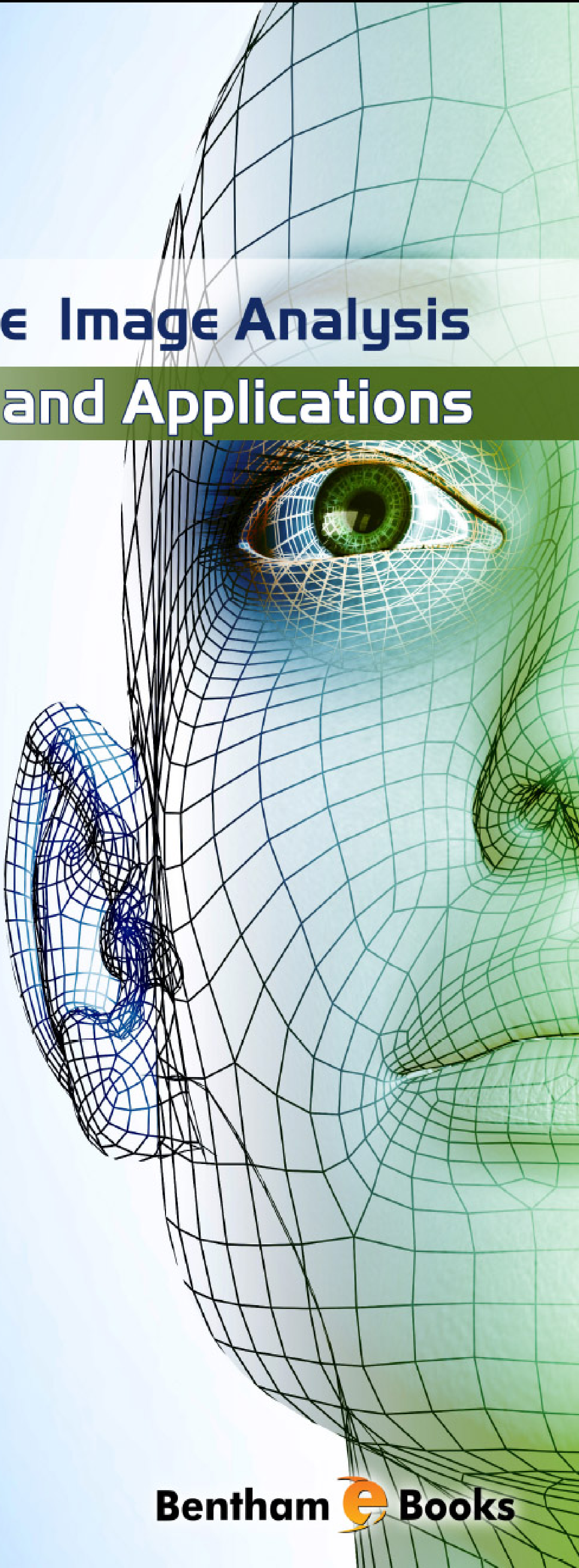
eISBN: 978-1-68108-110-6
ISBN: 978-1-68108-111-3

Advances in Face Image Analysis

Theory and Applications



Editor
Fadi Dornaika

Bentham e Books

Advances in Face Image Analysis: Theory and Applications

Edited By

Fadi Dornaika

University of the Basque Country (UPV/EHU)

IKERBASQUE Foundation for Science

Manuel Lardizabal,

1, 20018 San Sebastián,

Spain

advertisements or ideas contained in the Work.

Limitation of Liability:

In no event will Bentham Science Publishers, its staff, editors and/or authors, be liable for any damages, including, without limitation, special, incidental and/or consequential damages and/or damages for lost data and/or profits arising out of (whether directly or indirectly) the use or inability to use the Work. The entire liability of Bentham Science Publishers shall be limited to the amount actually paid by you for the Work.

General:

1. Any dispute or claim arising out of or in connection with this License Agreement or the Work (including non-contractual disputes or claims) will be governed by and construed in accordance with the laws of the U.A.E. as applied in the Emirate of Dubai. Each party agrees that the courts of the Emirate of Dubai shall have exclusive jurisdiction to settle any dispute or claim arising out of or in connection with this License Agreement or the Work (including non-contractual disputes or claims).
2. Your rights under this License Agreement will automatically terminate without notice and without the need for a court order if at any point you breach any terms of this License Agreement. In no event will any delay or failure by Bentham Science Publishers in enforcing your compliance with this License Agreement constitute a waiver of any of its rights.
3. You acknowledge that you have read this License Agreement, and agree to be bound by its terms and conditions. To the extent that any other terms and conditions presented on any website of Bentham Science Publishers conflict with, or are inconsistent with, the terms and conditions set out in this License Agreement, you acknowledge that the terms and conditions set out in this License Agreement shall prevail.

Bentham Science Publishers Ltd.
Executive Suite Y - 2
PO Box 7917, Saif Zone
Sharjah, U.A.E.
Email: subscriptions@benthamscience.org

BENTHAM SCIENCE

CONTENTS

FOREWORD ... i

PREFACE ... ii

LIST OF CONTRIBUTORS ... iii

CHAPTER 1 FACIAL EXPRESSION CLASSIFICATION BASED ON CONVOLUTIONAL NEURAL NETWORKS ... 3

- **INTRODUCTION** ... 3
  - Convolutional Neural Networks ... 4
  - Facial Expression Analysis ... 5
- **GRADIENT-BASED LEARNING FOR CNNS** ... 6
- **FEATURE GENERALIZATION** ... 10
- **EXPERIMENTS** ... 11
  - Datasets ... 11
    - *CK-Regianini Dataset* ... 11
    - *CK-Zheng Dataset* ... 13
    - *CMU-Pittsburgh dataset* ... 13
  - Experiments on CNN-based Facial Expression Classification ... 13
    - *Design* ... 13
    - *Results and Analysis* ... 14
  - Experiments on Feature Generalization ... 16
    - *Design* ... 16
    - *Results and Analysis* ... 18
- **DISCUSSION** ... 19
- **CONFLICT OF INTEREST** ... 20
- **ACKNOWLEDGEMENTS** ... 20
- **REFERENCES** ... 20

CHAPTER 2 SPARSITY PRESERVING PROJECTION BASED CONSTRAINED GRAPH EMBEDDING AND ITS APPLICATION TO FACE RECOGNITION ... 23

- **INTRODUCTION** ... 24
- **RELATED WORK** ... 25
  - Locality Preserving Projection ... 25
  - Neighborhood Preserving Embedding ... 27
  - Sparsity Preserving Projection ... 28
  - Constrained Graph Embedding ... 28
- **SPP BASED CONSTRAINED GRAPH EMBEDDING** ... 30
- **SPP-CGE** ... 30
  - Out-of-Sample Extension ... 31
- **EXPERIMENTAL RESULTS** ... 33
- **CONCLUSION** ... 36
- **NOTES** ... 37
- **CONFLICT OF INTEREST** ... 37
- **ACKNOWLEDGEMENTS** ... 37
- **REFERENCES** ... 37

CHAPTER 3 FACE RECOGNITION USING EXPONENTIAL LOCAL DISCRIMINANT EMBEDDING

.... 39
INTRODUCTION 40
Contribution and Related Work 42
REVIEW OF LOCAL DISCRIMINANT EMBEDDING (LDE) 44
Intrinsic Graph and Penalty Graph 44
Optimal Mapping 46
The Small Sample Size Problem 47
EXPONENTIAL LDE 47
Matrix Exponential 48
Exponential LDE 48
THEORETICAL ANALYSIS OF ELDE 50
Solving the SSS Problem 50
Distance Diffusion Mapping 50
PERFORMANCE EVALUATION 53
Face Databases 54
Recognition Accuracy 54
Comparison between Regularized LDE and ELDE 59
CONCLUSION 60
NOTES 61
CONFLICT OF INTEREST 61
ACKNOWLEDGMENTS 61
REFERENCES 62

CHAPTER 4 ADAPTIVE LOCALITY PRESERVING PROJECTIONS FOR FACE RECOGNITION 66
INTRODUCTION 67
LOCALITY PRESERVING PROJECTIONS 68
ENHANCED AND PARAMETERLESS LPP 71
PERFORMANCE EVALUATION 73
Face Databases 73
Experimental Results 74
Performance Comparison for OLPP and SLPP 81
CONCLUSION 82
NOTES 83
CONFLICT OF INTEREST 83
ACKNOWLEDGMENTS 83
REFERENCES 83

CHAPTER 5 FACE RECOGNITION USING 3D FACE RECTIFICATION 86
INTRODUCTION 86
PROPOSED METHOD 88
FACE DATABASE 89
PREPROCESSING 90
FACIAL FEATURE DETECTION 91
POSE ESTIMATION 93
IRAD Contours 94
Ellipse Fitting And Roll Correction 95
Yaw Correction 96
Pitch Correction 97
Accuracy Of The Pose Estimation Method 98
ROTATION AND POST PROCESSING 98
EXPERIMENTS 99

CONCLUSION 103
NOTES 104
CONFLICT OF INTEREST 104
ACKNOWLEDGMENTS 104
REFERENCES 104

CHAPTER 6 3D FACE RECOGNITION 109
INTRODUCTION 109
3D FACE ACQUISITION 111
3D FACE REPRESENTATION 116
PREPROCESSING 119
3D FACE ALIGNMENT 122
FACE RECOGNITION 124
CONCLUDING REMARKS 127
CONFLICT OF INTEREST 127
ACKNOWLEDGEMENTS 128
REFERENCES 128

CHAPTER 7 MODEL-LESS 3D FACE POSE ESTIMATION 132
INTRODUCTION 132
STATE OF THE ART 133
THE MACHINE LEARNING METHODOLOGY 134
Locality Preserving Projections 136
LPP Algorithm 136
Supervised Locality Preserving Projections 137
LABEL-SENSITIVE LOCALITY PRESERVING PROJECTION 138
Presetting: 138
Algorithm: 138
PROPOSED APPROACH: SPARSE GRAPH BASED LSLPP 139
EXPERIMENTAL RESULTS 141
CONCLUSION 150
NOTES 150
CONFLICT OF INTEREST 151
ACKNOWLEDGEMENTS 151
REFERENCES 151

CHAPTER 8 EFFICIENT DEFORMABLE 3D FACE MODEL FITTING TO MONOCULAR IMAGES ... 154
INTRODUCTION 155
LIGHTWEIGHT FACIAL FEATURE DETECTION 159
DEFORMABLE MODEL BACKPROJECTION 165
EXPERIMENTAL RESULTS AND DISCUSSION 167
CONCLUSION 173
APPENDIX A: MODIFICATIONS TO CANDIDE-3 175
NOTES 175
CONFLICT OF INTEREST 176
ACKNOWLEDGMENTS 176
REFERENCES 176

