Advances in Pattern Recognition

For other titles published in this series, go to www.springer.com/series/4205

Bir Bhanu • Ju Han

Human Recognition at a Distance in Video



Prof. Bir Bhanu Bourns College of Engineering University of California Riverside, CA 92521 USA bhanu@cris.ucr.edu

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

Dr. Ju Han Lawrence Berkeley National Laboratory University of California Cyclotron Road 1 Berkeley, CA 94720 USA jhan@lbl.gov

ISSN 1617-7916 ISBN 978-0-85729-123-3 e-ISBN 978-0-85729-124-0 DOI 10.1007/978-0-85729-124-0 Springer London Dordrecht Heidelberg New York

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

© Springer-Verlag London Limited 2011

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licenses issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

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

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

Cover design: VTEX, Vilnius

Printed on acid-free paper

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

Preface

Biometrics is the study of methods for uniquely recognizing humans based upon one or more intrinsic physical or behavioral traits, including fingerprint, face, voice, gait, iris, signature, hand geometry, palm, ear, etc. Each biometric has its own relative merits, and the choice of a biometric trait for a particular application depends on a variety of issues. The inherent limitation of a single biometric can be alleviated by fusing the information presented by multiple sources. For example, the face and gait traits, or multiple images of face, or the fingerprints of the right and left index fingers of an individual may be used together to resolve the identity of an individual. A system that consolidates the evidence presented by multiple biometric sources is expected to be more reliable.

Most of the current biometric systems for human recognition generally require a cooperative subject, views from certain aspects and physical contact or close proximity. These systems alone cannot reliably recognize non-cooperating individuals at a distance since it has been difficult to recognize a person from arbitrary views when one is walking at a distance in real-world changing environmental conditions. For optimal performance, a system should make use of as much information as can possibly be obtained from the available observations. Gait and face are the two available biometrics which can be easily captured on a video that is acquired from a distance.

Gait is defined as the manner in which a person walks, and is one of the few biometric traits that can be used to identify non-cooperating humans at a distance. Recent advances in human gait analysis introduced a new application area of recognizing and identifying individuals, their gender, age and ethnicity (soft biometrics) and activities in a surveillance environment for security purposes. Most gait recognition algorithms attempt to extract the human silhouette in order to derive the spatio-temporal attributes of a moving individual. Hence, the selection of a good model to represent the human body is pivotal to the efficient functioning of a gait recognition system. Gait-based systems also offer the possibility of tracking an individual over an extended period of time and performing gait-based recognition in 3D with data available from multiple video streams in a surveillance network. However, gait can be affected by clothing, shoes, or environmental context. Special physical conditions such as injury can also change a person's walking style. The large gait variation of the same person under different conditions (intentionally or unintentionally) reduces the discriminating power of gait as a biometric and it may not be as unique as fingerprint or iris, but the inherent gait characteristic of an individual still makes it irreplaceable and useful for human recognition in practical applications. Gait can be effectively utilized and combined with other biometrics for automatically detecting, recognizing and identifying individuals from a distance.

Face recognition is non-intrusive, and facial attributes are one of the most common features used to recognize an individual. The applications of facial recognition vary from a static, controlled "mug-shot" authentication to a dynamic, uncontrolled face identification in a cluttered background. Most of the face recognition approaches are based on either the location and shape of facial attributes, such as the eyes, eyebrows, nose, lips, and chin and their spatial relationships, or the overall analysis of the face image that represents a face as a weighted combination of a number of canonical faces. While face recognition is conveniently applied, it is easily affected by several factors including illumination, expression, pose, etc. It is also questionable whether the face itself has a sufficient basis for recognizing a person from a very large number of individuals with an extremely high level of confidence.

The general solution to analyze face and gait video data collected from arbitrary views is to estimate 3-D models. However, the problem of building reliable 3-D models of face and gait with non-rigid face, flexible neck and the articulated human body from low resolution video data is a challenging task. In this book, integrated face and gait recognition approaches are developed that exploit inherent characteristics of human signatures in video that is captured from a distance. Experimental results show the effectiveness of the current systems for human recognition at a distance in video.

This book addresses fundamental problems associated with gait, face and integrated gait and face based human recognition in color and infrared videos acquired at a distance under real-world environments.

