Mario Caparrini, Ivan Cester, and Giulio Ruffini</i>  |            |
| <b>19. A Multiresolution Analysis of the Effect of Face Familiarity on Human Event-Related Potentials</b>                             | <b>483</b> |
| <i>Brett DeMarco and Evangelia Micheli-Tzanakou</i>   |            |
| <b>20. On-Line Signature-Based Authentication: Template Security Issues and Countermeasures</b>                                       | <b>497</b> |
| <i>Patrizio Campisi, Emanuele Maiorana, and Alessandro Neri</i>   |            |
| <b>21. Unobtrusive Biometric Identification Based on Gait</b>   | <b>539</b> |
| <i>Xiayi Huang and Nikolaos V. Boulgouris</i>   |            |
| <b>22. Distributed Source Coding for Biometrics: A Case Study on Gait Recognition</b>   | <b>559</b> |
| <i>Savvas Argyropoulos, Dimosthenis Ioannidis, Dimitrios Tzovaras, and Michael G. Strintzis</i>                                       |            |
| <b>23. Measuring Information Content in Biometric Features</b>  | <b>579</b> |
| <i>Richard Youmaran and Andy Adler</i>  |            |
| <b>24. Decision-Making Support in Biometric-Based Physical Access Control Systems: Design Concept, Architecture, and Applications</b> | <b>599</b> |
| <i>Svetlana N. Yanushkevich, Vlad P. Shmerko, Oleg Boulanov, and Adrian Stoica</i>  |            |
| <b>25. Privacy in Biometrics</b>  | <b>633</b> |
| <i>Stelvio Cimato, Marco Gamassi, Vincenzo Piuri, Roberto Sassi, and Fabio Scotti</i>   |            |
| <b>26. Biometric Encryption: The New Breed of Untraceable Biometrics</b>  | <b>655</b> |
| <i>Ann Cavoukian and Alex Stoianov</i>  |            |
| <b>Index</b>  | <b>719</b> |

# Preface

---

The objective of biometric systems is the recognition or authentication of individuals based on some physical or behavioral characteristics that are intrinsically unique for each individual. Nowadays, biometric systems are fundamental components of advanced security architectures. The applications of biometrics range from access control, military, and surveillance to banking and multimedia copyright protection. Recently, biometric information has started to become an essential element in government issued authentication and travel documents. The large-scale deployment of biometrics sensors in a variety of electronic devices, such as mobile phones, laptops, and personal digital assistants (PDA), has further accelerated the pace at which the demand for biometric technologies has been growing. The immense interest in the theory, technology, applications, and social implications of biometric systems has created an imperative need for the systematic study of the use of biometrics in security and surveillance infrastructures.

This edited volume provides an extensive survey of biometrics theory, methods, and applications, making it a good source of information for researchers, security experts, policy makers, engineers, and graduate students. The volume consists of 26 chapters which cover most aspects of biometric systems. The first few chapters address particular recognition techniques that can be used in conjunction with a variety of biometric traits. The following chapters present technologies tailored to specific biometric traits, such as face, hand geometry, fingerprints, signature, electrocardiogram, electroencephalogram, and gait. The remaining chapters focus on both theoretical issues as well as issues related to the emerging area of privacy-enhancing biometric solutions.

**An overview of recent developments in discriminant analysis for dimensionality reduction** is presented in the first chapter. Specifically, a unified framework is presented for generalized linear discriminant analysis (LDA) via a transfer function. It is shown that various LDA-based algorithms differ in their transfer functions. This framework explains the properties of various algorithms and their relationship. Furthermore, the theoretical properties of various algorithms and their relationship are also presented. An emerging extension of the classical LDA is the **multilinear discriminant analysis (MLDA) for biometric signal recognition**. Biometric signals are mostly multidimensional objects, known as tensors. Recently, there has been a growing interest in MLDA solutions. In Chapter 2, the fundamentals of existing MLDA solutions are presented and then categorized according to the multilinear projection employed. At the same time, their connections with traditional linear solutions are pointed out. The next two chapters present classification issues in biometric identification. The problem of classification is extremely important because it

