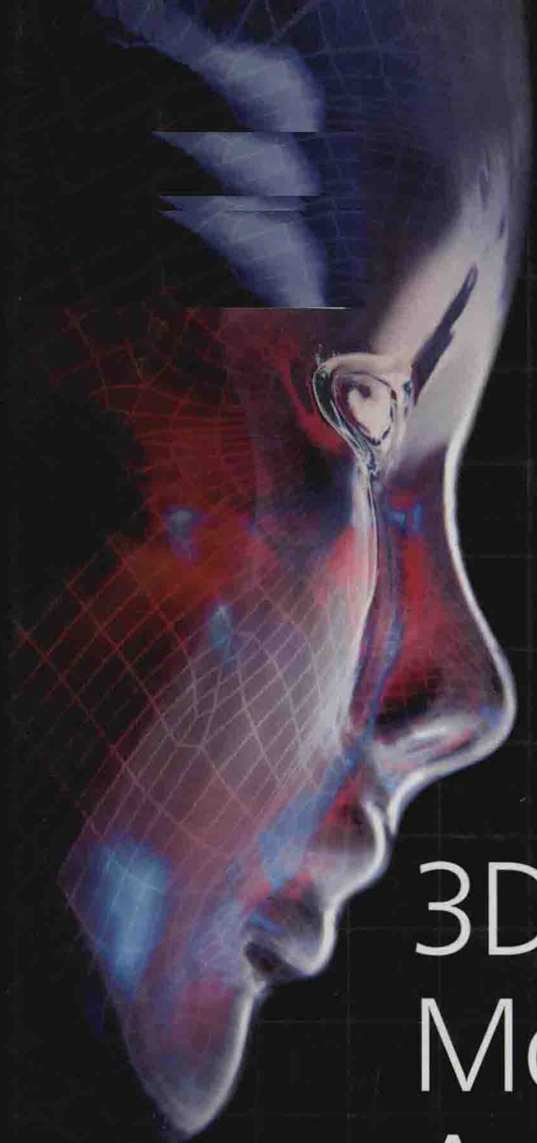

Editors

Mohamed Daoudi

Anuj Srivastava

Remco Veltkamp

3D Face Modeling, Analysis and Recognition

WILEY

3D FACE MODELING, ANALYSIS AND RECOGNITION

Editors

Mohamed Daoudi
TELECOM Lille 1/LIFL, France

Anuj Srivastava
Florida State University, USA

Remco Veltkamp
Utrecht University, The Netherlands

WILEY

This edition first published 2013

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication Data

Daoudi, Mohamed, 1964–
3D face modeling, analysis, and recognition / Mohamed Daoudi, Anuj Srivastava, Remco Veltkamp.
pages cm
Includes bibliographical references and index.
ISBN 978-0-470-66641-8 (cloth)
1. Three-dimensional imaging. 2. Human face recognition (Computer science) 3. Face–Computer simulation. I. Srivastava, Anuj, 1968– II. Veltkamp, Remco C., 1963– III. Title. IV. Title: Three dimensional face modeling, analysis, and recognition.
TA1637.D365 2013
006.6′93–dc23

2013005799

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

ISBN: 9780470666418

Typeset in 10/12pt Times by Aptara Inc., New Delhi, India
Printed and bound in Malaysia by Vivar Printing Sdn Bhd

1 2013

3D FACE MODELING, ANALYSIS AND RECOGNITION

Preface

Introduction

The human face has long been an object of fascination, investigation, and analysis. It is so familiar to our visual cognition system that we can recognize a person's face in difficult visual environments, that is, under arbitrary lighting conditions and pose variation. A common question to many researchers is whether a computer vision system can process and analyze 3D face as the human vision system does. In addition to understanding human cognition, there is also increasing interest in analyzing shapes of facial surfaces for developing applications such as biometrics, human–computer interaction (HCI), facial surgery, video communications, and 3D animation.