For gait-based human recognition the book addresses the problems associated with the representation, the large intra-person variation of gait appearance under different environmental conditions, the lack of discrimination analysis for gait-based human recognition, and the difficulties associated with reliable moving human detection in various situations. Both model-free and model-based approaches are considered for individual recognition under varying contextual, environmental and carrying conditions. This includes the newly developed techniques where the both the model and the data (obtained from multiple cameras) are in 3D. Bayesian-based statistical analysis is performed to evaluate the discriminating power of gait features for human recognition. To improve the performance of moving human detection for both model-free and model-based human recognition, the information from color and infrared videos is combined using automatic image registration methods.

For face recognition in video with people at a distance, the challenges are precise registration of faces in low resolution video data and the robustness of superresolution techniques to variations in pose, lighting, facial expression and the number of video frames. The book presents new video-based techniques for face profilebased recognition and three techniques for super-resolution of frontal and side facial imagery acquired from a video. These techniques are based on (a) closed-loop tracking, (b) free-form deformation, and (c) elastic registration. Objective measures, to evaluate the quality of super-resolved imagery for face recognition, are presented based on different conditions encountered during the video capture.

For integrated gait and face biometrics the challenges are the development of effective techniques at different levels of abstraction. The book presents several systems that integrate information of the side view of face and gait from video data. Several fusion schemes are introduced at the match score and feature levels for the integration of super-resolved face and gait. Both face and gait recognition systems integrate information over multiple frames in a video sequence for improved performance.

The first author would like to thank his Ph.D. students Ju Han, Xioli Zhou, Xiaotao Zou, Jiangang Yu, and Koichiro Yamauchi whose graduate research is incorporated in various ways in this book. The authors would also like to thank Jhon Gonzalez for technical support and Suresh Kumar for reviewing the document.

This research work was supported in part by various grants from NSF, ARO, ONR, and AFOSR over the years. The authors are grateful to these funding agencies. Finally, the authors would also like to thank Wayne Wheeler at Springer for his interest and support.

Riverside, CA, USA Berkeley, CA, USA Bir Bhanu Ju Han

Contents

Part I Introduction to Gait-Based Individual Recognition at a Distance

1	Intro	oductio	n	3
	1.1	Key Id	eas Described in the Book	5
	1.2	Organi	zation of the Book	7
Part	t II	Gait-B	ased Individual Recognition at a Distance	
2	Gait	Repres	sentations in Video	13
	2.1	Humar	n Motion Analysis and Representations	13
	2.2	Humar	n Activity and Individual Recognition by Gait	14
		2.2.1	Human Recognition by Gait	15
		2.2.2	Human Activity Recognition	17
	2.3	Gait E	nergy Image (GEI) Representation	17
		2.3.1	Motivation	18
		2.3.2	Representation Construction	18
		2.3.3	Relationship with MEI and MHI	18
		2.3.4	Representation Justification	19
	2.4	Frame	work for GEI-Based Recognition	21
		2.4.1	Silhouette Extraction and Processing	21
		2.4.2	Feature Extraction	22
	2.5	Summa	ary	24
3	Mod	lel-Free	Gait-Based Human Recognition in Video	25
	3.1	Statisti	ical Feature Fusion for Human Recognition by Gait	25
		3.1.1	Real and Synthetic Gait Templates	26
		3.1.2	Human Recognition	28
		3.1.3	Experimental Results	30
	3.2	Humar	n Recognition Based on Environmental Context	33
		3.2.1	Walking Surface Type Detection	34
		3.2.2	Classifier Design	37