CHAPTER 9 FACE DETECTION USING THE THEORY OF EVIDENCE 181
INTRODUCTION 182

STATE OF THE ART 183
THEORY OF BELIEF FUNCTIONS 185
Mass Sets 185
Modeling of Mass Functions 187
Combination of Beliefs 189
Management of Conflict 190
Combination Rules for Dependent Sources 190
Decision with Transferable Belief Model 191
APPLICATION TO EVIDENTIAL FACE MODEL 192
Information Sources 193
Mass Sets for VJ Face Detector 194
Color Mass Functions 195
Color Fusion by Cautious Rule 197
Illustrative Example 198
Global Fusion of Color and VJ Mass Sets by Conjunctive Rule 199
Computation of Pignistic Probabilities 201
PROBABILISTIC FACE TRACKING 202
Estimation of Ellipse Center 205
Estimation of Size and Pose 207
Visual Servoing 207
PERFORMANCE ANALYSIS 208
Qualitative Evaluation 208
Quantitative Evaluation 210
DISCUSSION 212
CONFLICT OF INTEREST 213
ACKNOWLEDGEMENTS 213
REFERENCES 213

CHAPTER 10 FUZZY DISCRIMINANT ANALYSIS: CONSIDERING THE FUZZINESS IN FACIAL AGE FEATURE EXTRACTION 217
INTRODUCTION 217
CHARACTERIZATION OF THE FUZZINESS IN FACIAL AGING 219
THE FUZZY DISCRIMINATIVE MANIFOLD ANALYSIS 221
Traditional Discriminative Manifold Methods 221
Fuzzy Discriminative Analysis for Age Estimation 222
EXPERIMENTAL RESULTS 225
A Toy Example 225
Facial Age Estimation Results 227
Databases 227
Experimental Setting 227
Experiments on the FGNET Database 228
Experiments on the MORPH Database 229
CONCLUSION 231
CONFLICT OF INTEREST 231
ACKNOWLEDGEMENTS 232
REFERENCES 232

CHAPTER 11 FACIAL IMAGE-BASED AGE ESTIMATION 234
INTRODUCTION 234
AGE ESTIMATION ALGORITHMS 235
Input Images Preprocessing 236

Feature Extraction 236
Distance Metric Adjustments 237
Dimensionality Reduction 238
Age Determination 238
AGE ESTIMATION FRAMEWORK AND EXPERIMENTS PERFORMED 239
EXPERIMENTAL SETTING 240
Datasets 240
Performance Scores 242
RESULTS 242
CONCLUSION 246
CONFLICT OF INTEREST 247
ACKNOWLEDGEMENTS 247
REFERENCES 247

SUBJECT INDEX 251

FOREWORD

Computer vision is one of the most active research fields in information technology, computer science and electrical engineering due to its numerous applications and major research challenges. Face image analysis constitutes an important field in computer vision and can be a key challenge in developing human-centered technologies. Face image analysis problems have been investigated in computer vision and Human Machine Interaction applications (*e.g.*, identity verification, eye typing, emotion recognition, m-commerce). Making computers understand the contents of images taken by cameras is very challenging, and therefore the computer vision technology faces a lot of challenges. Differed from the biometric problems, *e.g.*, finger-print or iris based recognition; face recognition inherently relies on the un-controlled environment and inevitably suffers from degrading factors such as illumination, expression, pose and age variations. Image-based age estimation is relatively a new research topic. Estimating human age automatically *via* facial image analysis has lots of potential real-world applications, such as human computer interaction and multimedia communication.

This book presents the reader with cutting edge research in the domain of face image analysis. Besides, the book includes recent research works from different world research groups, providing a rich diversity of approaches to the face image analysis. The material covered in the eleven chapters of the book presents new advances on computer vision and pattern recognition approaches, as well as new knowledge and perspectives.

The chapters, written by experts in their respective field, will make the reader acquainted with a number of topics and some trendy techniques used to tackle many problems related to face images. It is impressive to note that the editor and authors have tried to capture a wide and dynamic topic. I believe readers will not only learn from this book, but it will be of high reference value as well.

Denis Hamad
Université du Littoral Côte d'Opale,
Calais, France

PREFACE

Over the past two decades, many face image analysis problems have been investigated in computer vision and machine learning. The main idea and the driver of further research in this area are human-machine interaction and security applications. Face images and videos can represent an intuitive and non-intrusive channel for recognizing people, inferring their level of interest, and estimating their gaze in 3D. Although progress over the past decade has been impressive, there are significant obstacles to be overcome. It is not possible yet to design a face analysis system with a potential close to human performance. New computer vision and pattern recognition approaches need to be investigated. Face recognition as an essential problem in pattern recognition and social media computing, attracts many researchers for decades. For instance, face recognition became one of three identification methods used in e-passports and a biometric of choice for many other security applications.

The e-Book “Advances in Face Image Analysis: Theory and Applications” is oriented to a wide audience including: i) researchers and professionals working in the fields of face image analysis; ii) the entire pattern recognition community interested in processing and extracting features from raw face images; and iii) technical experts as well as postgraduate students working on face images and their related concepts. One of the key benefits of this E-Book is that the readers will have access to novel research topics. The book contains eleven chapters that address several topics including automatic face detection, 3D face model fitting, robust face recognition, facial expression recognition, face image data embedding, model-less 3D face pose estimation and image-based age estimation.

We would like to express our gratitude to all the contributing authors that have made this book a reality. We would also like to thank Prof. Denis Hamad for writing the foreword and Bentham Science Publishers for their support and efforts. A special thank goes to Dr. Ammar Assoum for providing the latex style file.

Fadi Dornaika
University of the Basque Country
Manuel Lardizabal, 1
20018 San Sebastián, Spain

LIST OF CONTRIBUTORS

Ammar Assoum	LaMA laboratory, Lebanese University, Tripoli, Lebanon
Alireza Behrad	Department of Electrical and Electronic Engineering, Shahed university, Tehran-Qom Exp. Way, 3319118651, Tehran, Iran
Alireza Bosaghzadeh	University of the Basque Country, Manuel Lardizabal, 1, 20018, San Sebastian, Spain
Fadi Dornaika	University of the Basque Country, Manuel Lardizabal, 1, 20018, San Sebastian, Spain Department of Computer Science and Artificial Intelligence, University of the Basque Country, UPV/EHU, San Sebastian, Spain IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
Jon Goenetxea	Vicomtech-IK4, Paseo Mikeletegi, 57, Parque Tecnológicoo, 20009, Donostia, Spain
Jouhayna Harmouch	LaMA Laboratory, Lebanese University, Tripoli, Lebanon
Zhong Jin	School of Computer Science & Engineering, Nanjing University of Sciences and Technology, Nanjing, China
Fawzi Khattar	LaMA Laboratory, Lebanese University, Tripoli, Lebanon
Franck Luthon	IUT de Bayonne Pays Basque, Université de Pau Pays d'Adour,, 2 allée du parc Montaury, 64600, Anglet, France
Waldir Pimenta	Departamento de Informática, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal
Luis P. Santos	Departamento de Informática, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal
Shenglan Ben	School of Electronic Science & Engineering, Nanjing University, Nanjing, 210094, China
Sun Wenyun	School of Computer Science & Engineering, Nanjing University, Nanjing, 210094, China
Luis Unzueta	Vicomtech-IK4, Paseo Mikeletegi, 57, Parque Tecnológicoo, 20009, Donostia, Spain
Libo Weng	Department of Computer Science and Artificial Intelligence, University of the Basque Country, UPV/EHU, San Sebastian, Spain School of Computer Science & Engineering, Nanjing University of Sciences and Technology, Nanjing, China

CHAPTER 1

Facial Expression Classification Based on Convolutional Neural Networks

Wenyun Sun, Zhong Jin*

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China

Abstract: Research trends in Convolutional Neural Networks and facial expression analysis are introduced at first. A training algorithm called stochastic gradient descent with l_2 regularization is employed for the facial expression classification problem, in which facial expression images are classified into six basic emotional categories of anger, disgust, fear, happiness, sadness and surprise without any complex pre-processes involved. Moreover, three types of feature generalization for solving problems with different classifiers, different datasets and different categories are discussed. By these techniques, pre-trained Convolutional Neural Networks are used as feature extractors which work quite well with Support Vector Machine classifiers. The results of experiments show that Convolutional Neural Networks not only have capability of classifying facial expression images with translational distortions, but also have capability to fulfill some feature generalization tasks.

Keywords: Alex-Net architecture, Backpropagation algorithm, CK-Regianini dataset, CK-Zheng dataset, Classification accuracy, CMU-Pittsburgh dataset, Combined features, Convolutional Neural Networks, Deep learning, Facial expression classification, Feature extraction, Feature generalization, Feature representation, Hidden layers, Pre-trained networks, Stochastic Gradient Descent, Supervised feature learning, Support Vector Machine, Trainable parameters, Translational invariance property.

INTRODUCTION

A feature extractor and a classifier are two essential modules in a conventional

* **Address to Corresponding Author Zhong Jin:** School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China; Tel: +86 25 84303280 Ext. 3071; Fax: +86 25 84317235; E-mail: zhongjin@njust.edu.cn

Fadi Dornaika (Ed)

image pattern recognition system. A good feature extractor of image could produce a feature representation which has more discriminant information and less correlations than the original pixel data. There are quite a few popular techniques recently, *e.g.*, the Scale-Invariant Feature Transform (SIFT) [1] and the Histogram of Oriented Gradients (HOG) [2]. On the other hand, a highly efficient classifier could perform its job well without any help of complex feature extractors. Nowadays, some highly efficient classifiers and good feature extractors based on deep learning have come out.

Convolutional Neural Networks

Remarkable achievements have been obtained by studies of classifiers for high dimensional image data in the last two decades. More and more attentions have been gotten by Convolutional Neural Networks (CNNs) which have become the representatives among other deep learning methods. Although CNNs were suggested in 1989 [3], efficient training algorithms were absent until the stochastic diagonal Levenberg-Marquardt algorithm for CNNs was proposed by LeCun *et al*. in 1998 [4]. A so-called LeNet-5 was designed by LeCun *et al*. It could classify handwritten digits and letters into categories without complex preprocesses.

There is some theoretical research which brings the state-of-the-art techniques to the classical CNNs recently, *e.g.* rectified linear unit [5], local contrast normalization [6], local response normalization [7] and dropout [8]. On the other hand, engineering studies have never been stopped. Handwritten character recognition [9, 10], natural image processing [7, 11], *etc*. are well-known engineering application of CNNs.

The most interesting work had been done by Krizhevsky *et al*. who won the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012 [7]. They achieved a top-5 error rate of 16.4% on the classification benchmark, which beat the second place result of 26.1% with handcrafted features.

In ILSVRC 2013, an approach from Zeiler *et al*. [11] improved the performance by visualizing hidden layers of CNNs. They found that Krizhevsky's network has the ability of extracting features of different scale and complexity. This phenomenon shows the feature representation capability of CNNs evidently. A reliable feature extractor could be easily got by cutting the Soft-Max layer off at the end of CNNs and keeping the rest layer's trainable parameters fixed. The features from one of each hidden layers, especially from the last one, could be used as the inputs of any other classifiers. In other words, when a classifier is trained, a feature extractor will be got at the same time. The extractor can be widely used for various purposes. Based on this view of point, Zeiler *et al*. proposed a theory of feature generalization. Abundant feature information is included in nature images which also have a large scale of categories. Thus, a pre-trained feature extracting network for natural images could be applied to the processes of specific data conveniently. Finally, it is notable that these methods we mentioned above are usually implemented and accelerated by Graphics Processing Unit (GPU) based high performance computing techniques.