Because facial *biometrics* is natural, contact free, nonintrusive, and of psychological interest, it has emerged as a popular modality in the biometrics community. Unfortunately, the technology for 2D image-based face recognition still faces difficult challenges. Face recognition is made difficult by data variability caused by pose variations, lighting conditions, occlusions, and facial expressions. Because of the robustness of 3D observations to lighting conditions and pose variations, face recognition using shapes of facial surfaces has become a major research area in the last few years. Many of the state-of-the-art methods have focused on the variability caused by facial deformations, for example, those caused by face expressions, and have proposed methods that are robust to such shape variations.

Another important use of 3D face analysis is in the area of *computer interaction*. As machines become more and more involved in everyday human life and take on increasing roles in both their living and work spaces, they need to become more intelligent in terms of understanding human moods and emotions. Embedding these machines with a system capable of recognizing human emotions and mental states is precisely what the HCI research community is focused on. Facial expression recognition is a challenging task that has seen a growing interest within the research community, impacting important applications in fields related to HCI. Toward building human-like emotionally intelligent HCI devices, scientists are trying to include identifiers of the human emotional state in such systems. Recent developments in 3D acquisition sensors have made 3D data more readily available. Such data help alleviate problems inherent in 2D data such as illumination, pose, and scale variations as well as low resolution.

The interest in 3D facial shape analysis is fueled by the recent advent of cheaper and lighter scanners that can provide high resolution measurements of both geometry and texture of human facial surfaces. One general goal here is to develop computational tools for analyzing 3D face data. In particular, there is interest in quantifiably comparing the shapes of facial surfaces. This can be used to recognize human beings according to their facial shapes, to measure changes in a facial shape following a surgery, or to study/capture the variations in facial shapes during conversations and expressions of emotions. Accordingly, the *main theme of this book is to develop computational frameworks for analyzing shapes of facial surfaces*. In this book, we use some basic and some advanced tools from differential geometry, Riemannian geometry, algebra, statistics, and computer science to develop the desired algorithms.

Scope of the book

This book, which focuses on 3D face modeling, processing, and applications, is divided into five chapters.

Chapter 1 provides a brief overview of successful ideas in the literature, starting with some background material and important basic ideas. In particular, the principles of depth from triangulation and shape from shading are explained first. Then, an original 3D face (static or dynamic) modeling-guided taxonomy is proposed. Next, a survey of successful approaches that have led to commercial systems is given in accordance with the proposed taxonomy. Finally, a general review of these approaches according to factors that are intrinsic factors (spatial and temporal resolutions, depth accuracy, sensor cost, etc.) and extrinsic (motion speed, illumination changes, face details, intrusion and need for user cooperation, etc.) are provided.

Chapter 2 discusses the state of the art in 3D surface features for the recognition of the human face. Particular emphasis is laid on the most prominent and recent contributions. The features extracted from 3D facial surfaces serve as means for dimensionality reduction of surface data and for facilitating the task of face recognition. The complexity of extraction, descriptiveness, and robustness of features directly affect the overall accuracy, performance, and robustness of the 3D recognition system.

Chapter 3 presents a novel geometric framework for analyzing 3D faces, with specific goals of comparing, matching, and averaging their shapes. In this framework, facial surfaces are represented by radial curves emanating from the nose tips. These curves, in turn, are compared using elastic shape analysis to develop a Riemannian framework for full facial surfaces. This representation, along with the elastic Riemannian metric, seems natural for measuring facial deformations and is robust to data issues such as large facial expressions. One difficulty in extracting *facial curves* from the surface of 3D face scans is related to the presence of noise. A possible way to smooth the effect of the noise without losing the effectiveness of representations is to consider aggregates of facial curves, as opposed to individual curves, called *iso-geodesic stripes*.

Chapter 4 presents an automatic and efficient method to fit a statistical deformation model of the human face to 3D scan data. In a global-to-local fitting scheme, the shape parameters of this model are optimized such that the produced instance of the model accurately fits the 3D scan data of the input face. To increase the expressiveness of the model and to produce a tighter fit of the model, the method fits a set of predefined face components and blends these components afterwards. In the case that a face cannot be modeled, the automatically acquired model coefficients are unreliable, which hinders the automatic recognition. Therefore, we present a bootstrapping algorithm to automatically enhance a 3D morphable face model with new face data. The accurately generated face instances are manifold meshes without noise and holes, and can be effectively used for 3D face recognition. The results show that model coefficient based face matching outperforms contour curve and landmark based face matching, and is more time efficient than contour curve matching.