essentially sets the framework regarding the way decisions are made once feature extraction and dimensionality reduction have taken place. A variety of classification approaches can be taken. One of these approaches is to use neural networks (NN). Chapter 3 is a **comparative survey on biometric identity authentication techniques based on neural networks**. This chapter presents a survey on representative NN-based methodologies for biometric identification. In particular, it captures the evolution of some of the representative NN-based methods in order to provide an outline of the application of neural nets in biometric systems. A specific, but far from uncommon, case of classification is that involving fusion of biometrics. The main task here is the **design of classifiers for fusion-based biometric verification**, which is addressed in Chapter 4. The chapter provides guidelines for optimal ensemble generation, where each classifier in the ensemble is a base classifier. Examples are shown for support vector machines and correlation filters. The chapter also focuses on decision fusion rules and the effect of classifier output diversity on their decision fusion accuracy is also analyzed.

Chapters 5–20 present systems based on specific biometric modalities. Methods for face recognition/verification are presented in Chapters 5–8. One of the most important problems in face recognition is feature selection. Chapter 5 presents a **person-specific characteristic feature selection for face recognition**. In this chapter, a new methodology for face recognition is introduced that detects and extracts unique features on a person's face and then uses those features for the purpose of recognition. Chapter 6 presents a different approach by performing **face verification based on elastic graph matching**. Using elastic graph matching, a face is represented as a connected graph. This approach endows the recognition process with robustness against geometric distortions of the facial image. Another challenging task in the area of face-based biometric systems is the efficient use of video sequences for face authentication. Chapter 7 presents a method for the **combination of geometrical and statistical models for video-based face authentication**. In this chapter, it is shown that it is possible to describe object appearance using a combination of analytically derived geometrical models and statistical data analysis. Specifically, a framework that is robust to large changes in facial pose and lighting conditions is presented for face recognition from video sequences. The method can handle situations where the pose and lighting conditions in the training and testing data are completely disjoint. Chapter 8 is about a **biologically inspired model for the simultaneous recognition of identity and expression**. This work builds upon the fact that faces can provide a wide range of information about a person's identity, race, sex, age and emotional state. In most cases, humans easily derive such information by processes that appear rapid and automatic. However, upon closer inspection, one finds these processes to be diverse and complex. This chapter examines the perception of identity and emotion. Next, it develops a computational model that is applied for identification based on face images with differing expression as well as for the classification of expressions.

Chapters 9–11 present some more advanced methods for face recognition. The first two of these chapters go beyond the conventional approach and are based on the realization that face recognition does not have to rely on an image taken using a

conventional camera. **Face recognition using infrared cameras** is a very interesting extension of conventional face recognition. In Chapter 9, a near-infrared (NIR) face-based approach is presented for multimodal biometric fusion. The NIR face is fused with the visible light (VL) face or iris modality. This approach has several advantages, including the fact that NIR face recognition overcomes problems arising from uncontrolled illumination in VL images and achieves significantly better results than when VL faces are used. Furthermore, the fusion of NIR face with VL face or iris is a natural combination for multibiometric solutions. A different, multimodal system based on the fusion of **2D and 3D face and hand geometry data** is presented in Chapter 10. This topic is of particular interest because recent advances in multimodal biometrics as well as the emergence of affordable 3D imaging technologies have created great potential for techniques that involve 3D data. The main advantage is the simultaneous acquisition of a pair of depth and color images of biometric information using low-cost sensors. Although the above face-based methodologies offer improved performance, they are not directly improving the resilience of visual biometric systems to the change that these biometrics undergo through time. Aging is a crucial factor for recognition applications and, in the case of face recognition, can be dealt with by using **learning facial aging models**. Such facial models studied in Chapter 11 can be used for the prediction of one's appearance across ages and, therefore, are of great importance for performing reliable face recognition across age progression. Chapter 12 is about **super-resolution** techniques, which can be used in conjunction with face recognition technologies.

The next three chapters are devoted to iris and fingerprint recognition. The technologies that are used in **iris recognition** systems are presented in Chapter 13. Iris recognition is an extremely reliable technique for identification of individuals, and this chapter reviews both its theoretical and practical aspects. Fingerprint recognition is another very important technology that has been reliably used in biometric systems for a many years. Chapter 14, entitled **learning in fingerprints**, gives a short introduction of the basic concepts and terminology. Furthermore, it provides a detailed review of the existing literature by discussing the most salient learning-based approaches applied to feature extraction, matching, and classification of fingerprints. Chapter 15 makes a **comparison of classification and indexing-based approaches for fingerprint recognition**. This chapter presents a comparison of two key approaches for fingerprint identification. These approaches are based on classification followed by verification and indexing followed by verification. The fingerprint classification approach is based on a feature-learning algorithm, while the indexing approach is based on features derived from triplets of minutiae.