Facial Expression Analysis

In another domain, the research of classifying facial expressions was started by psychologists. In 1978, facial action coding system (FACS) [12] was proposed by Ekman *et al*. The well-known facial expression image dataset, called Cohn-Kanade (CK) [13, 14], built by the Robotics Institute of Carnegie Mellon University and Department of Psychology of University of Pittsburgh, contains a set of facial expression image sequences and their corresponding action unit (AU) codes.

In 1984, Ekman *et al*. continued their studies, and classified facial expressions into six categories by different emotions, *i.e.*, anger, disgust, fear, happiness, sadness and surprise [15]. In the problem of facial expression analysis, especially in the classification case, feature extractors and classifiers should keep invariant to individuals and perspective projection distortions.

In the recent years, quite a few studies have been devoted to expression classification of static images or image sequences. Some are about specific handcrafted feature extraction algorithms [16, 17], and some are about classifiers which use plan 2-dimensional pixels data as their inputs [18, 19].

In the following sections, a gradient-based learning algorithm and a feature

CHAPTER 2

Sparsity Preserving Projection Based Constrained Graph Embedding and Its Application to Face Recognition

Libo Weng[1,2], Zhong Jin[1,*], Fadi Dornaika[2,3]

[1] *School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China*

[2] *Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU, San Sebastian, Spain*

[3] *IKERBASQUE, Basque Foundation for Science, Bilbao, Spain*

Abstract: In this chapter, a novel semi-supervised dimensionality reduction algorithm is proposed, namely Sparsity Preserving Projection based Constrained Graph Embedding (SPP-CGE). Sparsity Preserving Projection (SPP) is an unsupervised dimensionality reduction method. It aims to preserve the sparse reconstructive relationship of the data obtained by solving a L_1 objective function. Label information is used as additional constraints for graph embedding in the SPP-CGE algorithm. In SPP-CGE, both the intrinsic structure and the label information of the data are used. In addition, to deal with new incoming samples, out-of-sample extension of SPP-CGE is also proposed. Promising experimental results on several popular face databases illustrate the effectiveness of the proposed method.

Keywords: Affinity matrix, Constrained graph embedding, Dimensionality reduction, Eigenvalue problem, Face recognition, Graph embedding, ISOMAP, Laplacian eigenmaps, Laplacian matrix, Linear discriminant analysis, Locality preserving projection, Locally linear embedding, Multidimensional scaling, Neighborhood preserving embedding, Principal component analysis, Projection matrix, Recognition rate, Semi-supervised learning, Sparse representation, Sparsity preserving projection.

[*] **Address to Corresponding Author Zhong Jin:** School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China; Tel: +86 25 84303280 Ext. 3071; Fax: +86 25 84317235; E-mail: zhongjin@njust.edu.cn

Fadi Dornaika (Ed)

INTRODUCTION

In many real world applications, such as face recognition and text categorization, the data is usually provided in high dimension space. Moreover, the labels of the original data are usually inadequate and it will spend expensive human labor to acquire the labels. To deal with this problem, semi-supervised dimensionality reduction methods can be used to project the data in the high-dimensional space into a space of fewer dimensions.

In the recent years, researchers have proposed a lot of methods for dimension reduction. Principal Component Analysis [1] (PCA) and Multidimensional Scaling [2] (MDS) are two classic linear unsupervised dimensionality reduction methods. Linear Discriminant Analysis [1] (LDA) is a supervised method. In 2000, Locally Linear Embedding [3] (LLE) and ISOMAP [4] were separately proposed in *science* which laid a foundation of manifold learning. Soon afterwards, M. Belkin *et al.* proposed Laplacian Eigenmaps [5] (LE). He *et al.* proposed both Locality Preserving Projection [6] (LPP), essentially a linearized version of LE, and Neighborhood Preserving Embedding [7] (NPE), a linearized version of LLE. LPP and NPE can be interpreted in a general graph embedding framework with different choices of graph structure.

Sparsity Preserving Projection [8 - 11] (SPP) is an unsupervised learning method. It can be considered as an extension to NPE since the latter has similar objective function. However, SPP utilises sparse representation to obtain the affinity matrix. We extend SPP to the semi-supervised case by integrating the idea of Constrained Graph Embedding [12] (CGE). CGE tries to project the data point with the same labels into one single point in the projection space by a constraint matrix. We propose a method called sparsity preserving projection based constrained graph embedding that can combine the benefits of both SPP and CGE. In SPP-CGE algorithm, the construction of affinity matrix is parameter-free, local structure is preserved and data points with the same label are projected into one point in the projection space.

Since SPP-CGE is essentially a nonlinear method, the projection matrix of SPP-CGE could not be obtained in a direct manner. A traditional way to deal with a new incoming sample is to re-perform the whole algorithm one package again which will be quite time-consuming. A simple way to obtain an approximate mapping matrix is presented to replace the unknown mapping function for data projection.

In this chapter, we propose a new semi-supervised method for dimensionality reduction named SPP-CGE. The chapter is organized as follows. Firstly, the related methods including LPP, NPE, SPP and CGE are introduced. Then, the proposed method and the out-of-sample extension are presented. Finally, some experimental results for face recognition on three databases: Yale [13, 14], ORL and PIE are given.

RELATED WORK

Some mathematical notations are listed and will be used in the next several sections. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \in R^{m\times n}$ be the data matrix, where n is the number of training samples and m is the dimension of each sample. Let $\mathbf{y} = [y_1, y_2, \ldots, y_n]^T$ be a one-dimensional map of $\mathbf{X}$. Under a linear projection $\mathbf{y}^T = \mathbf{p}^T \mathbf{X}$, each data point $\mathbf{x}_i$ in the input space R^m is mapped into $y_i = \mathbf{p}^T \mathbf{x}_i$ in the real line. Here, $\mathbf{p} \in R^m$ is a projection axis. Let $\mathbf{Y} \in R^{d\times n}$ be the data projections into a d dimensional space.

Locality Preserving Projection

LPP aims to preserve the local structure of the data by keeping two sample points close in the projection space when they are similar in the original space.

The reasonable criterion of LPP is to optimize the following objective function under some constraints:

$$\min_{\mathbf{p}} \sum_{i,j} (y_i - y_j)^2 W_{ij}, \tag{1}$$

where $\mathbf{W}$ is the affinity matrix associated with the data.

The way to define $\mathbf{W}$ can be alterable. One simple definition is as follows:

CHAPTER 3

Face Recognition Using Exponential Local Discriminant Embedding

Alireza Bosaghzadeh[1], **Fadi Dornaika**[1,2,*]

[1] *University of the Basque Country, Manuel Lardizabal, 1, 20018, San Sebastian, Spain*

[2] *IKERBASQUE, Basque Foundation for Science, Bilbao, Spain*

Abstract: Local Discriminant Embedding (LDE) was recently proposed to overcome some limitations of the global Linear Discriminant Analysis (LDA) method. Whenever a small training data set is used, LDE cannot directly be applied to high-dimensional data. This case is the so-called small-sample-size (SSS) problem. The classic solution to this problem was applying dimensionality reduction on the raw data (*e.g.*, using Principal Component Analysis (PCA)). This chapter introduces a novel discriminant technique called "Exponential Local Discriminant Embedding" (ELDE). The proposed ELDE can be seen as an extension of LDE framework in two directions. Firstly, the proposed framework overcomes the SSS problem without discarding the discriminant information that was contained in the null space of the locality preserving scatter matrices associated with LDE. Secondly, the proposed ELDE is equivalent to transforming original data into a new space by distance diffusion mapping (similar to Kernel-based non-linear mapping), and then, LDE is applied in such a new space. As a result of diffusion mapping, the margin between samples belonging to different classes is enlarged, which is helpful in improving classification accuracy. The experiments are conducted on four public face databases, Extended Yale, PF01, PIE and FERET. The results show that the performances of the proposed ELDE are better than those of LDE and many state-of-the-art discriminant analysis techniques.

Keywords: Complete Kernel Fisher discriminant method, Distance diffusion mapping, Distance metric learning, Exponential discriminant analysis, Exponential locality preserving projections, Face recognition, Feature extraction, Generalized eigenvectors, Intrinsic graph, Kernel Fisher discriminant analysis, Kernel Principal component analysis, Linear discriminant analysis, Local discriminant embedding, Matrix exponential, Penalty graph, Principal component

* **Address to Corresponding Author Fadi Dornaika:** University of the Basque Country, Manuel Lardizabal, 1, 20018 San Sebastian, Spain; Tel: +34 943018034; Fax: +34 943 015590; E-mail: fadi.dornaika@ehu.es

analysis, Regularization, Regularized Kernel discriminant analysis, Singular matrix, Small sample size problem.

INTRODUCTION

In most computer vision and pattern recognition problems, the large number of sensory inputs, such as images and videos, are computationally challenging to analyze. In such cases it is desirable to reduce the dimensionality of the data while preserving the original information in the data distribution, allowing for more efficient learning and inference [1 - 4]. There are two main reasons for estimating a low-dimensional representation of high-dimensional data: reducing measurement cost of further data analysis and beating the curse of dimensionality. The dimensionality reduction can be achieved either by feature extraction or feature selection. Feature extraction (manifold learning) refers to methods that create a set of new features based on transformations and/or combinations of the original features, while feature selection methods select the most representative and relevant subset from the original feature set. Feature extraction methods can be classified into two main classes: i) linear methods; and ii) non-linear methods. Recently, manifold learning theory has received a lot of attention by researchers and practitioners. Exploiting the findings of manifold learning theory has led to many progresses in face recognition which is known to be a difficult problem in the domain of computer vision [5 - 7].

The non-linear methods such as Locally Linear Embedding (LLE) [8] and Laplacian Eigenmaps [9] focus on preserving the local structures. Isomap [10] is a non-linear projection method that globally preserves the data. It also attempts to preserve the geodesic distances between samples.

The linear techniques have been increasingly important in pattern recognition [2, 11 - 14] since they permit a relatively simple mapping of data onto a lower-dimensional subspace, leading to simple and computationally efficient classification strategies. The classical linear embedding methods (*e.g.*, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Locally LDA [15] are demonstrated to be computationally efficient and suitable for practical applications, such as pattern classification and visual recognition. PCA projects the samples along the directions of maximal variances and aims to preserve the Euclidean distances between the samples. Unlike PCA which is unsupervised, Linear Discriminant Analysis (LDA) is a supervised technique. One limitation of PCA and LDA is that they only see the linear global Euclidean structure.