Although there have been many research efforts in the area of 3D face analysis in the last few years, the development of potential applications and exploitation of face recognition tools is still in its infancy. Chapter 5 summarizes recent trends in 3D face analysis with particular emphasis on the application techniques introduced and discussed in the previous chapters. The chapter discusses how 3D face analysis has been used to improve face recognition in the presence of facial expressions and missing parts, and how 3D techniques are now being extended to process dynamic sequences of 3D face scans for the purpose of facial expression recognition.

We hope that this will serve as a good reference book for researchers and students interested in this field.

Mohamed Daoudi, *TELECOM Lille 1/LIFL, France*
Anuj Srivastava, *Florida State University, USA*
Remco Veltkamp, *Utrecht University, The Netherlands*

List of Contributors

Faisal Radhi M. Al-Osaimi, Department of Computer Engineering, College of Computer & Information Systems, Umm Al-Qura University, Saudi Arabia

Mohsen Ardabilian, Ecole Centrale de Lyon, Département Mathématiques – Informatique, France

Boulbaba Ben Amor, TELECOM Lille1, France

Mohammed Bennamoun, School of Computer Science & Software Engineering, The University of Western Australia, Australia

Stefano Berretti, Dipartimento di Sistemi e Informatica Universit degli Studi di Firenze, Italy

Alberto del Bimbo, Dipartimento di Sistemi e Informatica, Universit degli Studi di Firenze, Italy

Liming Chen, Ecole Centrale de Lyon, Département Mathématiques Informatique, France

Mohamed Daoudi, TELECOM Lille1, France

Hassen Drira, TELECOM Lille1, France

Frank B. ter Haar, TNO, Intelligent Imaging, The Netherlands

Pietro Pala, Dipartimento di Sistemi e Informatica, Universit di Firenze, Italy

Anuj Srivastava, Department of Statistics, Florida State University, USA

Remco Veltkamp, Department of Information and Computing Sciences, Universiteit Utrecht, The Netherlands

Contents

Preface ix

List of Contributors xiii

1 3D Face Modeling **1**
Boulbaba Ben Amor, Mohsen Ardabilian and Liming Chen

1.1 Challenges and Taxonomy of Techniques 2
1.2 Background 3
1.2.1 Depth from Triangulation 4
1.2.2 Shape from Shading 5
1.2.3 Depth from Time of Flight (ToF) 6
1.3 Static 3D Face Modeling 7
1.3.1 Laser-stripe Scanning 7
1.3.2 Time-coded Structured Light 8
1.3.3 Multiview Static Reconstruction 11
1.4 Dynamic 3D Face Reconstruction 14
1.4.1 Multiview Dynamic Reconstruction 14
1.4.2 Photometric Stereo 17
1.4.3 Structured Light 18
1.4.4 Spacetime Faces 24
1.4.5 Template-based Post-processing 27
1.5 Summary and Conclusions 31
Exercises 33
References 35

2 3D Face Surface Analysis and Recognition Based on Facial Surface Features **39**
Faisal Radhi M. Al-Osaimi and Mohammed Bennamoun

2.1 Geometry of 3D Facial Surface 39
2.1.1 Primary 3D Surface Representations 40
2.1.2 Rigid 3D Transformations 47
2.1.3 Decimation of 3D Surfaces 49
2.1.4 Geometric and Topological Aspects of the Human Face 51

2.2 Curvatures Extraction from 3D Face Surface 53
2.2.1 Theoretical Concepts on 3D Curvatures 53
2.2.2 Practical Curvature Extraction Methods 56
2.3 3D Face Segmentation 57
2.3.1 Curvature-based 3D Face Segmentation 57
2.3.2 Bilateral Profile-based 3D Face Segmentation 58
2.4 3D Face Surface Feature Extraction and Matching 59
2.4.1 Holistic 3D Facial Features 60
2.4.2 Regional 3D Facial Features 67
2.4.3 Point 3D Facial Features 68
2.5 Deformation Modeling of 3D Face Surface 71
Exercises 73
References 74