Chapters 16 and 17 present methods using electrocardiograms (ECG). ECG is essentially a medical diagnostic technique but more recently it has fulfilled a rather unlikely role, as a provider of security and privacy in the form of a biometric. Chapter 16, entitled **Electrocardiogram (ECG) Biometric for Robust Identification and Secure Communication**, examines the various implications and technical challenges of using the ECG as a biometric. Specifically, novel signal processing techniques are surveyed and proposed that seek to not only establish the status of the ECG as an indisputable biometric trait, but also reinforce its versatile utility, such as

in alleviating the resource consumption in certain communication networks. Chapter 17 discusses **the heartbeat as a living biometric**. Although previous research on the topic focused mainly on analysis of the electrocardiogram, this chapter extends the ECG results by applying processing methods to a larger and more diverse set of individuals, demonstrating that performance remains high for a larger and a more diverse population. Alternative sensing methods, using blood pressure and pulse oximetry, are presented and their corresponding performance is documented. The chapter also discusses the phenomenology and sensing modalities for monitoring cardiovascular function and, finally, examines the fusion of heartbeat information across the three modalities and quantifies its performance.

Chapters 18 and 19 explore methodologies mainly based on electroencephalograms (EEG). In Chapter 18, a method is proposed using **physiological signals for key features in high-security biometric systems**. The experimental protocol that is common for EEG and ECG recording is explained. EEG and ECG features as well as the authentication algorithms are presented and their efficiency is individually assessed. A fusion process carried out to achieve higher performance is also presented. Chapter 19 presents a **multiresolution analysis of the effect of face familiarity on human event-related potentials**. This method works by processing of the electroencephalograms (EEGs) in response to familiar and unfamiliar face stimuli. Stimuli were presented in successive trials and consisted of (a) multiple presentations of frontal, gray-scale images of one person known to the subject and (b) unique unknown images taken from multiple face databases. Coherent oscillations in phase were observed in the lower delta activity of ERPs in response to known stimuli but not in response to unknown stimuli.

Chapters 20 and 21 present methods and applications based on signature recognition and gait recognition. Although several approaches can be used in authentication systems, the most commonly used authentication method in everyday transactions is based on signature. The specific points of concern regarding **online signature-based authentication** have to do more with template security issues and countermeasures. Chapter 20 focuses on the security issues related to biometric templates, with application to signature based authentication systems. The main privacy and security issues are briefly summarized and some approaches that are used for the protection of biometric templates are discussed. Data hiding techniques are used to design a security scalable authentication system. The enrollment and the authentication procedure are detailed. In contrast to the signature, which is an established method for authentication, gait recognition is an emerging technology that is particularly attractive for biometric identification because the capturing of gait can take place in an unobtrusive manner. Chapter 21 presents the fundamental approaches for **unobtrusive biometric identification based on gait** and provides directions for future research.

Chapters 22–26 deal with biometric applications including issues related to the concept of biometric capacity. Chapter 22, presents a completely new framework for biometric authentication in secure environments as well as a relevant application based on gait recognition. The proposed framework is based on **distributed source coding for biometrics**. In this new framework, the problem of biometric recognition

is formulated as the dual of data communication over noisy channels. In such a system, the enrollment and authentication procedures are considered as the encoding and decoding stages of a communication system. The above approach is highly relevant to information theory. Further application of information theory in biometrics can be found in the assessment of the information content of biometric traits. The discriminating ability of biometric features is usually estimated by means of experimentation. However, the information carried by biometrics, their uniqueness, and their fusion prospects can be studied based on concepts from information theory. This is the topic of Chapter 23, which deals with **measuring information content in biometric features**. Next, in Chapter 24, a summary is presented of the theoretical results and design experience obtained during the development of a next generation physical access security system (PASS). The main feature of this PASS is its efficient **decision-making support** of security personnel enhanced with the situational awareness paradigm and intelligent tools.