In addition to the Linear Discriminant Analysis (LDA) technique and its variants [16 - 18], there is recently a lot of interests in graph-based linear dimensionality reduction. Many dimensionality reduction techniques can be derived from a graph whose nodes represent the data samples and whose edges quantify the similarity among pairs of samples [19, 20]. Recently proposed methods attempt to linearize some non-linear embedding techniques. This linearization is obtained by forcing the mapping to be explicit, *i.e.*, performing the mapping by a projection matrix. For example, Locality Preserving Projection (LPP) [21 - 23] and Neighborhood Preserving Embedding (NPE) [21] can be seen as linearized versions of LE and LLE, respectively. The main advantage of the linearized embedding techniques is that the mapping is defined everywhere in the original space. Some researchers tried to remedy to the global nature of the linear methods (*e.g.*, PCA, LDA and LPP) by proposing localized models [24]. In this work, localized PCA, LDA, or LPP models are built using the neighbors of a query sample. The authors have shown that the obtained localized linear models can outperform the global models for face recognition and coarse head pose estimation problems. In [25], the authors have extended the LPP to the supervised case by adapting the entries of the similarity matrix according to the labels of the sample pair. In [26], the authors have proposed an enhanced supervised variant of LPP. The affinity matrix weights are modified in order to take into account label information as well as the similarities between pairs of samples. Moreover, the optimized criterion integrates uncorrelation and orthogonality constraints. In [27], the authors assessed the performance of the quotient and difference criteria used in LDA.

They also proposed a unified criterion that combines Quotient-LDA and Difference-LDA criteria.

In [28], the authors have proposed a discriminant method called Average Neighborhood Margin Maximization (ANMM). It associates to every sample a

CHAPTER 4

Adaptive Locality Preserving Projections for Face Recognition

Fadi Dornaika[1,2,*], **Ammar Assoum**[3]

[1] *University of the Basque Country, San Sebastian, Spain*

[2] *IKERBASQUE, Basque Foundation for Science, Bilbao, Spain*

[3] *LaMA Laboratory, Lebanese University, Tripoli, Lebanon*

Abstract: This chapter addresses the graph-based linear manifold learning for object recognition. In particular, it introduces an adaptive Locality Preserving Projections (LPP) which has two interesting properties: (i) it does not depend on any parameter, and (ii) there is no correlation between mapped data. The main contribution consists in a parameterless computation of the affinity matrix built on the principle of meaningful and Adaptive neighbors. In addition to the framework of LPP, these two properties have been integrated to the framework of two graph-based embedding techniques: Orthogonal Locality Preserving Projections (OLPP) and Supervised LPP (SLPP). After introducing adaptive affinity matrices and the uncorrelated mapped data constraint, we perform recognition tasks on six public face databases. The results show improvement over those of classic methods such as LPP, OLPP, and SLPP. The proposed method could also be applied to other kinds of objects.

Keywords: Affinity matrix, Classification, Dimensionality reduction, En- hanced Locality Preserving Projections, Face recognition, Graph-based linear embedding, Label information, Laplacian eigenmaps, Laplacian matrix, latent points, Linear discriminant analysis, Locality preserving projections, Nearest neighbor classifier, Orthogonal locality preserving projections, Parameter-less locality preserving projections, Pearson's coefficient, principal component analysis, Projection directions, Recognition rate, Supervised locality preserving projections.

[*] **Address to Corresponding Author Fadi Dornaika:** University of the Basque Country, Manuel Lardizabal, 1, 20018 San Sebastian, Spain; IKERBASQUE, Basque Foundation for Science, Bilbao, Spain; Tel: +34 943018034; Fax: +34 943 015590; E-mail: fadi.dornaika@ehu.es

INTRODUCTION

When dealing with pattern recognition and computer vision problems, the large amount of input data, such as images and videos, are computationally challenging to manipulate. In these cases it is preferable to reduce the dimensionality of the data without radically altering the original information in the data distribution. This would mostly result in more efficient learning and inference. If the variance of the multivariate data is faithfully represented as a set of parameters, the data can be considered as a set of geometrically related points lying on a smooth low-dimensional manifold. The fundamental issue in dimensionality reduction is how to model the geometry structure of the manifold and produce a faithful embedding for data projection. A large number of approaches have been proposed for computing the embedding. The linear methods, such as multidimensional scaling (MDS) [1] and principal component analysis (PCA) are characterized by their efficiency in observing the Euclidean structure. Unlike PCA which is unsupervised, linear discriminant analysis (LDA) is a supervised technique.

In pattern recognition, linear dimensionality reduction (LDR) techniques have been increasingly used [2, 3] since they allow to build a relatively simple mapping of data onto a lower-dimensional subspace. Recently, manifold learning theory has received a lot of attention by researchers and practitioners. Exploiting the findings of manifold learning theory has led to many progresses in face recognition which is known to be a difficult problem in the domain of computer vision [4 - 6]. The linear methods of dimensionality reduction have the advantage over the non-linear ones in that the embedding function of the former is defined everywhere in the input space, while for the latter, it is only defined for a set of data samples. Several linear dimensionality reduction algorithms are based on a graph whose nodes consist of the data samples and whose edges represent the similarity among these samples [7]. For example, the Locality Preserving Projections (LPP) [8, 9] is a typical graph-based method used for Linear Dimensionality Reduction. LPP is a linearized version of Laplacian Eigenmaps [10] and has been successfully applied in many practical problems such as speech recognition [11], face recognition [12] and age estimation [13]. In order to avoid possible singularities, LPP is generally preceded by a PCA step.

Several extensions to the LPP method have been proposed in the literature. For example, Xu *et al.* [14] added new features to improve the original LPP by i) introducing linear transforms prior to LPP, and ii) changing the original quotient-based criterion to a new difference-based one. Furthermore, although LPP is a unsupervised LDR technique, it is shown that, in some cases it can outperform some supervised techniques such as LDA [15]. This is mainly due to the fact that LPP preserves the locality structures of data. Orthogonal LPP (OLPP) [16] is an other extension of LPP that uses its criterion in order to provide orthogonal projection directions. Moreover, [8] proposes a regularized version of LPP that solves the singularity of the matrix $\mathbf{XDX}^T$. Although the LPP method intrinsically preserves the manifold structure of the input data, its ability to discriminate between different classes is little because of its unsupervised nature. Indeed, the estimation of the linear transform that maps the input data through the LPP framework does not take into account the label information. That is why, a supervised version of LPP (SLPP) is proposed in order to overcome this limitation [17]. The key idea of SLPP is that the computation of the affinity matrix is based on the intraclass neighborhood of each point, *i.e.*, the nearest neighbors belonging to the same class as its.

In this chapter, we propose an improved version of LPP that offers two interesting advantages: i) it is entirely parameter-free, and ii) there is no correlation between mapped data. The chapter is organized as follows. Section 1 describes briefly the original linear mapping of LPP. Section 2 depicts the proposed adaptive LPP. Section 3 summarizes the experimental results for face recognition experiments conducted on five face databases. In the sequel, capital bold letters denote matrices and small bold letters denote vectors.

LOCALITY PRESERVING PROJECTIONS

We assume that we have a set of N samples $\{\mathbf{x}_i\}_{i=1}^{N} \subset \mathbb{R}^D$. We Construct a neighborhood graph on these data, such as a full mesh, a k-nearest-neighbor or ε-ball graph. The weight A_{ij} of each edge $\mathbf{x}_i \sim \mathbf{x}_j$ is computed by a symmetric affinity function $A_{ij}= K(\mathbf{x}_i; \mathbf{x}_j)$, typically Gaussian, *i.e.*,

CHAPTER 5

Face Recognition Using 3D Face Rectification

Alireza Bosaghzadeh[1,*], Mohammadali Doostari[2], Alireza Behrad[2]

[1] *University of the Basque Country, San Sebastian, Spain*

[2] *Shahed University, Tehran, Iran*

Abstract: While face recognition algorithms have shown promising results using gray level face images, their accuracy deteriorate if the face images are not frontal. As the head can move freely, it causes a key challenge in the problem of face recognition. The challenge is how to automatically and without manual intervention recognize non-frontal face images in a gallery with frontal face images. The rotation is a linear problem in 3D space and can be solved easily using the 3D face data. However, the recognition algorithms based on 3D face data gain less recognition rates than the methods based on 2D gray level images. In this chapter, a sequential algorithm is proposed which uses the benefits of both 2D and 3D face data in order to obtain a pose invariant face recognition system. In the first phase, facial features are detected and the face pose is estimated. Then, the 3D data (Face depth data) and correspondingly the 2D image (Gray level face data) are rotated in order to obtain a frontal face image. Finally, features are extracted from the frontal gray level images and used for classification. Experimental results on FRAV3D face database show that the proposed method can drastically improve the recognition accuracy of non-frontal face images.

Keywords: 3D rotation, Biometric, Depth data, Dimensionality reduction, Ellipse fitting, Eigen problem, Eigenface, Face recognition, Facial features, Fisherface, Feature extraction, Gray level image, IRAD contours, Linear Discriminant Analysis, Least mean square, Manifold learning, Mean filter, Mean curvature, Nearest Neighbor classifier, Pose estimation.

INTRODUCTION

A lot of researchers has focused on face recognition as it has a variety of applications in surveillance systems, law enforcement, human-computer interfaces

* **Address to Corresponding author Alireza Bosaghzadeh:**University of the Basque Country, Paseo Manuel Lardizabal, 1, 20018, San Sebastian, Spain; Tel: +34 943015060; Fax: +34 943015590; E-mail: alireza.bosaghzadeh@ehu.es

Fadi Dornaika (Ed)

and access controls [1]. Also, it observes a lot of attention in the biometric area since capturing a face image is passive and non intrusive compared to other biometrics like fingerprint or iris. Variations in illumination conditions, aging, facial expressions and face pose make this task difficult [2 - 10].

Different algorithms were proposed to overcome these variations. Wang *et al.* [11] used a set of images with different variations in lightening and head movement for training. They model each image set as a manifold and proposed a manifold-manifold distance to compare the distance between each two manifolds. Arandjelovic *et al.* [12] proposed a "re-illumination" algorithm to solve the lightening condition variations. They obtained high recognition rates on databases with large variation in capturing conditions.

Among different variations, since the head can move easily, the face pose variation is one of the most important challenges. Although, most of the available face recognition techniques obtain high recognition rates on frontal faces [13, 14], their results deteriorate if the face is not frontal. Thus, handling the pose variations between the gallery face images and the test image is an extremely important factor for many practical application scenarios. A variety of methods have been proposed specifically to address the pose invariance face recognition [15 - 19].

To cover the possible appearances under all horizontal rotations, Singh *et al.* [20] construct a panoramic view of a face by using a mosaicing scheme. They use frontal view and rotated views to generate the panoramic view. In the recognition phase, the synthesised face mosaics were matched with face images in arbitrary poses by a SVM classifier.

In [18], a 3D Generic Elastic Model is used to construct the 3D model of each subject using only a single 2D image, which can match the images in the same pose as the test image. Before matching, a linear regression approach is used to estimate the pose of the test image. Then, all 3D models are rendered in different poses around the estimated pose and the rendered images are used for matching against the test query. Finally, by normalized correlation matching, the distances between the test query and the synthesized images are computed.