		3.2.3 Ex	perimental Results	39
	3.3	View-Inse	nsitive Human Recognition by Gait	40
		3.3.1 Vi	ew-Insensitive Gait Templates	40
		3.3.2 Hu	man Recognition	42
		3.3.3 Ex	perimental Results	43
	3.4	Human Re	epetitive Activity Recognition in Thermal Imagery	45
		3.4.1 Ot	bject Detection in Thermal Infrared Imagery	46
		3.4.2 Hu	uman Repetitive Activity Representation	
		an	d Recognition	47
		3.4.3 Ex	perimental Results	48
	3.5	Human Re	ecognition Under Different Carrying Conditions	50
		3.5.1 Te	chnical Approach	50
		3.5.2 Ex	perimental Results	53
	3.6	Summary		55
4	Disc	rimination	Analysis for Model-Based Gait Recognition	57
	4.1	Predicting	Human Recognition Performance	57
	4.2	Algorithm	Dependent Prediction and Performance Bounds	58
		4.2.1 Bo	bdy Part Length Distribution	58
		4.2.2 Al	gorithm Dependent Performance Prediction	60
		4.2.3 Up	oper Bound on PCR	61
	4.3	Experimen	ntal Results	62
	4.4	Summary		63
5	Moo	lel-Based F	Juman Recognition—2D and 3D Gait	65
-	5.1	2D Gait R	ecognition (3D Model, 2D Data)	65
		5.1.1 3D	Human Modeling	66
		5.1.2 Hu	Iman Recognition from Single Non-calibrated	
		Ca	mera	70
		5.1.3 Hu	uman Recognition from Multiple Calibrated	
		Ca	meras	76
	5.2	Gait Reco	gnition in 3D	80
		5.2.1 Inc	dividual Recognition by Gait in 3D	80
		5.2.2 Re	lated Work	81
		5.2.3 Te	chnical Approach	83
		5.2.4 Ex	perimental Results	89
	5.3	Summary		94
6	Fusi	on of Colo	r/Infrared Video for Human Detection	95
	6.1	Related W	⁷ ork	97
	6.2	Hierarchic	al Image Registration and Fusion Approach	99
		6.2.1 Im	age Transformation Model	100
		6.2.2 Pro	eliminary Human Silhouette Extraction	
		an	d Correspondence Initialization	101

	6.2.3	Automatic Image Registration
	6.2.4	Sensor Fusion
	6.2.5	Registration of EO/IR Sequences with Multiple
		Objects
6.3	Exper	imental Results
	6.3.1	Image Registration Results
	6.3.2	Sensor Fusion Results
6.4	Summ	nary

Part III Face Recognition at a Distance in Video

7	Sup	er-Reso	olution of Facial Images in Video at a Distance
	7.1	Closed	d-Loop Super-Resolution of Face Images in Video 118
		7.1.1	Related Work
		7.1.2	Technical Approach
		7.1.3	Experimental Results
	7.2	Super	Resolution of Facial Images with Expression Changes
		in Vid	eo
		7.2.1	Related Work
		7.2.2	Technical Approach
		7.2.3	Experimental Results
	7.3	Const	ructing Enhanced Side Face Images from Video
		7.3.1	Enhanced Side Face Image (ESFI) Construction 139
		7.3.2	Technical Approach
	7.4	Summ	nary
8	Eva	luating	Quality of Super-Resolved Face Images
	8.1	Image	Quality Indices
	8.2	Integr	ated Image Quality Index
		8.2.1	Gray Scale Based Quality (Q_g)
		8.2.2	Structure Based Quality (Q_e)
		8.2.3	Similarity Between Input Images (Q_i)
		8.2.4	Integrated Quality Measure (Q_{int})
	8.3	Exper	imental Results for Face Recognition in Video
		8.3.1	Experiment 1: Influence of Pose Variation on the Super-
			Resolved Face Image
		8.3.2	Experiment 2: Influence of Lighting Variation
			on the Super-Resolved Face Image
		8.3.3	Experiment 3: Influence of Facial Expression Variation
			on the Super-Resolved Face Image
		8.3.4	Experiment 4: Influence of the Number of Images Used
			for Constructing the Super-Resolved Face Image for Face
			Recognition
		8.3.5	Discussion
	8.4	Summ	nary

Part IV	Integrated Face and (Gait for Human	Recognition at a	Distance
in V	ideo			

9	Integ	grating Face Profile and Gait at a Distance	57
	9.1	Introduction	57
	9.2	Technical Approach	59
		9.2.1 High-Resolution Image Construction for Face Profile 16	59
		9.2.2 Face Profile Representation and Matching	73
		9.2.3 Gait Recognition	78
		9.2.4 Integrating Face Profile and Gait for Recognition	
		at a Distance	79
	9.3	Experimental Results	79
		9.3.1 Face Profile-Based Recognition	79
		9.3.2 Integrating Face Profile With Gait	31
	9.4	Summary	34
10	Mat	ch Score Level Fusion of Face and Gait at a Distance	35
	10.1	Introduction	36
	10.2	Related Work	37
	10.3	Technical Approach	38
		10.3.1 Enhanced Side Face Image Construction	39
		10.3.2 Gait Energy Image Construction) 0
		10.3.3 Human Recognition Using ESFI and GEI) 0
	10.4	Experimental Results and Performance Analysis) 3
		10.4.1 Experiments and Parameters) 3
		10.4.2 Performance Analysis)1
	10.5	Summary)6
11	Foot	ure Level Fusion of Face and Cait at a Distance	იი
11	11 1	Introduction)9)0
	11.1	Technical Approach	12
	11.2	11.2.1 Human Identification Using ESEL and CEL	12
	112	The Deleted Eusion Schemes	14
	11.5	11.2.1 Eusion at the Match Score Level [200]	10
		11.3.1 Fusion at the Fracture Level [209]	10
	11 /	11.5.2 Fusion at the Feature Level [207]	10
	11.4	Experimental Results and Comparisons	18
		11.4.1 Experiments and Parameters	18
	11 7	11.4.2 Discussion on Experiments	50
	11.5	Summary	32
Part	t V Reco	Conclusions for Integrated Gait and Face for Human ognition at a Distance in Video	
12	Con	clusions and Future Work	35