Despite the increasing use of biometric features for authentication and identification purposes in a broad variety of institutional and commercial systems, the adoption of biometric techniques is restrained by a rising concern regarding the protection of the biometrics templates. In fact, people are not generally keen to give out biometric traits unless they are assured that their biometrics cannot be stolen or used without their consent. Recent results showed that it is feasible to generate a unique identifier by combining biometric traits. This approach makes it impossible to recover the original biometric features and, thus, ensures the privacy of the biometrics. Chapter 25, entitled **Privacy in Biometrics**, reviews the privacy issues related to the use of biometrics, presents some of the most advanced techniques available up to date, provides a comparative analysis, and gives an overview of future trends. A particular system that builds privacy into an information system is presented in the final chapter, entitled **Biometric Encryption**. In this chapter, the emerging area of privacy-enhancing biometric technologies, referred to as “untraceable biometrics,” makes it possible to enhance both privacy and security in a positive-sum model.

By its nature, an edited volume covers only a limited number of works and initiatives in the area of biometric systems. Researchers and practitioners are introducing new developments at a very fast pace, and it would be impossible to cover all of them in a single volume. However, we believe that the collection of chapters presented here cover sufficiently well the theory, methods, and applications of biometrics. Readers who wish to further explore the fascinating area of biometrics can find additional information using the bibliographic links that are provided in each one of the chapters of this volume.

We thank all those who have helped to make this edited volume possible, especially the contributors who spent much of their precious time and energy in preparing their chapters. We are really grateful for their enthusiasm and devotion to this project. We thank the contributors and other experts who served as reviewers. Special thanks should go to Dr. Qinghan Xiao, the reviewer assigned by IEEE Press, for providing lots of useful suggestions for the improvement of the book. Our deep feelings of appreciation go to John Wiley & Sons for the impeccable processing of the authors' contributions and the final production of the book. Last, but certainly not least, we

**xii** Preface

would like to thank Jeanne Audino of IEEE Press for her professionalism and her continuous support and assistance during all stages of the preparation and publication of the manuscript.

NIKOLAOS V. BOULGOURIS  
KONSTANTINOS N. PLATANIOTIS  
EVANGELIA MICHELI-TZANAKOU

*London, United Kingdom*  
*Toronto, Ontario, Canada*  
*New Brunswick, New Jersey*  
*July 2009*

# Contributors

---

- ANDY ADLER**, Carleton University, Ottawa, Ontario, Canada
- FOTEINI AGRAFIOTI**, University of Toronto, Toronto, Ontario, Canada
- SAVVAS ARGYROPOULOS**, Informatics and Telematics Institute, Thessaloniki, Greece
- VINEETH BALASUBRAMANIAN**, Arizona State University, Tempe, Arizona
- BIR BHANU**, University of California, Riverside, California
- JOHN BLACK**, Arizona State University, Tempe, Arizona
- OLEG BOULANOV**, University of Calgary, Calgary, Alberta, Canada
- NIKOLAOS V. BOULGOURIS**, King's College, London, United Kingdom
- NIKOLAOS BOURBAKIS**, Wright State University, Dayton, Ohio
- FRANCIS MINHTHANG BUI**, University of Toronto, Toronto, Ontario, Canada
- PATRIZIO CAMPISI**, University of Rome, Rome, Italy
- MARCO CAPARRINI**, Starlab, Barcelona, Spain
- ANN CAVOUKIAN**, Office of Information and Privacy Commissioner Ontario, Toronto, Ontario, Canada
- IVAN CESTER**, Starlab, Barcelona, Spain
- RAMA CHELLAPPA**, University of Maryland, College Park, Maryland
- STELVIO CIMATO**, University of Milan, Milan, Italy
- BRETT DEMARCO**, Rutgers University, New Brunswick, New Jersey
- MARCO GAMASSI**, University of Milan, Milan, Italy
- DIMITRIOS HATZINAKOS**, University of Toronto, Toronto, Ontario, Canada
- XIAXI HUANG**, King's College, London, United Kingdom
- DIMOSTHENIS IOANNIDIS**, Informatics and Telematics Institute, Thessaloniki, Greece
- JOHN M. IRVINE**, Science Applications International Corporation (SAIC), Arlington, Virginia
- STEVEN A. ISRAEL**, Science Applications International Corporation (SAIC), Arlington, Virginia
- SHUIWANG JI**, Arizona State University, Tempe, Arizona
- RAGHUDEEP KANNAVARA**, Wright State University, Dayton, Ohio
- SREEKAR KRISHNA**, Arizona State University, Tempe, Arizona