In [21], the authors proposed a method to synthesize a frontal view from a non-frontal face image. They divide the face into overlapping patches and estimate the local warp in order to obtain the frontal view of the patches. To find the optimal warps, they use a discrete Markov random fields and a belief propagation method in order to formulate the optimization problem as a discrete labeling algorithm.

In [17], a multi-pose face recognition approach using fusion of scale invariant features (FSIF) is proposed. They fuse some scale invariant features extracted by scale invariant features transforms (SIFT) from several different poses of 2D face images. Finally, Linear Discriminant Analysis (LDA) dimensionality reduction is used to extract the optimum features for classification.

The authors in [22] combine two feature descriptors namely Gabor Wavelet and enhanced LBP features for feature extraction. Then, they use a generalized neural network for classification. Their method is efficient only for slight variations in pose and deteriorates with large pose variations. Later in [19], they use higher-order moments of curvelet as features and a curvelet neural network for classification. The experimental results show that their proposed method achieve higher accuracy for pose invariant face recognition than standard back propagation neural network.

PROPOSED METHOD

In this chapter, we introduce a novel pose invariant face recognition technique which can overcome the rotation problem in face recognition. As pose variation is non-linear in two-dimensional (2D) space but linear in three-dimensional (3D) space, we solve the pose variation problem in the 3D space and then use it to map 2D data. The proposed method uses only frontal images for training and performs matching on test images with pose variations. It has a sequential procedure to eliminate the effect of pose variation in the test data. In the first step, preprocessing tasks are done on the data to remove noises from the 3D face data. In the second phase, facial features in the 3D image are located. In the third step, the 3D rotation of the face (*i.e.*, face pose) is estimated. Knowing the face pose, 3D face data and correspondingly the gray level face image are rotated in order to

CHAPTER 6

3D Face Recognition

Alireza Behrad*

Department of Electrical and Electronic Engineering, Shahed University, Tehran-Qom Exp. Way, 3319118651, Tehran, Iran

Abstract: 3D face recognition algorithms are a group of methods which utilize 3D geometry of face and facial feature for recognition. In comparison with 2D face recognition algorithms that employ intensity or color based features, they are generally robust against lighting condition, head orientation, facial expression and make-up. 3D face recognition has several advantages. Firstly, the shape and the related features of 3D face can be acquired independent from lighting condition. Secondly, the pose of 3D face data can be easily corrected and used for subsequent pose invariant feature extraction. Thirdly, 3D face data are less affected by skin color, face cosmetic and similar face reflectance factors. 3D face recognition may include several stages such as 3D image acquisition, face localization, feature extraction and face recognition. In this article, different algorithms and the pipeline for 3D face recognition are discussed.

Keywords: 3D Face matching, 3D Face recognition, 3D Face registration, 3D Image acquisition, 3D Surface descriptors active triangulation, Biometric identifiers, Curvature descriptors, Face alignment, Face analysis, Face recognition face segmentation, Facial features, Facial landmarks, Feature extraction, Head pose estimation, Human identification, Laser scanner, Nose tip detection, Range images.

INTRODUCTION

Recently the use of biometric information for verification and identification has been getting more and more importance. Generally, biometric information includes various information regarding the human's body, feature or characteristics like height, hair color, finger print, face shape, *etc*. However, some

* **Address to Corresponding Author Alireza Behrad:** Electrical and Electronic Engineering Department, Faculty of Engineering, Shahed University, Tehran-Qom Exp. Way, 3319118651, Tehran, Iran; Tel: (+98-21) 51212070; Fax: (+98-21) 51212020; E-mail: behrad@shahed.ac.ir

Fadi Dornaika (Ed)

biometrics can be easily changed and cannot be used as unique feature for identification or verification. A biometric identifier is defined as a unique, distinctive and measurable identifier that can be used as a means to describe, verify or identify a person. Fingerprint, iris, palm print and hand motion are general biometric identifiers.

Recently face recognition is increasingly employed as a tool for human identification and verification. Face recognition is defined as a process for identification or verification of a human person based on the characteristics of face. For humans, faces are the most important characteristic for human recognition and we generally know and recognize each other based on the face information. Face recognition is an identification tool at distance. It means that the subject under the identification can be in an arbitrary place unaware of identification process. Because of exclusive characteristics of face recognition systems, they are widely used in security system for human identification and verification. In addition to human recognition, facial features can be used in various applications like human computer interface, age, gender or race analysis, expression determination and so forth.

Face recognition algorithms may be based on 2D or 3D information of the face. 2D face recognition algorithms use intensity or textured based information for the sake of face recognition [1 - 3]. On the contrary, 3D face recognition algorithm use 3D model of face or 3D shape information for face recognition [4 - 6]. Some algorithms also employ mixture of 2D-3D information to enhance the efficiency of face recognition algorithms [7 - 9]. 3D face recognition has several advantages. Firstly, the shape and the related features of 3D face can be acquired independent from lighting condition. Secondly, the pose of 3D face data can be easily corrected and used for subsequent pose invariant feature extraction. Thirdly, 3D face data are less affected by skin color, face cosmetic and similar face reflectance factors.

There are some limitations with 3D face recognition comprising 3D face acquisition, change of face shape with age and expression to name a few. The acquisition of 3D or range images is the main limitation of 3D face recognition algorithms. Laser scanners are general equipment to capture 3D images; however they may damage eyes and are not proper for real-time applications. Therefore, stereo imaging or methods like shape from shading are generally employed to capture 3D face images.

Face localization is the first stage in 3D face recognition. Depending on 3D face acquisition equipment, the 3D face image may include various noises like spikes, holes and distortions; therefore different preprocessing algorithms may be applied to improve the quality of 3D data for further processing. When the noise-free 3D image is obtained, the registration algorithm is employed for the alignment of the 3D faces. The registration stage generally requires facial feature extraction which is an important step for efficient face recognition. Finally, feature extraction algorithm is utilized to extract proper features for recognition.

In this article, various stages for 3D face recognition are discussed and different algorithms for each stage are explained. The rest of this article is organized as follows. In the next section, 3D face acquisition systems are discussed. Section 3 describes various 3D face representation approaches. Section 4 deals with some preprocessing steps for 3D face recognition. Section 4 presents some existing algorithms for 3D face alignment. Various 3D face recognition algorithms appear in section 5 and we conclude the article in section 6.

3D FACE ACQUISITION

Laser scanners are mostly used equipment for capturing 3D images. Laser scanners may be constructed based on the phase shift, time-of-flight and active triangulation principles. In phase shift approach, laser beam with sinusoidally modulated optical power is sent toward the target. Then the phase shift between the sent and reflected light is measured to calculate the distance between laser source and the target. In addition to phase shift, it is also necessary to have the number of full cycles that a light wave undergoes during the transmitted path. In time-of-flight approach, a very short laser pulse is sent to the target and the return light is detected to measure the round-trip time. The round-trip time together with light speed in the medium is utilized to measure the distance between the target and laser emitting device. Time-of-light approach is a suitable method for long-distance measurement. However for near distances, laser scanning based on active

CHAPTER 7

Model-Less 3D Face Pose Estimation

Fawzi Khattar[1], Fadi Dornaika[2,3], Ammar Assoum[1,*]

[1] *LaMA Laboratory, Lebanese University, Tripoli, Lebanon*

[2] *University of the Basque Country, San Sebastian, Spain*

[3] *IKERBASQUE, Basque Foundation for Science, Bilbao, Spain*

Abstract: Automatic head pose estimation consists of using a computer to predict the pose of a person based on a given facial image. Fast and reliable algorithms for estimating the head pose are essential for many applications and higher-level face analysis tasks. Many of machine learning-based techniques used for face detection and recognition can also be used for pose estimation. In this chapter, we present a new dimensionality reduction algorithm based on a sparse representation that takes into account pose similarities. Experimental results conducted on three benchmarks face databases are presented.

Keywords: Age classification, Age estimation, Age prediction, Dimensionality reduction, Facial feature extraction, Gabor filter, K-nearest neighbors, Label-sensitive, Local binary pattern, Local regression, Locality preserving projections, Machine learning, Marginal fisher analysis, Mean absolute error, Partial least square regression, Preprocessing, Recognition rate, Support vector regression.

INTRODUCTION

In the domain of computer vision, the identification of specific objects within an image is a well-known and typical task. The aim of the identification is to determine the objects' position and orientation with respect to a given coordinate system. The obtained information may then be used for several purposes, for example to manipulate an object by a robot or to prevent this latter from hitting obstacles. In the related terminology, the pose of an object refers to the combination of its position and orientation, even though it is sometimes used in line with the orientation by itself.

* **Address to Corresponding author Ammar Assoum:** LaMA Laboratory, Lebanese University, Faculty of Science, Section III, Tripoli, Lebanon; Tel: +9613069591; Fax: +9616386365; E-mail: a.assoum@ul.edu.lb

Particularly, human's facial pose is considered as an important cue of non-verbal communication. Indeed, humans can discover and understand other people's intentions easily by interpreting their head pose. However, in order to make a machine capable of interacting with human's head movements and expressions, huge effort has to be done to estimate the pose from the pixel representation of a facial image in a robust and efficient way. The estimation process requires a series of processing steps to transform a pixel-based representation of a head into a high-level concept of direction.

STATE OF THE ART

The head pose estimation refers to the specific task consisting of determining the position and/or the orientation of the head in an image (*e.g.* facial one). This task is a challenging problem because there are many degrees of freedom that should be estimated. During the past years many techniques and approaches have been proposed to solve this problem [1, 2]. Appearance template methods use similarity algorithms and compare a given image to a set of exemplars in order to discover which labelled image (template) is the most similar to the one the pose of which is to be estimated [3, 4]. Nevertheless, even if these methods have the advantage of not requiring a features extraction step, they may suffer from noise caused by illumination and expression changes in addition to the need of high computational power since the matching process they use is based on pair-wise similarities. Classification-based methods [5] operate by training head pose classifiers through the distribution of the training images into a a set of discretized poses. However, both the appearance template and classification-based methods can only return discrete poses and are sensitive to non-uniform sampling in the learn data. Regression-based methods [6] allow to obtain continuous pose estimates. Indeed, they use regression techniques [7, 8] in order to find the relationship between the face image and its corresponding pose and to earn continuous mapping functions between the face image and the pose space. The high dimensionality of the data represents an important challenge in this kind of methods because of the well-known "curse of dimensionality problem" [9]. Many researchers use a dimensionality reduction step before the regression [10, 11]. The main disadvantage of these methods is that their performance deteriorates with bad head localization. The manifold embedding methods [12] consider face images as samples of a low-dimensional manifold embedded in the high-dimensional observation space (the space of all possible images). They try to find a low dimensional representation that is linked to the pose. After that, classification or regression techniques are applied to discover the pose. The main weakness of manifold embedding methods is that appearance variation is not only affected by pose changes but also by other factors such as lighting changes and identity. Geometric methods [13, 14] rely heavily on the estimation of facial features, such as eyes, mouth corners, nose tip, etc. and use their relative position to estimate the pose using projective geometry. For example, if the eyes and the mouth form an isosceles triangle, then the image corresponds to a frontal view. The major disadvantage of these methods is to locate the features needed for estimation in a very precise and accurate way. They also need to handle missing facial features in some poses. The tracking methods track the head and use the relative movement of the face with temporal continuity and smooth motion constraint to estimate the pose. These methods require an initialization step for the 3D head pose parameters [15]. The detector array methods train different detectors for different head poses. These detectors try to detect the face in a new image. It is supposed that only the detector trained for the same exact pose will be able to detect the face in the image. The detector that succeeds to detect the face assigns its pose to the image [16].