12.1 Summary
12.1.1 Gait-Based Human Recognition at a Distance

Contents

		12.1.2 12.1.3	Vid Fus	eo-H ion	Base of H	ed Fac	Hu e a	ım and	an G	Ro	ecc t fo	ogi or 1	nit Hu	ioı ım	n a nai	at a 1 R	ı E lec	Dis cos	ta 2n	nc iti	e or	1	•	 •	•	236
			at E	Dista	nce																					237
1	12.2	Future	Res	earc	h D	ire	cti	on	s	•					•		•	•	•	•		•	•		•	238
Refer	ence	s				•			•						•			•	•		•		•	 •	•	241
Index	ζ																									251

List of Figures

Fig. 2.1	Examples of normalized and aligned silhouette frames in different human walking sequences. The rightmost image	
	<i>in each row</i> is the corresponding gait energy image (GEI)	19
Fig. 2.2	Diagram of the proposed GEI-based framework for the purpose	
	of recognition	21
Fig. 2.3	Frequency and phase estimation of human walking	22
Fig. 3.1	System diagram of human recognition using the proposed	
	statistical feature fusion approach	26
Fig. 3.2	An example of real gait template set generated from a long	
	silhouette sequence of an individual	27
Fig. 3.3	Procedure of generating synthetic GEI templates from an	
	original template. The <i>leftmost</i> template is the original template,	
	other templates are generated by gradually cutting the bottom	
	portion (templates in the first row) and fitting it to the original	
	template size (templates in the first row). Synthetic templates	
	in the second row are used as the synthetically generated gait	
	templates	27
Fig. 3.4	Pseudo code for generating synthetic GEI templates	28
Fig. 3.5	GEI examples in USF HumanID database	30
Fig. 3.6	Context-based classifier combination	34
Fig. 3.7	Diagram of context-based classifier combination for individual recognition by gait. The context investigated in this diagram is	
	the walking surface type	35
Fig. 3.8	GEI examples of three people (rows) walking on different	
	surface types. First four examples in each row are on the grass	
	surface, and <i>the others</i> are on the concrete surface	35
Fig. 3.9	(a) The Bhattacharyya distance of the two distributions with	
	respect to different N_{TOP} values. (b) The estimated distributions	
	of $p(s \text{grass})$ and $p(s \text{concrete})$ for $N_{\text{TOP}} = 6$	36
Fig. 3.10	Examples of human walking along different directions	40
Fig. 3.11	The graphic illustration (<i>from the top view</i>) of human walking along different directions	41
		• 1

Fig. 3.12	GEIs generated from three sequences of the same person	40
Fig. 3.13	Experiment design: <i>Left</i> —walking left; <i>Right</i> —walking right; positive angle—walking toward the camera; negative	42
Fig. 3.14	angle—walking away from the camera (all angles in $^{\circ}$) An example of human walking at different time of a day recorded from a thermal infrared sensor: noon (<i>first row</i>), late	43
Fig. 3.15	afternoon (second row) and night (third row)Examples of normalized and aligned silhouette frames indifferent human motion sequences. The rightmost image in each	45
Fig. 3.16	<i>row</i> is the corresponding gait energy image (GEI) Thermal images of the background at different time of a day recorded from the thermal infrared sensor: noon, late afternoon and night, each of which is normalized by the temperature range	47
Fig. 3.17	GEI examples of the 54 human motion sequences used in our	48
Fig. 3.18	experiments	49
Fig. 3.19	of the people without a briefcase (<i>below</i>)	51
Fig. 3.20	(b) testing phase	52
Fig. 4.1	Anthropometric estimates of individual body part length distribution for British adults 19–65 years [136]	59
Fig. 4.2	Uncertainty computation for the given silhouette resolution r (in mm/pixel)	61
Fig. 4.3	Predicted PCR with regard to different database size and different within-class standard deviation values of the features	60
Fig. 4.4	Predicted upper bound on PCR with regard to different database size and different human silhouette resolution values	63
Fig 51	3D human kinematic model	67
Fig. 5.2	Body part geometric representation	67
Fig. 5.3	Camera geometry with perspective projection and radial lens distortion [168]	68
Fig. 5.4	Diagram of the proposed approach for human gait analysis	72
Fig. 5.5	Sample sequences in our gait database recorded from single camera view	74
Fig. 5.6	Sample frames of three camera views in Mobo database	79
Fig. 5.7	Gait cycle expressed by swing phase and stance phase	83
Fig. 5.8	3D human body model. (a) Tapered cylinder with two angular DOF. (b) Body model approximated by 12 segments.	
	(c) Hierarchical structure	84
Fig. 5.9	Body axes estimation by PCA	86