3 3D Face Surface Analysis and Recognition Based on Facial Curves 77
Hassen Drira, Stefano Berretti, Boulbaba Ben Amor, Mohamed Daoudi, Anuj Srivastava, Alberto del Bimbo and Pietro Pala

3.1 Introduction 77
3.2 Facial Surface Modeling 78
3.3 Parametric Representation of Curves 80
3.4 Facial Shape Representation Using Radial Curves 81
3.5 Shape Space of Open Curves 81
3.5.1 Shape Representation 82
3.5.2 Geometry of Preshape Space 84
3.5.3 Reparametrization Estimation by Using Dynamic Programming 86
3.5.4 Extension to Facial Surfaces Shape Analysis 88
3.6 The Dense Scalar Field (DSF) 90
3.7 Statistical Shape Analysis 94
3.7.1 Statistics on Manifolds: Karcher Mean 94
3.7.2 Learning Statistical Models in Shape Space 96
3.8 Applications of Statistical Shape Analysis 98
3.8.1 3D Face Restoration 98
3.8.2 Hierarchical Organization of Facial Shapes 101
3.9 The Iso-geodesic Stripes 103
3.9.1 Extraction of Facial Stripes 107
3.9.2 Computing Relationships between Facial Stripes 109
3.9.3 Face Representation and Matching Using Iso-geodesic Stripes 113
Exercises 114
Glossary 116
References 117

4 3D Morphable Models for Face Surface Analysis and Recognition 119
Frank B. ter Haar and Remco Veltkamp

4.1 Introduction 120
4.2 Data Sets 121

4.3 Face Model Fitting 122
4.3.1 Distance Measure 122
4.3.2 Iterative Face Fitting 123
4.3.3 Coarse Fitting 124
4.3.4 Fine Fitting 124
4.3.5 Multiple Components 125
4.3.6 Results 126
4.4 Dynamic Model Expansion 129
4.4.1 Bootstrapping Algorithm 131
4.4.2 Results 136
4.5 Face Matching 141
4.5.1 Comparison 141
4.5.2 Results 142
4.6 Concluding Remarks 144
Exercises 145
References 146

5 Applications 149
Stefano Berretti, Boulbaba Ben Amor, Hassen Drira, Mohamed Daoudi, Anuj Srivastava, Alberto del Bimbo and Pietro Pala

5.1 Introduction 149
5.2 3D Face Databases 150
5.3 3D Face Recognition 157
5.3.1 Challenges of 3D Face Recognition 158
5.3.2 3D Face Recognition: State of the Art 159
5.3.3 Partial Face Matching 162
5.3.4 Comparison of State-of-the-Art Methods 168
5.4 Facial Expression Analysis 170
5.4.1 3D Facial Expression Recognition: State of the Art 171
5.4.2 Semi-automatic 3D Facial Expression Recognition 173
5.4.3 Fully Automatic 3D Facial Expression Recognition 180
5.5 4D Facial Expression Recognition 184
5.5.1 The BU-4DFE Database 186
5.5.2 3D Shape Motion Analysis 187
5.5.3 Discussion and Comparative Evaluation 192
Exercises 192
Glossary 193
References 198

Index 203

1

3D Face Modeling

Boulbaba Ben Amor,[1] Mohsen Ardabilian,[2] and Liming Chen[2]
[1] *Institut Mines-Télécom/Télécom Lille 1, France*
[2] *Ecole Centrale de Lyon, France*

