

2D/3D FACE RECOGNITION

by

Xin Guan

A Dissertation Submitted to the Faculty of

College of Engineering & Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Florida Atlantic University

Boca Raton, FL

May 2012

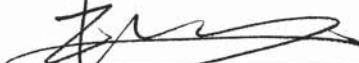
2D/3D FACE RECOGNITION

by

Xin Guan

This dissertation was prepared under the direction of the candidate's dissertation advisor, Dr. Hanqi Zhuang, Department of Computer & Electrical Engineering and Computer Science, and has been approved by the members of his supervisory committee. It was submitted to the faculty of College of Engineering & Computer Science and was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

SUPERVISORY COMMITTEE:

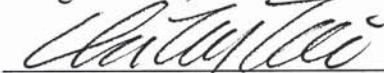


Hanqi Zhuang, Ph.D.

Dissertation Advisor



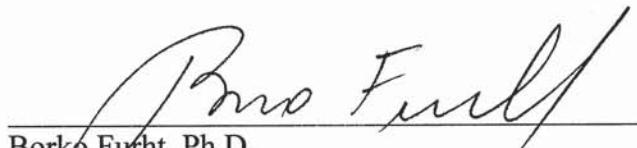
Nurgun Erdol, Ph.D.



Chi-Tay Tsai, Ph.D.

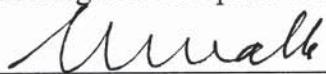


Xingquan Zhu, Ph.D.



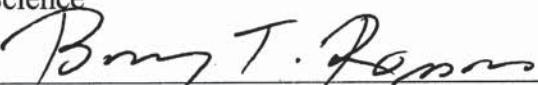
Borko Furht, Ph.D.

Chair, Department of Computer & Electrical
Engineering and Computer Science



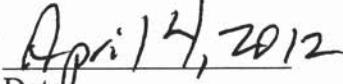
Mohammad Ilyas, Ph.D.

Interim Dean, College of Engineering & Computer
Science



Barry T. Rosson, Ph.D.

Dean, Graduate College



Date

ACKNOWLEDGEMENTS

The author wishes to express his sincere thanks and love to his wife, children, and parents for their support and encouragement throughout the writing of this manuscript. The author is grateful to his advisor, Dr. Hanqi Zhuang for providing knowledgeable advice and consistent support to the research. The efforts of the support face capture subjects and computer resources provided by Florida Atlantic University also are greatly appreciated.

ABSTRACT

Author: Xin Guan
Title: 2D/3D Face Recognition
Institution: Florida Atlantic University
Dissertation Advisor: Dr. Hanqi Zhuang
Degree: Doctor of Philosophy
Year: 2012

This dissertation introduces our work on face recognition using a novel approach based on creating 3D face model from 2D face images. Together with the pose angle estimation and illumination compensation, this method can be used successfully to recognize 2D faces with 3D recognition algorithms. The results reported here were obtained partially with our own face image database, which had 2D and 3D face images of 50 subjects, with 9 different pose angles. It is shown that by applying even the simple PCA algorithm, this new approach can yield successful recognition rates using 2D probing images and 3D gallery images.

The insight gained from the 2D/3D face recognition study was also extended to the case of involving 2D probing and 2D gallery images, which offers a more flexible approach since it is much easier and practical to acquire 2D photos for recognition. To test the effectiveness of the proposed approach, the public AT&T face database, which

had 2D only face photos of 40 subjects, with 10 different images each, was utilized in the experimental study. The results from this investigation show that with our approach, the 3D recognition algorithm can be successfully applied to 2D only images.

The performance of the proposed approach was further compared with some of the existing face recognition techniques. Studies on imperfect conditions such as domain and pose/illumination variations were also carried out. Additionally, the performance of the algorithms on noisy photos was evaluated. Pros and cons of the proposed face recognition technique along with suggestions for future studies are also given in the dissertation.

DEDICATION

This manuscript is dedicated to my understanding and supportive wife, Xinai Zhu, who has taken a lot of household burdens so that I can focus on my research, and to my lovely children, Nathan and Isabella, who are the little angels in my life and the resource of my happiness. I also dedicate this work to my parents, Zuoyao Guan and Xiaoyuan Zhou, both of whom always encouraged me to pursue my dreams.