Fig. 5.10	3D human body data. Representative poses of walking	00
E. 511	$\mathbf{F}'_{44} = 1 2 \mathbf{D}_{44} 1 1 1 1 1 1 1 1$	90
Fig. 5.11	Fitted 3D models to four poses of gait cycle	91
Fig. 5.12	Gait sequence composed of 20 frames recovered	0.1
	by four poses	91
Fig. 5.13	Examples of training data for six subjects	92
Fig. 5.14	Examples of testing data for six subjects	93
Fig. 6.1	Human silhouette extraction results from color images (<i>first two</i> rows) and thermal images (<i>last two rows</i>) using the background subtraction method with increasing thresholds from left to right. The leftmost image is the original image for each row	96
Fig. 6.2	Proposed hierarchical genetic algorithm based multi-modal	
U	image registration approach	99
Fig. 6.3	IR and EO camera set-up	100
Fig. 6.4	Different object appearances in color and thermal images are due to the phenomenological differences between the image formation process of EO and IR cameras. The images are at different resolutions and the field-of-views (FOVs)	
	of the two cameras overlap (EO camera FOV contains	
	IR camera FOV)	103
Fig. 6.5	Illustration of the HGA-based search scheme to estimate the model parameters—coordinate values of the two points (<i>black and white</i>) simultaneously in the color image plane. The search windows here are only for illustration whose sizes are much	
	larger than the real window sizes	105
Fig. 6.6	Illustration of initial control point selection in both color and thermal image planes	106
Fig 67	Pseudo code for parameter estimation based on hierarchical	100
119.0.7	genetic algorithm	107
Fig. 6.8	Examples of registration results: <i>first row</i> —original color images. <i>second row</i> —original thermal images. <i>third</i>	107
	<i>row</i> —transformed color images	109
Fig. 6.9	Examples of estimated transformation results from good initial correspondence: (a) original silhouettes from thermal images, (b) original silhouettes from color images, (c) matching error of the initial transformation, (d) matching error after the first search level, (e) after the second search level, (f) after the third search level, (g) after the fourth search level, (h) after the 12th search level	110
Fig. 6.10	Examples of estimated transformation results from bad initial correspondence: (a) original silhouettes from thermal images, (b) original silhouettes from color images, (c) matching error of the initial transformation, (d) matching error after the first search level, (e) after the second search level, (f) after the third search level, (g) after the fourth search level, (h) after the 23th	110
	search level	111

List of Figures
Variation of fitness values from bad initial correspondence. The <i>vertical line</i> corresponds to the last generation at each search
level. The curve between two adjacent vertical lines indicate the
variation of GA fitness values at a search level
Variation of fitness values from good initial correspondence.
The vertical line corresponds to the last generation at each
search level. The curve between two adjacent vertical lines
indicate the variation of GA fitness values at a search level 112
ROC curves for detection performance evaluation of different
fusion strategies for silhouette detection
Examples of fusion results: (a) transformed color images.
(b) original thermal images. (c) silhouette from (a). (d) silhouette
from (b), (e) silhouette from product rule fusion, (f) silhouette

	from sum rule fusion, (g) silhouette from max rule fusion,
	(h) silhouette from min rule fusion
Fig. 7.1	Block diagram of the closed-loop approach for super-resolution
	of facial images
Fig. 7.2	Block diagram of the super-resolution algorithm
Fig. 7.3	Results on synthetic video with ground-truth poses. The first row
	shows the original low-resolution (LR) frames, and the second
	row shows the bicubically interpolated ones. Reconstructed SR
	images are shown in the third row. The last row shows pose and

