

# SpringerBriefs in Computer Science

## *Series Editors*

Stan Zdonik

Shashi Shekhar

Jonathan Katz

Xindong Wu

Lakhmi C. Jain

David Padua

Xuemin (Sherman) Shen

Borko Furht

V.S. Subrahmanian

Martial Hebert

Katsushi Ikeuchi

Bruno Siciliano

Sushil Jajodia

Newton Lee

More information about this series at <http://www.springer.com/series/10028>



Ran He • Baogang Hu • Xiaotong Yuan  
Liang Wang

# Robust Recognition via Information Theoretic Learning

 Springer

Ran He  
National Laboratory  
of Pattern Recognition  
Institute of Automation  
Chinese Academy of Sciences  
Beijing, China

Baogang Hu  
National Laboratory  
of Pattern Recognition  
Institute of Automation  
Chinese Academy of Sciences  
Beijing, China

Xiaotong Yuan  
School of Information  
and Control  
Nanjing University of Information  
Science and Technology  
Nanjing, China

Liang Wang  
National Laboratory  
of Pattern Recognition  
Institute of Automation  
Chinese Academy of Sciences  
Beijing, China

ISSN 2191-5768

ISBN 978-3-319-07415-3

DOI 10.1007/978-3-319-07416-0

Springer Cham Heidelberg New York Dordrecht London

ISSN 2191-5776 (electronic)

ISBN 978-3-319-07416-0 (eBook)

Library of Congress Control Number: 2014944531

© The Author(s) 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*To my parents Guocai He and Shuhua Wang,  
my elder cousin Jie He,  
and my wife Linlin Chi*

*— Ran He*

*To my parents Zhaosen Hu and Huayun Deng  
and my wife Wenzheng Zhang*

*— Baogang Hu*



# Acknowledgements

The completion of this book owes to not only the work of the authors but also many more persons and groups. First and foremost, we would like to thank Dr. Tieniu Tan for always trusting and supporting us. Without his help and understanding, we could not finish this book. We also thank Dr. Jose C. Principe for his 20 years' research efforts in information theoretic learning, which makes this book theoretically solid. We thank our friends, especially Dr. Wei-Shi Zheng, Dr. Zhennan Sun, and Dr. Kaiqi Huang, who showed us the ropes in machine learning and computer vision. Special thanks to Mrs. Hegde Ritya, Dr. Xiang-Wei Kong, Dr. Jing Dong, and Miss. Jessica Y. Zhang for helping us revise this book. Second, we want to give thanks to the assistance provided by Courtney Clark and the publication team at SpringerBriefs. We are very thankful to all the members of CRIPAC ([www.cripac.ia.ac.cn](http://www.cripac.ia.ac.cn)) and NLPR ([www.nlpr.ia.ac.cn](http://www.nlpr.ia.ac.cn)) for their help and hard work. And many thanks to the administrative and support staff (Ling Di, Lei Li, Lixia Ma, Wei Zhao, and Guozhen Lian) who makes living and research easier.

Research efforts summarized in this book were supported by the National Basic Research Program of China (Grant No. 2012CB316300), National Natural Science Foundation of China (Grant No. 61103155, 61135002, 61075051, 61175003), and Nanjing University of Information Science and Technology Faculty Startup (Grant No. S8113029001).





# Contents

<b>1</b>	<b>Introduction</b>	1
1.1	Outline	2
<b>2</b>	<b>M-Estimators and Half-Quadratic Minimization</b>	
2.1	M-Estimation	3
2.2	Half-Quadratic Minimization	4
2.2.1	Iterative Minimization	5
2.2.2	The Additive and Multiplicative Forms	6
2.3	Phase Transition Diagrams	11
2.4	Summary	11
<b>3</b>	<b>Information Measures</b>	13
3.1	Shannon Entropy	13
3.2	Renyi's Quadratic Entropy	18
3.2.1	Robust PCA	18
3.2.2	Robust Discriminant Analysis	24
3.3	Normalized Information Measures and Classification Evaluations	30
3.3.1	Confusion Matrix in Abstaining Classifications	32
3.3.2	Meta-measures in Classification Evaluations	33
3.3.3	Normalized Information (NI) Measures	34
3.4	Summary	44
<b>4</b>	<b>Correntropy and Linear Representation</b>	45
4.1	Correntropy	45
4.1.1	Properties of Correntropy	46
4.1.2	Correntropy Minimization	47
4.2	Correntropy Induced Metric (CIM)	48
4.2.1	CIM and $\ell_0$ -Norm	49
4.2.2	CIM and M-Estimation	50

4.3	Linear Representation .....	51
4.3.1	Linear Least Squares .....	51
4.3.2	Linear Representation Classification .....	52
4.4	Robust Linear Representation via Correntropy .....	54
4.4.1	$\ell_1$ Estimator and Huber M-Estimator .....	55
4.4.2	Parameter Selection .....	57
4.4.3	Stability of Linear Representation Methods .....	58
4.5	Summary .....	60
<b>5</b>	<b><math>\ell_1</math> Regularized Correntropy</b> .....	<b>61</b>
5.1	Sparse Signal Reconstruction .....	61
5.1.1	$\ell_1$ Minimization .....	61
5.1.2	$\ell_1$ -Minimization via Half-Quadratic Optimization .....	62
5.1.3	Numerical Results .....	64
5.2	Robust Sparse Representation .....	67
5.2.1	Error Correction .....	67
5.2.2	Error Detection .....	67
5.3	Robust Sparse Representation via Correntropy .....	68
5.3.1	Error Correction .....	69
5.3.2	Error Detection .....	71
5.3.3	An Active Set Algorithm .....	73
5.4	Numerical Results .....	73
5.4.1	Sparse Representation Algorithms .....	74
5.4.2	Phase Transition Diagrams .....	78
5.4.3	Sunglasses Disguise .....	79
5.4.4	Parameter Setting, Sparsity, and Robustness .....	80
5.5	Summary .....	83
<b>6</b>	<b>Correntropy with Nonnegative Constraint</b> .....	<b>85</b>
6.1	Nonnegative Sparse Coding .....	85
6.2	Robust Nonnegative Sparse Representation .....	87
6.3	Two-Stage Recognition for Large-Scale Problems .....	90
6.3.1	Outlier Detection via Correntropy .....	91
6.3.2	Efficient Nonnegative Sparse Coding .....	94
6.3.3	Two-Stage Sparse Representation .....	95
6.4	Numerical Results .....	97
6.4.1	Sunglasses Disguise .....	98
6.4.2	Scarf Occlusion .....	100
6.4.3	Large-Scale Problems .....	101
6.5	Summary .....	102
	<b>References</b> .....	<b>103</b>

# Acronyms

CESR	Correntropy-based sparse representation
CIM	Correntropy induced metric
CS	Compressed sensing
EM	Expectation maximization
FN	False negative
FP	False positive
GMM	Gaussian mixture model
HQ	Half quadratic
KKT	Karush–Kuhn–Tucker
LDA	Linear discriminant analysis
LPP	Locality preserving projections
MaxEnt	Maximum entropy
MCC	Maximum correntropy criterion
mELE	Minimum entropy linear embedding
MSE	Mean square error
NFL	Nearest feature line
NI	Normalized information
NSR	Nonnegative sparse representation
ICA	Independent component analysis
IRLS	Iteratively reweighted least squares
IP	Information potential
ITL	Information theoretic learning
PCA	Principal component analysis
PDF	Probability density function
RN	Reject negative
RP	Reject positive
SRC	Sparse representation classification
SVD	Singular value decomposition
TN	True negative
TP	True positive
TSR	Two-stage sparse representation