THE MACHINE LEARNING METHODOLOGY

The machine learning pipeline used in our work is summarized in Fig. (**1**). It is divided into the following steps:

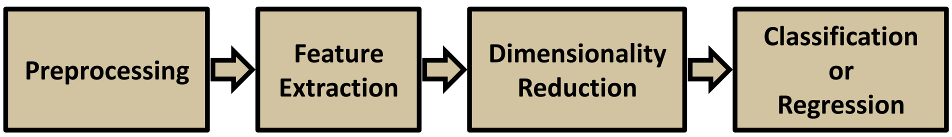


Fig. 1. The proposed machine learning pipeline.

Preprocessing: In this step, image processing techniques are used to prepare the images and put them in a sort of normalized way. The RGB images are converted to grayscale because colors may not bring pertinent information about the pose. During this step we may perform a face alignment and cropping in order to

CHAPTER 8

Efficient Deformable 3D Face Model Fitting to Monocular Images

Luis Unzueta[1,*], Waldir Pimenta[2], Jon Goenetxea[1], Luís Paulo Santos[2], Fadi Dornaika[3,4]

[1] *Vicomtech-IK4, Paseo Mikeletegi, 57, Parque Tecnológico, 20009, Donostia, Spain*

[2] *Departamento de Informática, University of Minho. Campus de Gualtar, 4710-057, Braga, Portugal*

[3] *Computer Engineering Faculty, University of the Basque Country EHU/UPV, Manuel de Lardizabal, 1, 20018, Donostia, Spain*

[4] *Ikerbasque, Basque Foundation for Science, Alameda Urquijo, 36-5, Plaza Bizkaia, 48011, Bilbao, Spain*

Abstract: In this work, we present a robust and lightweight approach for the automatic fitting of deformable 3D face models to facial pictures. Well known fitting methods, for example those taking into account statistical models of shape and appearance, need a training stage based on a set of facial landmarks, manually tagged on facial pictures. In this manner, new pictures in which to fit the model cannot differ excessively in shape and appearance (including illumination changes, facial hair, wrinkles, and so on) from those utilized for training. By contrast, our methodology can fit a generic face model in two stages: (1) the localization of facial features based on local image gradient analysis; and (2) the backprojection of a deformable 3D face model through the optimization of its deformation parameters. The proposed methodology preserves the advantages of both learning-free and learning-based methodologies. Subsequently, we can estimate the position, orientation, shape and actions of faces, and initialize user-specific face tracking approaches, such as Online Appearance Models (OAMs), which have demonstrated to be more robust than generic user tracking methodologies. Experimental results demonstrate that our strategy outperforms other fitting methods under challenging illumination conditions and with a computational footprint that permits its execution in gadgets with reduced computational power, such as cell phones and tablets. Our proposed methodology fits well with numerous systems addressing semantic inference in face images and videos.

[*] **Address to Corresponding Author Luis Unzueta:** Vicomtech-IK4, Paseo Mikeletegi, 57, Parque Tecnológico, 20009, Donostia, Spain; Tel: +[34] 943 30 92 30; Fax: +[34] 943 30 93 93; E-mail:lunzueta@vicomtech.org

Fadi Dornaika (Ed)

Keywords: 2D shape landmarks, 3D face model, Deformable model back projection, facial actions, Facial expression recognition, Facial feature extraction, facial parts, Face gesture analysis, Face model fitting, Face recognition, Face tracking, Gradient maps, Head pose estimation, Learning-free, Levenberg-Marquardt, Monocular image, Online appearance model, Shape variations, Sigmoidal filter, Weak perspective.

INTRODUCTION

Generic face model fitting has been a hot research topic during the last decade. It can be seen as an essential part in numerous Human-Computer Interaction applications since it allows face tracking, head pose estimation, identification, and face gesture analysis. In general terms, two types of methods have been proposed: (i) learning-free and (ii) learning-based. The latter require a training stage with many pictures to construct the model, and therefore rely on the choice of pictures for a good fitting in unseen pictures. Learning-free methodologies depend intensely on some radiometric and geometric properties present in face pictures. These methodologies rely on generic knowledge about faces, which usually incorporates the position, symmetry, and edge profile of facial organs. They can place facial features using low-level methods (*e.g.* filtering, gradients), typically relying on recognizing individual face features (lips, nose, irises, ...) [1 - 4]. A large portion of the learning-free methodologies do not produce a full collection of extracted face features, contrary to learning-based strategies.

For example, in [5], the authors exploit a range facial scan in order to automatically distinguish the nose tip for both frontal and non frontal poses. In [7], an incremental certainty methodology regarding the extraction of facial features over real video frames is explained. The proposed procedure adapts to large varieties of subject appearances, including frame-to-frame changes within video sequences. The framework identifies the zones of the face that are measurably exceptional and assembles an initial set of regions that are expected to incorporate data about the features of interest. In this methodology, core facial features, for example the eyes and the mouth, are in effect reliably identified. In [6], the authors try to recognize the eyes and mouth utilizing the separation vector field that is structured by attributing a vector to every pixel indicating its nearest edge. Separation vector fields are based on geometrical structure, and consequently can help in evading illumination issues in the location of the eyes and mouth areas. In [9], the authors demonstrated that the eyes and mouth in facial pictures can be robustly identified. They used their locations to normalize the pictures, assuming affine transformation, which can make up for different viewpoints. In [10], real-time face detection algorithm for searching faces, eyes and lips in pictures and videos is explained. The calculation builds upon the extraction of skin pixels based on rules derived from a straightforward quadratic polynomial model in a normalized color space. In [8], the authors separated the facial feature extraction into three core steps. The initial step is preprocessing. The objective of this step is to get rid of high intensity noise and to binarize the input picture. The second step incorporates a labeling procedure and an aggregation procedure. This step tries to create facial feature candidates block by block. Finally, a geometrical face model is utilized to detect the face position.

As can be seen, learning-free methodologies have appealing characteristics. Nonetheless, they present a few deficiencies. Firstly, the majority of them makes the assumption that a few conditions are met (for instance, that face pictures are taken in controlled conditions and in an upright orientation). Furthermore, they usually depend on the discovery of few facial features (primarily the eyes and the mouth). Almost no consideration is given to the assembly of an extensive collection of facial features. Thirdly, accurate localization of the detected face features is still faulty.

Learning-based methodologies, on the other hand, aim to overcome these deficiencies. Three subcategories can be identified: parameterized appearance models, part-based deformable models and discriminative methodologies.

Parameterized appearance models generate a statistical model of shape and appearance from a collection of manually marked data [11 - 15]. In the 2D data domain, Active Shape Models (ASM) [11, 16], Active Appearance Models (AAM) [13, 14] and more recently, Active Orientation Models (AOM) [15] have been proposed. The ASM methodology generates 2D shape models and relies on motion constraints in conjunction with some image data from the regions near the 2D shape landmarks to find features on new pictures. The AAM uses both the

CHAPTER 9

Face Detection Using the Theory of Evidence

Franck Luthon*

Computer Science Lab., University of Pau, 2 allée parc Montaury, 64600, Anglet, France

Abstract: Face detection and tracking by computer vision is widely used for multimedia applications, video surveillance or human computer interaction. Unlike current techniques that are based on huge training datasets and complex algorithms to get generic face models (*e.g.* active appearance models), the proposed approach using evidence theory handles simple contextual knowledge representative of the application background, *via* a quick semi-supervised initialization. The transferable belief model is used to counteract the incompleteness of the prior model due to lack of exhaustiveness in the learning stage.

The method consists of two main successive steps in a loop: detection, then tracking. In the detection phase, an evidential face model is built by merging basic beliefs carried by a Viola-Jones face detector and a skin color detector. The mass functions are assigned to information sources computed from a specific nonlinear color space. In order to deal with color information dependence in the fusion process, a cautious combination rule is used. The pignistic probabilities of the face model guarantee the compatibility between the belief framework and the probabilistic framework. They are the inputs of a bootstrap particle filter which yields face tracking at video rate. The proper tuning of the few evidential model parameters leads to tracking performance in real-time. Quantitative evaluation of the proposed method gives a detection rate reaching 80%, comparable to what can be found in the literature. Nevertheless, the proposed method requires a scanty initialization only (brief training) and allows a fast processing.

Keywords: Belief function, Cautious rule, Classification, Computer vision, Conjunctive rule, Dempster-Shafer, Face tracking, Fusion of information, LUX color space, Mass set, Particle filter, Pattern recognition, Pignistic probability, Region of interest, Skin hue, Source of information, Transferable belief model, Uncertainty management, Viola-Jones detector, Visual servoing.

* **Address to Corresponding Author F. Luthon:** IUT de Bayonne Pays Basque, Université de Pau Pays d'Adour, 2 allée du parc Montaury, 64600, Anglet, France; Tel: +33(0)5.59.57.43.44; Fax: +33(0)5.59.57.43.49; E-mail: Franck.Luthon@univ-pau.fr

Fadi Dornaika (Ed)

INTRODUCTION

Real-time face detection and tracking in video sequences has been studied for more than twenty years by the computer vision and pattern recognition community, owing to the multiplicity of applications: teleconferencing, closed-circuit television (CCTV), human machine interface and robotics. Despite the ongoing progress in image processing and the increase in computation speed of digital processors, the design of generic and robust algorithms is still the object of active research. Indeed, face image analysis (either detection, recognition or tracking) is made difficult by the variability of appearance of this deformable moving object due to many factors: individual morphological differences (nose shape, eye color, skin color, beard), presence of visual artifacts (glasses, occlusions, make-up), illumination variations (shadow, highlight) and facial expression changes depending on context (social, cultural, emotional). Those are difficult to model and do not easily cope with real-time implementations. Moreover, the scene background might disturb detection, in case of foreground-background similarity or background clutter.

To handle the face specificity, a semi-supervised learning method is presented here, where the user selects manually a zone of the face in the first image of the video. This rapid initializing step constitutes the learning stage which yields simply a prior model for face class and background class. It is however dependent on the user subjectivity while selecting the face zone and it suffers from incompleteness because of lack of exhaustiveness of this short training. In this context, a probabilistic modeling is not relevant. Therefore, the proposed approach is based on belief functions: indeed the transferable belief model (TBM) is well suited to model partial knowledge in a complex system [1]. It was successfully applied to classification of emotions and facial expressions, or to human activity recognition [2].