Fig. 7.4 Evaluation of super-resolved video as measured by the peak Fig. 7.5 Real video with expression changes. The first row shows the original LR frames, and the second row shows the reconstructed

illumination normalized, reconstructed SR images with respect

	ones using the global method. The third row shows the SR
	images of our locally-based method
Fig. 7.6	An example of facial images with expression changes:
	(a) high-resolution (92×73) images, (b) low-resolution
	(30×24) images $\ldots \ldots 124$
Fig. 7.7	Block diagram of our approach
Fig. 7.8	Tracking results for one sequence. The first row shows the video
	frames. The second row shows the tracked facial regions 127
Fig. 7.9	The resolution aware incremental free form deformation: (a) the
	input LR image, (b) bicubically interpolated image of (a), (c) the
	reference LR image, (d) the interpolated image (b) overlaid with
	control lattice, (e) the deformed LR image
Fig. 7.10	Examples of our resolution aware IFFD local registration:
-	

Fig. 7.10	Examples of our resolution aware IFFD local registration:
	(a) source frames which need to be deformed, (b)–(c) illustration
	of the deformation process from coarse to fine control lattice,
	(d) representation of the final deformed LR images warped
	to the reference frames, (e) representation of the reference
	frames

Fig.

Fig.

Fig.

Fig.

Fig. 7.11	Super-resolution results: (a)-(b) low-resolution images,
	(c) reconstructed SR images using global registration,
	(d) reconstructed SR images using global and our RAIFFD local
	deformation approach
Fig. 7.12	Super-resolution results: (a)-(b) low-resolution images,
	(c) reconstructed SR images using global registration,
	(d) reconstructed SR images using global and our RAIFFD local
	deformation approach
Fig. 7.13	Marked regions for calculating peak signal-to-noise ratio
	(PSNR): (a) reconstructed SR image using global registration,
	(b) reconstructed SR image using global + local deformation
	approach, (c) original high-resolution image
Fig. 7.14	Comparison of PSNR values between globally reconstructed SR
8	images and our globally + locally reconstructed SR images \dots 136
Fig. 7.15	Super-resolution results using SR algorithms (I): (a) SR results
8	of globally aligned data using the method of [43]. (b) SR results
	of globally $+$ locally aligned data using the method of [43]
	(c) SR results of globally aligned data using the method of
	[212] (d) SR results of globally \pm locally aligned data using the
	method of [212] 137
Fig. 7.16	Super-resolution results using SR algorithms (II): (a) SR results
11g. 7.10	of globally aligned data using the method of $[A3]$ (b) SR results
	of globally \perp locally aligned data using the method of [43].
	(a) SP results of globally aligned data using the method of
	(c) SK results of globally angled data using the method off
	[212], (d) SR results of globally + locally aligned data using the
F' 7 17	method of [212]
F1g. /.1/	Comparison of PSNR values between globally reconstructed
	SR images and our globally + locally reconstructed SR images
	using the SR methods in $[43, 212]$
Fig. 7.18	Pseudo code for low-resolution image alignment
Fig. 7.19	Two examples of alignment results with the match statistic <i>S</i> :
	the reference image (<i>left</i>), the image to be aligned (<i>middle</i>) and
	the aligned image $(right)$
Fig. 7.20	Pseudo code for high-resolution image construction 145
Fig. 7.21	Four examples of resized low-resolution face images (top) and
	constructed high-resolution face images (bottom)
Fig. 7.22	The extracted face profile and the absolute values of the
	curvature
Fig. 7.23	Examples of 4 people: (a) resized OSFIs, (b) ESFIs 147
Fig. 8.1	Quality evaluation of a super-resolved image
Fig. 8.2	Three input face images with the same pose
Fig. 8.3	Three input face images with different poses
Fig. 8.4	(a) The super-resolved face image from the input images in
-	Fig. 8.2. (b) The super-resolved face image from the input
	images in Fig. 8.3. (c) The ideal reference image shown for
	comparison. It is directly obtained from the original video 157