- B. V. K. VIJAYA KUMAR**, Carnegie Mellon University, Pittsburgh, Pennsylvania  
**ZHEN LEI**, National Laboratory of Pattern Recognition, Institute of Automation,  
Chinese Academy of Sciences, Beijing, China  
**STAN Z. LI**, National Laboratory of Pattern Recognition, Institute of Automation,  
Chinese Academy of Sciences, Beijing, China  
**YUNG-HUI LI**, Carnegie Mellon University, Pittsburgh, Pennsylvania  
**SHENGCAI LIAO**, National Laboratory of Pattern Recognition, Institute of  
Automation, Chinese Academy of Sciences, Beijing, China  
**HAIPING LU**, University of Toronto, Toronto, Ontario, Canada  
**ALESSANDRA LUMINI**, Biometric System Laboratory, University of Bologna,  
Bologna, Italy  
**EMANUELE MAIORANA**, University of Rome, Rome, Italy  
**SOTIRIS MALASSIOTIS**, Informatics and Telematics Institute, Thessaloniki,  
Greece  
**DAVIDE MALTONI**, Biometric System Laboratory, University of Bologna,  
Bologna, Italy  
**ALEX M. MARTINEZ**, Ohio State University, Columbus, Ohio  
**EVANGELIA MICHELI-TZANAKOU**, Rutgers University, New Brunswick,  
New Jersey  
**LORIS NANNI**, Biometric System Laboratory, University of Bologna, Bologna,  
Italy  
**ALESSANDRO NERI**, University of Rome, Rome, Italy  
**DONALD NETH**, Ohio State University, Columbus, Ohio  
**SETHURAMAN PANCHANATHAN**, Arizona State University, Tempe, Arizona  
**SUNG WON PARK**, Carnegie Mellon University, Pittsburgh, Pennsylvania  
**IOANNIS PITAS**, University of Thessaloniki, Thessaloniki, Greece  
**VINCENZO PIURI**, University of Milan, Milan, Italy  
**KONSTANTINOS N. PLATANIOTIS**, University of Toronto, Toronto, Ontario,  
Canada  
**NARAYANAN RAMANATHAN**, University of Maryland, College Park,  
Maryland  
**ALEJANDRO RIERA**, Starlab, Barcelona, Spain  
**AMIT K. ROY-CHOWDHURY**, University of California, Riverside, California  
**GIULIO RUFFINI**, Starlab, Barcelona, Spain  
**ROBERTO SASSI**, University of Milan, Milan, Italy  
**MARIOS SAVVIDES**, Carnegie Mellon University, Pittsburgh, Pennsylvania  
**FABIO SCOTTI**, University of Milan, Milan, Italy  
**VLAD P. SHMERKO**, University of Calgary, Calgary, Alberta, Canada

- AURELI SORIA-FRISCH**, Starlab, Barcelona, Spain
- ALEX STOIANOV**, Office of Information and Privacy Commissioner Ontario, Toronto, Ontario, Canada
- ADRIAN STOICA**, Jet Propulsion Laboratory, NASA, USA
- MICHAEL G. STRINTZIS**, Informatics and Telematics Institute, Thessaloniki, Greece
- XUEJUN TAN**, University of California, Riverside, California
- ANASTASIOS TEFAS**, University of Thessaloniki, Thessaloniki, Greece
- FILARETI TSALAKANIDOU**, Informatics and Telematics Institute, Thessaloniki, Greece
- DIMITRIOS TZOVARAS**, Informatics and Telematics Institute, Thessaloniki, Greece
- ANASTASIOS N. VENETSANOPOULOS**, University of Toronto, Toronto, Ontario, Canada
- KRITHIKA VENKATARAMANI**, Carnegie Mellon University, Pittsburgh, Pennsylvania
- RONG WANG**, University of California, Riverside, California
- RUI WANG**, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China
- BRENDA K. WIEDERHOLD**, Science Applications International Corporation (SAIC), Arlington, Virginia
- MARK D. WIEDERHOLD**, Science Applications International Corporation (SAIC), Arlington, Virginia
- YILEI XU**, University of California, Riverside, California
- SVETLANA N. YANUSHKEVICH**, University of Calgary, Calgary, Alberta, Canada
- JIEPING YE**, Arizona State University, Tempe, Arizona
- RICHARD YOUMARAN**, Carleton University, Ottawa, Ontario, Canada