The goal of the application is to automatically track the face of a person placed in the field of view of a motorized pan-tilt-zoom camera (or simply a webcam). The tracking technique should be as robust as possible to occlusions, pose, scale, background and illumination changes. It should take control of the camera to perform an automatic centering of the face in the image plane during the whole video sequence. The algorithm consists of two main steps: face detection and then tracking (Fig. **1**). An elliptical region of interest (ROI) including the face is computed by particle filtering, and held at the center of the image by visual servoing. The context of application is indoor environments, typically a laboratory or an office. As regards acquisition conditions, the distance between user and sensor ranges from about 50 cm to a few meters. Ordinary lighting conditions prevail (uncontrolled illumination context), possibly in the presence of additional light sources, like a desk lamp or the influence of outside light entering through a window.

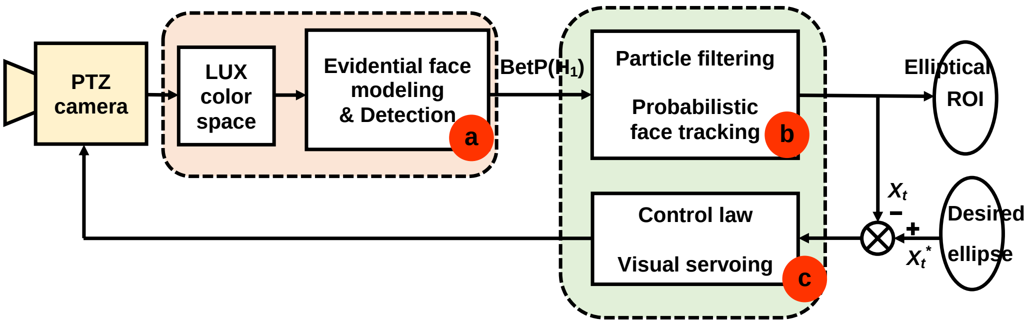


Fig. (1). Overview of the processing with feedback loop: a) face detection by evidential modeling; b) face tracking by particle filtering; c) camera control by visual servoing.

After a state of the art about face detection, the theory of belief functions is briefly exposed. The proposed evidential model for face detection is then detailed in the application section. The tracking with particle filter and visual servoing of the camera are described. Performance analysis, both qualitative and quantitative, is presented. The chapter ends with a discussion.

STATE OF THE ART

Face detection methods may be grouped into two categories differing in the way of processing prior information [3]. It is also worth making a difference between detection methods dedicated to still images, where complex algorithms can be used, and methods dedicated to video sequences where the computation cost is of major concern for real-time processing.

CHAPTER 10

Fuzzy Discriminant Analysis: Considering the Fuzziness in Facial Age Feature Extraction

Shenglan Ben*

School of Electronic Science and Engineering, Nanjing University, Nanjing, 210023, China

Abstract: In traditional age estimation methods which utilize discriminative methods for feature extraction, the biological age labels are adopted as the ground truth for supervision. However, the appearance age, which is indicated by the facial appearance, is intrinsically a fuzzy attribute of human faces which is inadequate to be labeled as a crisp value. To address this issue, this paper firstly introduces a fuzzy representation of age labels and then extends the LDA into fuzzy ones. In the definition of fuzzy labels, both the ongoing property of facial aging and the ambiguity between facial appearance and biological age are considered. By utilizing the fuzzy labels for supervision, the proposed method outperforms the crisp ones in both preserving ordinal information of aging faces and adjusting the inconsistency between the biological age and appearance. Experiments on both FG-NET and MORPH databases confirm the effectiveness of the proposed method.

Keywords: Age information, Aging pattern subspace, Appearance age, Biological age, Conformal embedding analysis, Cumulative score, Facial age estimation, Facial appearance, Facial landmark, Facial shape, Fuzziness in Facial aging, Fuzzy discriminant analysis, Fuzzy LDA, Fuzzy representation of age labels, Intra-class and inter-Class neighbors, Linear discriminant analysis, Marginal Fisher analysis, Mean absolute error, Ordinary preserving LDA, Ordinary preserving MFA.

INTRODUCTION

The objective of facial age estimation is to evaluate a person's age from facial images [1]. It has wide applications in human-computer interactions such as

* **Address to Corresponding Author Shenglan Ben:** School of Electronic Science and Engineering, Nanjing University, Nanjing, 210023, China; Tel: +86 25 89686705; Fax: +86 25 84317235; E-mail: benshenglan@mail.njust.edu.cn

internet access control, underage alcohol vending machine and age specific advertising. Attracted by the potential applications, many researchers have devoted to investigate techniques to achieve reliable age estimation. Among the research topics, feature extraction plays a crucial role in determining the performance of the system since facial appearance is an intricate composition of age, identity, pose, expression and illumination.

Early works utilized the ratios of the distance between facial landmarks as the age related features [2, 3]. Obviously, such anthropometric features are sensitive to the accuracy of landmark localization. And, they are only effective in classifying the face images into children and adult because the facial shape does not change too much during adult aging.

To solve such problem, Lanitis [4 - 6] utilized the active appearance model (AAM) [7], which combines both the shape and texture information of a face image, in regression based age estimation methods. Also using AAM based face encoding, Geng *et al*. [8] proposed aging pattern subspace (AGES) method, which learns a subspace from sequences of individual aging images. The AAM based methods utilized PCA to do dimension reduction and thus may omit some useful information such as facial wrinkles. The supervised methods, which benefit from the information of age labels, are also widely adopted to extract the discriminative features for facial age estimation. Fu *et al*. [9, 10] utilized manifold embedding techniques such as Marginal Fisher Analysis (MFA) [11] and Conformal Embedding Analysis (CEA) [12], to learn discriminative subspace. In [13, 14], the biologically- inspired features (BIF) were extended to age estimation and combined with manifold learning to get even lower age estimation errors.

To preserve the ordinal information of age labels, Li [15] formulated feature extraction as an optimization problem which considered the temporally ordinal and continuous characteristic of facial aging in the objective function. Lu [16, 17] constructed an age locality graph for the labels of training samples and proposed ordinary preserving LDA (OPLDA) and ordinary preserving MFA (OPMFA) to project the samples with similar ages closer than those with dissimilar ages.

In the above methods, crisp biological ages/age groups are used as the ground truth for supervision. We notice that the fuzziness is an intrinsic feature of facial aging. Firstly, the fuzziness also exists in continuous age labeling because facial aging is an ongoing process. It is ambiguous to label a face image with an exact value of ‘year’. Secondly, the fuzziness exists in the relationship between the biological age and facial appearance since facial aging is a personalized procedure. Some people may look younger or older than their biological age. Thus, it is inadequate to use the biological age as the ground truth of appearance age.

Considering the above observations, we propose methods to solve the fuzziness in extracting the age related features from facial image. In the proposed method, the fuzziness in age labeling is firstly tackled by viewing the age labels as fuzzy sets, and the fuzziness between facial appearance and biological age can then be handled by combining the fuzzy labels of a sample’s neighborhoods. Using the fuzzy age labels for supervision, we extend LDA into fuzzy one. Experimental results on FGNET [18] and MORPH [19] database demonstrate that the proposed method can extract age features from face images efficiently and outperforms several state of the art methods.

The remainder of the paper is organized as follows. The fuzzy representation of age labels is firstly illustrated. A fuzzy discriminant analysis approach is then described. Experimental results and comparisons are made and conclusions are given.

CHARACTERIZATION OF THE FUZZINESS IN FACIAL AGING

In traditional age estimation methods, each sample is labeled with a crisp age value which is typically consistent with the biological age. Supervised feature extractions can then be conducted by viewing each age as a class. However, the appearance age, which is indicated by the facial appearance, is intrinsically a fuzzy attribute of human faces, and is inadequate to be labeled as a crisp value.

The fuzziness of facial age is in two folds. Firstly, facial aging is an ongoing process. Each person is in the period of transition from one exact age to another in most of time. It is inadequate to label a face image with an exact age. Secondly, facial aging is an intricate progress depending on various factors, including human

CHAPTER 11

Facial Image-Based Age Estimation

Ammar Assoum*, Jouhayna Harmouch

LaMA Laboratory, Lebanese University, Tripoli, Lebanon

Abstract: Automatic age estimation consists of using a computer to predict the age of a person based on a given facial image. The age prediction is built on distinct patterns emerging from the facial appearance. The interest of such process has increasingly grown due to the wide range of its potential applications in law enforcement, security control, and human-computer interaction. However, the estimation problem remains challenging since it is influenced by a lot of factors including lifestyle, gender, environment, and genetics. Many recent algorithms used for automatic age estimation are based on machine learning methods and have proven their efficiency and accuracy in this domain. In this chapter, we present an empirical study on a complete age estimation system built around label sensitive learning [1]. Experimental results conducted on FG-NET and MORPH Album II face databases are presented.

Keywords: Age classification, Age estimation, Age prediction, Dimensionality reduction, Facial feature extraction, Gabor filter, K-nearest neighbors, Label-sensitive, Local binary pattern, Local regression, Locality preserving projections, Machine learning, Marginal fisher analysis, Mean absolute error, Partial least square regression, Preprocessing, Recognition rate, Support vector regression.

INTRODUCTION

Despite of its relative newness, automatic age estimation from facial images has recently emerged among the interesting new technologies due to its multiple potential applications. Indeed, taking advantage of rapid progress in computer vision, pattern recognition and machine learning, this technique can be widely used in such areas as age-based access control, age-adaptive human computer interaction, person identification, data mining and organization, multimedia communication and age-targeted advertising and entertainment [2, 3]. Automatic

* **Address to Corresponding Author Ammar Assoum:** Lebanese University, Faculty of Science, Section III, Tripoli, Lebanon; Tel: +9613069591; Fax: +9616386365; E-mail: a.assoum@ul.edu.lb

Fadi Dornaika (Ed)

age estimation is achieved using computers and is useful in scenarios where one does not need to explicitly find the identification of the individual, but wants to know his or her age.

However, estimating accurately the age from facial images is a particularly challenging task since it depends on many complicated and unrelated factors, such as ethnic origin, working environment, living style, health condition and sociality [4, 5]. On the other hand, visual facial features used in the age estimation process are affected by pose, illumination and imaging conditions [6, 7]. Thus, face aging is uncontrollable and personalized [8, 9].

In addition to the challenges cited above, three other factors should also be considered when developing a consistent age estimation system. Firstly, the ordinal relationship that exists between age labels; for instance, age 40 is closer to age 35 than to age 15. This fact makes the age determination learning task more complicated than the traditional classification problem since this latter assumes there is no correlations between classes. Indeed, most of the dimensionality reduction algorithms and distance metric learning techniques ignore this kind of correlations and are well suited, due to their original design, to the traditional classification problems. That is why, techniques like regression and cost-sensitive learning are usually used in order to consider the ordinal relationship between the ages in their objective functions. Secondly, it may not be easy to detect the aging effects on human faces by a single classifier or regressor [10]. Finally, it is often hard to gather a large aging database that contains chronometrical image series for the same individuals which implies that the number of images available for each age label can be very different. This would lead to a serious imbalance during the learning phase and to a degradation in the global performance of the estimation process.