Acquiring, modeling, and synthesizing realistic 3D human faces and their dynamics have emerged as an active research topic in the border area between the computer vision and computer graphics fields of research. This has resulted in a plethora of different acquisition systems and processing pipelines that share many fundamental concepts as well as specific implementation details. The research community has investigated the possibility of targeting either end-to-end consumer-level or professional-level applications, such as facial geometry acquisition for 3D-based biometrics and its dynamics capturing for expression cloning or performance capture and, more recently, for 4D expression analysis and recognition. Despite the rich literature, reproducing realistic human faces remains a distant goal because the challenges that face 3D face modeling are still open. These challenges include the motion speed of the face when conveying expressions, the variabilities in lighting conditions, and pose. In addition, human beings are very sensitive to facial appearance and quickly sense any anomalies in 3D geometry or dynamics of faces. The techniques developed in this field attempt to recover facial 3D shapes from camera(s) and reproduce their actions. Consequently, they seek to answer the following questions:

☞ How can one recover the facial shapes under pose and illumination variations?
☞ How can one synthesize realistic dynamics from the obtained 3D shape sequences?

This chapter provides a brief overview of the most successful existing methods in the literature by first introducing basics and background material essential to understand them. To this end, instead of the *classical* passive/active taxonomy of 3D reconstruction techniques, we propose here to categorize approaches according to whether they are able to acquire faces in action or they can only capture them in a static state. Thus, this chapter is preliminary to

3D Face Modeling, Analysis and Recognition, First Edition.
Edited by Mohamed Daoudi, Anuj Srivastava and Remco Veltkamp.

the following chapters that use static or dynamic facial data for face analysis, recognition, and expression recognition.

1.1 Challenges and Taxonomy of Techniques

Capturing and processing human geometry is at the core of several applications. To work on 3D faces, one must first be able to recover their shapes. In the literature, several acquisition techniques exist that are either dedicated to specific objects or are general. Usually accompanied by geometric modeling tools and post-processing of 3D entities (3D point clouds, 3D mesh, volume, etc.), these techniques provide complete solutions for 3D full object reconstruction. The acquisition quality is mainly linked to the accuracy of recovering the z-coordinate (called depth information). It is characterized by loyalty reconstruction, in other words, by data quality, the density of 3D face models, details preservation (regions showing changes in shapes), etc. Other important criteria are the acquisition time, the ease of use, and the sensor's cost. In what follows, we report the main *extrinsic* and *intrinsic* factors which could influence the modeling process.

- ☞ *Extrinsic factors*. They are related to the environmental conditions of the acquisition and the face itself. In fact, human faces are globally similar in terms of the position of main features (eyes, mouth, nose, etc.), but can vary considerably in details across (i) their variabilities due to facial deformations (caused by expressions and mouth opening), subject aging (wrinkles), etc, and (ii) their specific details as skin color, scar tissue, face asymmetry, etc. The environmental factors refer to lighting conditions (controlled or ambient) and changes in head pose.
- ☞ *Intrinsic factors*. They include sensor cost, its intrusiveness, manner of sensor use (cooperative or not), spatial and/or temporal resolutions, measurement accuracy and the acquisition time, which allows us to capture moving faces or simply faces in static state.

These challenges arise when acquiring static faces as well as when dealing with faces in action. Different applications have different requirements. For instance, in the computer graphics community, the results of performance capture should exhibit a great deal of spatial fidelity and temporal accuracy to be an authentic reproduction of a real actors' performance. Facial recognition systems, on the other hand, require the accurate capture of person-specific details. The movie industry, for instance, may afford a 3D modeling pipeline system with special purpose hardware and highly specialized sensors that require manual calibration. When deploying a 3D acquisition system for facial recognition at airports and in train stations, however, cost, intrusiveness, and the need of user cooperation, among others, are important factors to consider. In ambient intelligence applications where a user-specific interface is required, facial expression recognition from 3D sequences emerges as a research trend instead of 2D-based techniques, which are sensitive to changes and pose variations. Here, also, sensor cost and its capability to capture facial dynamics are important issues. Figure 1.1 shows a new 3D face modeling-guided taxonomy of existing reconstruction approaches. This taxonomy proposes two categories: The first category targets 3D static face modeling, while the approaches belonging to the second category try to capture facial shapes in action (i.e., in 3D+t domain). In the level below, one finds different approaches based on concepts presented