2D/3D FACE RECOGNITION

List of Tables	xi
List of Figures	xii
1. Introduction.....	1
1.1. Motivation, Objectives and Background.....	1
1.2. Solution Strategies.....	4
1.3. Contributions	5
1.4. Structure of the Dissertation.....	7
2. Review of the Related Work.....	9
2.1. Introduction.....	9
2.2. 2D Face Recognition Methods.....	10
2.2.1. Geometric, Feature-based Approach	11
2.2.2. Template Matching	12
2.2.3. Hybrid Template Methods	15
2.3. 3D Face Recognition.....	16
2.4. 3D Face Models	21
2.4.1. Simple Models	21
2.4.2. Biomechanical Models.....	21
2.4.3. Morphable Models	21
2.5. Challenges to Face Recognition.....	22

2.6. Types of 3D/2D Face Recognition	23
2.6.1. Using 3D Shape Only	23
2.6.2. 3D Shape Assisted 2D Recognition.....	23
2.6.3. 3D Shape and Intensity	24
2.7. Summary and Conclusions	24
3. Preliminaries	27
3.1. Introduction.....	27
3.2. Preprocessing Techniques.....	29
3.2.1. Parameterized Model	29
3.2.2. Masking.....	32
3.2.3. 3D Rotation.....	32
3.2.4. 3D Data Fitting	32
3.2.5. Depth Parameter Extraction.....	34
3.2.6. Intensity Parameter Extraction.....	36
3.3. 3D Face Recreation Method	37
3.4. Initial Pose Determination	38
3.5. Illumination Compensation.....	45
3.6. Principal Component Analysis	48
3.7. Linear Discriminant Analysis	52
3.8. Optimization Method	55
3.8.1. Non-Linear Least Square Problem.....	55
3.8.2. Quasi-Newton Method.....	55
3.8.3. Levenberg-Marguardt Method.....	57

3.8.4. Hybrid Method.....	57
3.8.5. Discrete Approximation.....	62
3.9. The Summary of the Training Process.....	62
3.10. The Summary of the Recognition Process.....	65
3.11. Summary	67
4. The Proposed 2D/3D Recognition Approaches	68
4.1. Introduction.....	68
4.2. System Setup.....	69
4.3. 3D Shape Assisted 2D Recognition.....	70
4.3.1. Solution Method.....	70
4.3.2. Pre-Processing.....	73
4.3.3. Pose Acquisition	73
4.3.4. Illumination Compensation.....	74
4.3.5. Optimization Method	75
4.3.6. 2D PCA for Dimensional Reduction and Parameter Extraction.....	76
4.4. 2D Assisted 3D Recognition.....	79
4.4.1. Solution Strategy.....	79
4.4.2. Enrollment Stage.....	79
4.4.3. Verification Stage	80
4.4.4. 3D Face Recreation	82
4.5. Summary	84
5. Experiment Configurations	85
5.1. Introduction.....	85

5.2. Experiment Setup.....	85
5.3. Raw 3D Shape Data.....	88
5.4. Raw 2D Intensity Data.....	92
5.5. Model Fitting	94
5.6. Summary	97
6. 2D/3D Face Recognition Experiment Results	98
6.1. Introduction	98
6.2. 3D Shape Assisted 2D Recognition	98
6.2.1. Introduction.....	98
6.2.2. Experiment with Simple 2D PCA Face Recognition.....	99
6.2.3. Experiment with Pose Compensation	101
6.2.4. Experiment with Both Pose and Illumination Compensation.....	103
6.2.5. Summary	105
6.3. 2D Assisted 3D Recognition.....	105
6.3.1. Introduction.....	105
6.3.2. Using Single 2D Image.....	106
6.3.3. Using Multiple 2D Photos	107
6.3.4. Noisy Input Images	107
6.3.5. Experiment Using PCA for Feature Extraction	109
6.3.6. Experiment Using FLD for Feature Extraction	111
6.3.7. Summary	112
6.4. Summary	113
7. Enhanced Recognition Method Using 3D Model.....	114

7.1.	Introduction.....	114
7.2.	Solution Strategy.....	116
7.2.1.	Modified Enrollment Stage.....	116
7.2.2.	Verification Stage	119
7.3.	Experiment with Simple 2D PCA Face Recognition.....	119
7.4.	Experiment with 3D Modeling and Recognition	120
7.5.	Summary	123
8.	Conclusions.....	124
8.1.	Review of Methods.....	124
8.2.	Future Works	125
	Bibliography	127

TABLES

Table 6.1. 3D Shape Assisted 2D Recognition Rate Comparison.....	105
Table 6.2. 2D Assisted 3D Recognition Rate Comparison.....	113
Table 7.1. 3D Recognition Rate Comparison	123