AGE ESTIMATION ALGORITHMS

An age estimation algorithm is composed mainly of two processing stages: feature extraction and age determination. In the first stage, a compact representation of facial images is built by extracting facial features related to human ages or facial appearance change; in the second stage, the age corresponding to the input facial image is estimated by an age determination function based on the extracted features. We can eventually add other processing blocks such as input images preprocessing, distance metric adjustment and dimensionality reduction in order to solve some aspects related to the algorithm or to improve its performance (see Fig. (**1**)). In the following, we present a review of the previous work related to each of the processing stages.

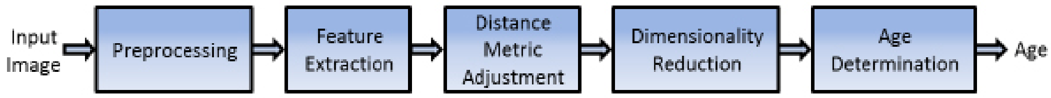


Fig. 1. The full flowchart of age estimation process (some of the blocks may not exist).

Input Images Preprocessing

In the context of automatic age estimation, the preprocessing step consists in preparing the input images and putting them in a common and homogeneous form. This process includes operations such as cropping, resizing, reshaping of the set of input images. The output of this step is generally composed of an ($m \times n \times p$) array corresponding to the preprocessed image ($p = 1$ for grayscale images and $p = 3$ for color ones). Depending on how the next processing methods are implemented, the preprocessed images can be reshaped as 1D vectors that contain the pixels and that can be gathered to form a huge matrix corresponding to the input data.

Feature Extraction

In the terminology related to image processing and computer vision domains, a feature is a particular information suitable for solving the computational task related to a given application. In general, features are intended to be informative and non redundant. They serve as a starting point for numerous computer vision algorithms and should facilitate the subsequent learning and generalization steps, in some cases leading to better human interpretations. In practice, features may consist of specific structures in the image such as points, edges or objects but may also be the result of an intermediate processing stage such as general neighborhood operation or feature detection applied to the input image [11]. During the feature extraction process, it is important to define a set of features, or

SUBJECT INDEX

2

2D shape landmarks 155, 156

3

3D Face matching 109, 126, 127
3D Face model 119, 154, 155, 159, 160, 166, 173, 174
3D Face recognition 106, 119, 120, 124, 125, 127-130
3D Face Registration 107, 114, 129
3D Rotation 86, 88, 166

A

Affinity matrix 27, 29, 30, 41, 66, 68, 69, 71, 72, 81, 82
Age classification 232
Age estimation 67, 84, 132, 152, 227, 228, 242, 246-249
Age Information 217, 220
Age prediction 132, 234
Aging pattern subspace 217, 218
Alex-Net architecture 3, 13
Appearance age 217, 219, 220, 222, 224

B

Backpropagation algorithm 3
Belief function 181, 188, 189, 194
Biological Age 217, 224, 225, 228, 229, 231
Biometric 86, 87, 109, 110, 128

C

Cautious rule 181, 197, 198, 200
CK-Regianini dataset 3, 11, 12, 14-17
CK-Zheng dataset 3, 13-16
Classification accuracy 3, 39, 42, 101, 102
CMU-Pittsburgh dataset 3, 13, 17
Combined features 3
Complete Kernel Fisher Discriminant method 39, 43, 55
Computer vision 20, 21, 37, 40, 63, 64, 67, 84, 129, 130, 132, 151, 152, 181, 182, 236, 238, 250
Conformal embedding analysis 64, 217, 218, 233
Conjunctive rule 181, 189, 190, 197, 199-201
Convolutional neural networks 3, 4, 21
Cumulative score 217, 227, 228, 231
Curvature descriptors 109, 120

D

Deep learning 3, 4
Deformable model backprojection 155, 165, 166
Dempster-Shafer 181, 185, 191, 203, 214, 215
Depth data 86, 89, 90, 93, 94, 99, 104
Dimensionality reduction 33, 36, 37, 62, 63, 66, 67, 83, 84, 86, 88, 99, 102, 107, 132, 133, 135, 136, 138, 143, 144, 146, 150, 240, 244-249
Distance diffusion mapping 39, 42, 48, 50, 53, 61
Distance metric learning 39, 42, 64, 107, 235

E

Eigenface 86, 102
Eigen problem 86, 100, 101
Eigenvalue problem 23, 30, 31, 70, 71
Ellipse fitting 86, 95
Exponential discriminant analysis 39, 43, 54

F

Face alignment 97, 109, 111, 122, 123, 127, 134, 178
Face analysis 109, 124
Face gesture analysis 155
Face model fitting 155, 165, 171, 176
Face recognition 21, 47, 54, 70, 73, 80, 91, 93, 116, 151, 152, 155, 158, 177, 178, 213, 233, 248
Face tracking 154, 155, 158, 159, 165, 167, 169, 178, 181, 183, 202, 209, 212
Facial actions 155, 158, 159, 172
Facial Age Estimation 84, 152, 217, 218, 223, 224, 227, 247
Facial appearance 229, 231, 234, 235, 237
Facial expression classification 3, 13
Facial expression recognition 21, 105, 155, 158
Facial feature extraction 111, 132, 155, 156, 177, 234
Facial features 86, 88, 91, 93, 103, 104, 109, 110, 121, 123, 134, 162, 165, 173, 177, 231, 235
Facial Landmarks 109, 122, 154, 218
Facial parts 155, 159, 170
Facial shape 159, 217, 218
Feature extraction 3, 5, 6, 10, 39, 40, 63, 64, 83, 84, 86, 88, 89, 99, 104, 106, 116, 119, 132, 135, 143, 146, 150, 155, 156, 177, 217, 218, 224, 231, 248
Feature generalization 3, 5, 10, 11, 13, 15-19
Feature representation 16, 127, 238
Fisherface 86, 130
Fusion of information 181, 185, 197
Fuzziness in facial aging 217, 219
Fuzzy Discriminant Analysis i, 217, 219
Fuzzy LDA 217, 229, 231
Fuzzy representation of age labels 217, 219, 220

G

Gabor filter 132, 234, 240, 246, 247
Generalized eigenvectors 39, 47, 49, 59, 70, 72
Gradient maps 128, 155, 158, 162
Graph-based linear embedding 66
Graph embedding 23, 24, 28, 30, 36, 38, 62, 63, 83, 84, 232, 248
Gray level image 86, 89, 104

H

Head pose estimation 41, 93, 94, 106, 107, 109, 126, 132, 133, 151, 152, 155, 233, 248, 250
Hidden layers 16, 19
Human Identification 109, 110, 127

I

Intra-class and Inter-class neighbors 217, 228
Intrinsic graph 39, 44
IRAD contours 86, 94
ISOMAP 23, 24, 40, 238

K

K-nearest neighbors 132, 136, 141, 234
Kernel Fisher discriminant analysis 39, 43

L

Label-sensitive 84, 132, 138, 140, 143, 152, 234, 238, 247
Label information 23, 28, 34, 36, 41, 43, 66, 68
Laplacian Eigenmaps 23, 24, 37, 40, 62, 66, 67, 84, 107, 238, 249
Laplacian matrix 23, 26, 46, 66, 69, 73, 137
Laser Scanner 109, 112
Latent points 66, 69
Learning-free 158, 159, 166, 174
Least mean square 86, 96
Levenberg- Marquardt 155

Linear discriminant analysis 23, 24, 63, 66, 67, 86, 88, 100, 125, 217
Local binary pattern 104, 105, 132, 234, 247
Local Discriminant Embedding 39, 42, 44, 45, 61, 62, 64, 107
Locality preserving i, 37, 39, 47, 54, 73, 132, 152, 234, 238, 247, 248
Locally linear embedding 23, 24, 37, 40, 62, 238, 249
Local regression 132, 136, 141, 234, 239
LUX color space 181, 193, 213

M

Machine learning 21, 43, 54, 73, 106, 132, 134, 150, 151, 234, 238, 247-249
Manifold learning 24, 40, 66, 67, 71, 86, 99, 102, 218, 232, 247, 248
Marginal fisher analysis 44, 132, 217, 218, 234, 238, 247
Mass set 181, 187, 194, 199, 206
Matrix exponential 43, 48, 60, 63
Mean absolute error 132, 143, 146, 217, 227, 234, 242
Mean curvature 86, 92, 93
Mean filter 86, 91, 92
Monocular image 152, 155, 160
Multidimensional scaling 23, 24, 37, 67, 83, 248

N

Nearest Neighbor classifier 33, 55, 66, 74, 83, 86, 101, 102
Neighborhood preserving embedding 23, 24, 27, 37, 41
Nose tip detection 106, 109, 120, 121, 130, 177

O

Online appearance model 155
Ordinary preserving LDA 217, 218
Ordinary preserving MFA 217, 218

P

Partial least square regression 132, 136, 141, 234, 239, 247
Particle filter 181, 183, 207, 209, 211, 212, 215
Pattern recognition 4, 20, 21, 40, 50, 54, 63, 64, 67, 73, 84, 106, 107, 129, 130, 151, 152, 181, 182, 188, 238, 249, 250
Penalty graph 39, 44
Pignistic probability 181, 191, 192, 199, 202, 209, 210
Pose estimation 41, 86, 91, 93, 94, 97, 98, 106, 107, 109, 126, 130, 132, 133, 135, 146, 151, 152, 155, 178, 206, 233, 248, 250
Pre-trained networks 3, 17
Preprocessing 74, 88, 90, 111, 119, 126, 132, 134, 135, 156, 177, 179, 234, 236
Principal component analysis 23, 24, 39, 40, 66, 67, 100, 124, 135, 237, 238, 247
Projection directions 66, 68, 70, 73, 100, 225
Projection matrix 23, 24, 26, 31, 41, 166, 221, 222, 224, 226
Projections 25, 31, 37, 39, 43, 49, 50, 54, 64, 83, 84, 132, 136, 137, 141, 152, 167, 225, 234, 238, 247, 248

R

Range Images 106, 107, 109, 110, 129, 130, 177
Recognition rate 23, 66, 74, 82, 102, 103, 132, 150, 234, 242, 244-246
Region of interest 12, 159, 183, 210
Regularization 3, 8, 13, 32, 40, 59, 60, 62, 140
Regularized kernel discriminant analysis 40, 55

S

Semi-supervised learning 23, 36, 62, 182, 196
Shape variations 155, 158
Sigmoidal filter 155
Singular matrix 40, 59
Skin hue 192, 194, 196, 198, 199, 202
Small sample size problem 40, 47, 49, 55, 61, 64
Source of information 189
Sparse representation 23, 24, 28, 30, 36, 38, 63, 65, 128, 132, 139, 140, 143, 150, 153
Sparsity preserving projection 23, 24, 28, 36
Stochastic gradient descent 3, 8
Supervised feature learning 3, 10
Support vector machine 3, 11, 121
Support vector regression 132, 136, 151, 227, 234, 239, 247, 249, 250

T

Trainable parameters 3, 5
Transferable belief model 181, 182, 191, 213
Translational invariance property 3, 14

U

Uncertainty management 181, 192

V

Viola-Jones detector 181
Visual servoing 181, 183, 199, 207, 208

W

Weak perspective 155, 165, 174