Fig. 8.5	Three input face images with the same lighting conditions 158
Fig. 8.6	Three input face images with different lighting conditions 158
Fig. 8.7	(a) The super-resolved face image from the input images in
•	Fig. 8.5. (b) The super-resolved face image from the input
	images in Fig. 8.6. (c) The ideal reference image shown for
	comparison. It is directly obtained from the original video 159
Fig. 8.8	Three input face images with the same expression
Fig. 8.9	Three input face images with different expressions
Fig. 8.10	(a) The super-resolved face image from the input images in
•	Fig. 8.8. (b) The super-resolved face image from the input
	images in Fig. 8.9. (c) The ideal reference image shown for
	comparison. It is directly obtained from the original video 160
Fig. 8.11	Sample low-resolution input face images taken from videos of
-	45 people. They are numbered 1 to 45 (from left to right and top
	<i>to bottom</i>)
Fig. 8.12	Examples of super-resolved face images of people from number
	13 to 24. They are constructed from nine low-resolution input
	images
Fig. 8.13	Results from 90 video sequences of 45 people: (a) recognition
	rate vs. number of input images, (b) quality vs. number of input
	images
Fig. 9.1	Technical approach for integrating face profile and gait
	in video
Fig. 9.2	Six low-resolution face profile images resized by using bilinear
	interpolation (a–f)
Fig. 9.3	The edge images of six low-resolution face profile images
	shown in Fig. 9.2
Fig. 9.4	The reconstructed high-resolution face profile and
	its edge image
Fig. 9.5	Technical approach for face profile recognition
Fig. 9.6	The extracted face profile and the absolute values of curvature 176
Fig. 9.7	Four examples of curvature features on face profiles 17
Fig. 9.8	The similarity matrix (<i>left</i>) and the dynamic programming
	matrix ($right$)
Fig. 9.9	Examples of face profile images from the University
	of Bern
Fig. 9.10	Examples of face profile images from the University
	of Stirling
Fig. 9.11	Four examples of video sequences
Fig. 9.12	GEIs of two people misclassified by the gait classifier. For each
	person, the training GEI and the testing GEI are shown for
	comparison
Fig. 9.13	Face profile of two people misclassified by the gait classifier.
	For each person, the training profile and the testing profile are
	shown for comparison

Fig. 10.1	Technical approach for integrating side face and gait	188
Fig. 10.2	(a) One resized low-resolution face image $(left)$ and one reconstructed high-resolution face image $(right)$. (b) Resized	100
Fig. 10.3	OSFI (<i>left</i>) and ESFI (<i>right</i>)	189
$\Gamma = 10.4$	Image (GEI)	190
Fig. 10.4	Data in Experiment 1. Video sequences 1 through 23	195
Fig. 10.5 Fig. 10.6	Experiment 1. Video sequences 24 through 45 Experiment 1. GEIs of people misclassified by the gait classifier (see Table 10.3). For each person, two GEIs of the training video sequence and two GEIs of the testing video sequence are	190
Fig. 10.7	shown for comparison	197
	for comparison	198
Fig. 10.8	Data in Experiment 2. 10 updated video sequences {1, 2, 5, 6, 8, 9, 10, 13, 19, 40}	198
Fig. 10.9	Experiment 2. GEIs of people misclassified by the gait classifier (see Table 10.5). For each person, two GEIs of the training video sequence and two GEIs of the testing video sequence are shown for comparison	200
Fig. 10.10	Experiment 2. ESFIs of people misclassified by the face classifier (see Table 10.5). For each person, two ESFIs of the training video sequence and two ESFIs of the testing video sequence are shown for comparison	201
Fig. 10.11	Experiment 2. People misclassified by the integrated classifier based on ESFI and GEI using different fusion rules (see Table 10.6). For each person, one frame of the training video sequence and one frame of the testing video sequence are shown	201
Fig. 10.12	for comparison	202
Fig. 10.13	Experiment 2. (a) Correlation of the normalized match scores of the two classifiers using GEI and OSFI. (b) Correlation of the normalized match scores of the two classifiers using GEI and ESFI	205
Fig. 11.1	The basic processes of the fusion schemes for comparison	213
Fig. 11.2	The proposed feature level fusion scheme for integrating side face and gait in video	214

Fig. 11.3	Two examples of video sequences	218
Fig. 11.4	The top 70 eigenvectors of face (from left to right and top to	220
Fig. 11.5	First 70 eigenvectors of gait (<i>from left to right</i> and	220
Fig. 11.6	top to bottom)	220
Fig. 11.7	in Experiment 2	221
Fig. 11.8	Experiment 1. ESFIs of people misclassified by the face classifier (see Table 11.3). For each person, two ESFIs of the training video sequence and two ESFIs of the testing video	223
Fig. 11.9	sequence are shown for comparison	223
Fig. 11.10	for comparison	224
Fig. 11.11	and ESFI	225
Fig. 11.12	shown for comparison	226
Fig. 11.13	sequence are shown for comparison	227
Fig. 11.14	for comparison	228 229
		/