# Chapter 1

---

## Discriminant Analysis for Dimensionality Reduction: An Overview of Recent Developments

Jieping Ye and Shuiwang Ji

### 1.1 INTRODUCTION

Many biometric applications such as face recognition involve data with a large number of features [1–3]. Analysis of such data is challenging due to the *curse-of-dimensionality* [4, 5], which states that an enormous number of samples are required to perform accurate predictions on problems with a high dimensionality. Dimensionality reduction, which extracts a small number of features by removing irrelevant, redundant, and noisy information, can be an effective solution [6]. The commonly used dimensionality reduction methods include supervised approaches such as linear discriminant analysis (LDA) [7, 8], unsupervised ones such as principal component analysis (PCA) [9], and additional spectral and manifold learning methods [10–14]. When the class label information is available, supervised approaches, such as LDA, are usually more effective than unsupervised ones such as PCA for classification.

Linear discriminant analysis (LDA) is a classical statistical approach for supervised dimensionality reduction and classification [8, 15–18]. LDA computes an optimal transformation (projection) by minimizing the within-class distance and maximizing the between-class distance simultaneously, thus achieving maximum

class discrimination. The optimal transformation in LDA can be readily computed by applying an eigendecomposition on the so-called scatter matrices. It has been used widely in many applications involving high-dimensional data [19–24]. However, classical LDA requires the so-called *total scatter matrix* to be nonsingular. In many applications involving high-dimensional and low sample size data, the total scatter matrix can be singular since the data points are from a very high-dimensional space, and in general the sample size does not exceed this dimension. This is the well-known *singularity or undersampled problem* encountered in LDA.

In recent years, many LDA extensions have been proposed to deal with the singularity problem, including PCA+LDA [19, 23], regularized LDA (RLDA) [21], null space LDA (NLDA) [20], orthogonal centroid method (OCM) [25], uncorrelated LDA (ULDA) [24], orthogonal LDA (OLDA) [24], and LDA/GSVD [26]. A brief overview of these algorithms is given in Section 1.2. Different algorithms have been applied successfully in various domains, such as PCA+LDA in face recognition [19, 23], OCM in text categorization [25], and RLDA in microarray gene expression data analysis [21]. However, there is a lack of a systematic study to explore the commonalities and differences of these algorithms, as well as their intrinsic relationship. This has been a challenging task, since different algorithms apply completely different schemes when dealing with the singularity problem.

Many of these LDA extensions involve an eigenvalue problem, which is computationally expensive to solve especially when the sample size is large. LDA in the binary-class case, called Fisher LDA, has been shown to be equivalent to linear regression with the class label as output. Such regression model minimizes the sum-of-squares error function whose solution can be obtained efficiently by solving a system of linear equations. However, the equivalence relationship is limited to the binary-class case.

In this chapter, we present a unified framework for generalized LDA via a transfer function. We show that various LDA-based algorithms differ in their transfer functions. The unified framework elucidates the properties of various algorithms and their relationship. We then discuss recent development on establishing the equivalence relationship between multivariate linear regression (MLR) and LDA in the multiclass case. In particular, we show that MLR with a particular class indicator matrix is equivalent to LDA under a mild condition, which has been shown to hold for most high-dimensional data. We further show how LDA can be performed in the semisupervised setting, where both labeled and unlabeled data are provided, based on the equivalence relationship between MLR and LDA. We also extend our discussion to the kernel-induced feature space and present recent developments on multiple kernel learning (MKL) for kernel discriminant analysis (KDA).

The rest of this chapter is organized as follows. We give an overview of classical LDA and its generalization in Section 1.2. A unified framework for generalized LDA as well as the theoretical properties of various algorithms and their relationship is presented in Section 1.3. Section 1.4 discusses the least squares formulation for LDA. We then present extensions of the discussion to semisupervised learning and kernel-induced feature space in Sections 1.5 and 1.6, respectively. This chapter concludes in Section 1.8.