FIGURES

Figure 3.1. Conceptual Illustration of Linear Combination of Faces	28
Figure 3.2. Cubic Representation of 3D Object.....	31
Figure 3.3. An Example of 3D Images	31
Figure 3.4. Triangulate Interpolation	34
Figure 3.5. The Yaw, Pitch, and Roll Angle.....	39
Figure 3.6. 3D Data Image of the Fronto-Parallel Pose.....	41
Figure 3.7. The 2D Data Image with Slightly Side View.....	42
Figure 3.8. Sample Image Rotation	45
Figure 3.9. 3D Face Shape with Phong Reflectance Model	46
Figure 4.1. A Snapshot of the 3D Camera System in the FAU DSP Lab.....	70
Figure 4.2. Face Recognition Diagram	72
Figure 4.3. Example of 2D Depth and Intensity Image	73
Figure 4.4. Sample Image Rotation	74
Figure 4.5. Illumination Compensation Example	75
Figure 4.6. Verification Stage	81
Figure 4.7. 3D Face Recreation	83
Figure 5.1. Database Construction Flowchart.....	87
Figure 5.2. Raw 3D Data	88
Figure 5.3. Original 3D Shape Data.....	89

Figure 5.4. Data after Transformation, Rotation, Scaling and Crop.....	89
Figure 5.5. The 3D Shape Image Re-generated by the NURBS.....	90
Figure 5.6. 3D Shape Image (amit-20041028-1.stl)	91
Figure 5.7. After the Pre-processing (amit-20041028-1-output.stl)	91
Figure 5.8. The Generated Point Clouds (temp.txt).....	92
Figure 5.9. Original High Definition Intensity Image	93
Figure 5.10. Original Grayscale 2D Intensity Image	93
Figure 5.11. Generated Grayscale Image.....	94
Figure 5.12. Average Face	95
Figure 5.13. Recreated 3D Faces from 2D Images	96
Figure 5.14. First Three Iterations of Recreated 3D Faces	97
Figure 6.1. Example of 2D Probing Image	100
Figure 6.2. Error Norms between Each of the Probe and Gallery Images.....	101
Figure 6.3. Example of the 2D Probing Image in Semi-profile View	102
Figure 6.4. Error Norms between Each of the Probe and Gallery Images.....	103
Figure 6.5. Illumination Compensation Example	104
Figure 6.6. Error Norms between Each of the Probe and Gallery Images.....	104
Figure 6.7. Recreated 3D Faces from 2D Images	106
Figure 6.8. Recreated 3D Faces from Different Number of Input Photos	107
Figure 6.9. Recreated 3D Faces from “Unfocused Images”	108
Figure 6.10. Recreated 3D Faces from Images under “Camera Motion”	109
Figure 6.11. Example of 3D PCA Test Result.....	110
Figure 6.12. Error Norms between Each of the Probe and Gallery Images.....	111

Figure 6.13. Error Norms between Each of the Probe and Gallery Images.....	112
Figure 7.1. Preview of AT&T Database of Faces.....	115
Figure 7.2. Enrollment Stage	118
Figure 7.3. Example of Probing and Gallery Images.....	119
Figure 7.4. Error Norms between Each of the Probe and Gallery Images.....	120
Figure 7.5. Recreated 3D Faces from 2D Images	121
Figure 7.6. Error Norms between Each of the Probe and Gallery Images.....	122

Chapter 1

INTRODUCTION

1.1 Motivation, Objectives and Background

Face recognition plays a very important role in human activities. The ability to recognize other people is the firm basic of the way human beings interact with each other. Humans are able to identify distorted images [1], coarsely quantized images, faces with occluded detail, and even inverted face images [2].

With the advances in automated technology, researchers developed numerous applications for a robust automated recognition system. Biometric systems are automated methods for identifying people through physiological or behavioral characteristics. A face recognition system provides a least obtrusive method to identify users, which only needs a user simply walking past a surveillance camera. There are many other biometric methods, which are briefly discussed below.

Fingerprint is a well-known method of biometric measure. FBI and other institutions around the world have carried out research in this field. It is relatively cheap and unobtrusive to buy a fingerprint system. Fowler [3] has produced a short summary of the available systems.

Iris of the eye is also used as biometric measure recently due to its uniqueness to each human being. Like fingerprints, it displays patterns and textures unique to each

person and it remains stable over decades of life, details of which are described by Siedlarz [4].

One of the most natural and less obtrusive biometric measures is speech recognition. AT&T have produced a prototype to identify a user through his or her spoken words, as detailed by Mandelbaum [5].

In this dissertation, we focus mainly on issues related to 3-Dimensional (3D) face recreation and 2-Dimensional (2D)/3D face recognition. Face recognition has been an interesting and difficult problem. Each face has special features that define a particular individual, and meanwhile they can also be very similar in a particular feature space. It is not only technically but also economically feasible to make automatic biometric authentication based on face images.