List of Tables

Table 3.1	Twelve experiments designed for human recognition on USF HumanID database	30
Table 3.2	Comparison of recognition performance (rank 1) using different approaches on silhouette sequence version 1.7. (Legends: USF—direct frame shape matching [150]; CMU—key frame shape matching [166]; UMD—HMM framework [85]; Real—proposed real gait feature classifier only;	50
	Synthetic—proposed synthetic gait feature classifier only;	21
Table 3.3	Comparison of recognition performance on silhouette sequence version 2.1. (Legends: USF—direct frame shape matching	31
	[150]; Real—proposed real gait feature classifier only;	
	Synthetic—proposed synthetic gait feature classifier only;	
	Fusion—proposed gait feature fusion)	31
Table 3.4	Comparison of recognition performance (rank 5) using different	
	approaches on silhouette sequence version 1.7. (Legends:	
	USF—direct frame shape matching [150]; CMU—key	
	frame shape matching [166]; UMD—HMM framework	
	[85]; Real—proposed real gait feature classifier only;	
	Synthetic—proposed synthetic gait feature classifier only;	22
TT 1 1 2 7	Fusion—proposed gait feature fusion)	33
Table 3.5	Comparison of recognition performance using different	
	approaches on sinouelle sequence version 2.1. (Legends:	
	Baseline—USF baseline algorithm [157]; Real—Ieal	
	context based approach)	30
Table 3.6	Recognition performance using sequences of one specific	57
1000 5.0	direction for training (all performance in %)	44
Table 3.7	Recognition performance using sequences of two specific	
	directions for training	44
Table 3.8	An example of Hu's moment for the GEI of a human with a	
	briefcase and without a briefcase	51

xxiii

Table 3.9	Description of experimental data
Table 4.1	Anthropometric estimates of individual body part length
	distribution for British adults 19–65 years [136]
Table 4.2	Anthropometric estimates of their correlation coefficients
	Australia elderly population [90]
Table 4.3	Resolution (mm/pixel) for a 1675 mm (population average
	height) person occupying different vertical portions of the
	frame with different video formats
Table 5.1	Comparison of performance using different number of
	stationary features
Table 5.2	Comparison of performance using mean and standard deviation
	features computed from each body part angle variation
	sequences over a single-cycle sequence
Table 5.3	Comparison of performance using different number of
	kinematic features
Table 5.4	Comparison of performance using different combination
	strategies
Table 5.5	Summary of 3D biometrics approaches for human
	recognition
Table 5.6	Identification rate and average pose error
Table 6.1	Confusion matrix
Table 7.1	Sample work for face SR (L-learning-based method,
	R-reconstruction-based method)
Table 8.1	A few recent works for the image quality evaluation vs. the
	proposed approach
Table 8.2	The effect of pose variation on quality
Table 8.3	The effect of light variation on quality
Table 8.4	The effect of facial expression variation on quality
Table 8.5	The effect of the number of input images on the quality of
	super-resolved image
Table 9.1	Experimental results for face profile recognition
Table 9.2	Experimental results for integrated face profile and gait 182
Table 10.1	Approaches for integrating face and gait for human
	recognition
Table 10.2	Summary of three experiments 194
Table 10.3	Experiment 1. Single biometric performance and error index of
1000 1000	individuals 196
Table 104	Experiment 1. Fused biometric performance and error index of
10010 10.1	individuals 107
Table 10.5	Experiment 2 Single biometric performance and error index of
1010 1010	individuals

Table 10.8	Experiment 1. <i>Q</i> statistics
Table 10.9	Experiment 2. <i>Q</i> statistics
Table 11.1	The recent work for feature level fusion
Table 11.2	The related work for integrating face and gait for human
	recognition vs. this chapter
Table 11.3	Experiment 1. Single biometrics performance and error index
	of individuals
Table 11.4	Experiment 1. Fused biometrics performance and error index of
	individuals
Table 11.5	Experiment 2. Single biometrics performance and error index
	of individuals
Table 11.6	Experiment 2. Fused biometrics performance and error index of
	individuals