Over the past 30 years, the problem of 2D face recognition has received significant attention from researchers in the areas of pattern recognition, computer vision and biometrics. Early methods focused on 2D face recognition, and the classical 2D face recognition approaches mainly include the following 6 algorithms:

1. Eigenfaces [6]
2. Fisherfaces or Linear Discriminant Analysis (LDA) [7]
3. Independent Component Analysis (ICA) [8]
4. Support Vector Machine (SVM) [9]
5. Neural network [10]
6. Hidden Markov Model (HMM) [11]

The reports of major evaluation studies of 2D face recognition technology can be found at Face Recognition Vendor Test [12] and Face Verification Competition [13]. A major problem associated with any 2D approach is that the 2D appearance of a face changes significantly when either pose or illumination condition varies. In a 2D face recognition system, even only the illumination condition changes or only the pose changes, the performance of state-of-the-art systems can be greatly decreased [14]. Therefore for robust face recognition, it is necessary to utilize a 3D face model to compensate for the variation of pose and illumination differences prior to feature extraction.

In order to effectively conducting a face recognition task, processing of face images is essential. Early work included developing algorithms to match a small number of feature vertices to image positions, to interpolate deformations of the surface in between [15], using deformation models [16], and to analyze images with shape-from-shading [17]. An excellent deformable 3D model was presented in [18], which was combined with a computer graphics simulation of projection and illumination. Given a single 2D image of a person, the algorithm proposed in [18] was able to estimate 3D depth, intensity, and other relevant 3D scene parameters. To cope with pose variations in 2D face recognition, Hager and Belhumeur devised a method to deform the target image into the fronto-parallel pose using a 2D motion model [19]. In this method, projection of the target image into a standard image space would be problematic due to large pose variation. To describe an illumination condition with only a small number of parameters is difficult because of the infinite degrees of freedom in types or amount of light sources.

However, previous works showed that face images obtained from various illumination conditions can be sufficiently approximated by a low-dimensional linear subspace [20, 21]. The problem of face recognition under varying pose and illumination conditions was addressed in [22]. Robustness to appearance variations is achieved mainly by using face geometry information to cope with pose and illumination variations. A face recognition system that is invariant to both viewing directions and facial expressions was proposed in [23]. Landmarks are used to relate 2D features with 3D features of the face. To classify test faces under varying views or varying facial expressions, a Structural Hausdorff Distance is proposed to deal with the case of matching incomplete data under some structural constraints.

Nowadays it's not too difficult to obtain 3D face images in a relatively small quantity. However, 3D face enrollment and verification could become problematic as the method could be time consuming, computationally expensive and impractical in implementation.

Our research seeks a new method for face recreation in 3D by using 2D probing images, with the aid of using our own 3D face database, and then performs the face recognition either in 2D or 3D. These 2D *probing images* (image to be recognized) can be taken from any direction and under any illumination, and the *gallery images* (image enrolled in the system) can be either 2D or 3D. Our own 3D face database is taken by Genex 3D facecam system and includes both the intensity and the depth information of that human face. And we use the AT&T 2D face database in one of our research project as well.

1.2 Solution Strategies

As has been mentioned, we are interested in recreating 3D face images from 2D images with the aid of a 3D database. Then the face recognition is done either in the 2D plane or 3D space. For 2D face recognition, pose and illumination compensation is first applied to the 3D face and then it is projected to the 2D plane. For 3D recognition, only illumination compensation is needed, and both the depth and the intensity information are used in recognition. The following are specific strategies adopted in this research.

1. 3D face images are preprocessed through filtering, rotation, scaling, translation and interpolation.
2. A feature extraction technique such as the Principle Component Analysis (PCA) is applied to extract the depth and intensity features of 3D images, and these features are combined for the creation of 3D faces of the subject under investigation.
3. Features extracted from the 3D face images are utilized to span the feature space. 3D feature vectors of 2D probing photos are estimated in the 3D feature space, utilizing the 3D model feature vectors. These estimated 3D feature vectors are then used to create the 3D face image of the subject, and later are also used to perform 3D face recognition.

1.3 Contributions

Face recognition poses a challenging problem for many researchers around the world. Due to the change of lighting, pose angle, appearance and background, it may be very difficult to match similar images using computers although they look alike to

humans. Many researchers in the computer vision literature have focused on improving the basic methods of either geometric features or template matching. In this dissertation, a novel approach of using 3D feature vector to recreate a 3D face based on 2D face images, and then performing the face recognition in either the 2D plane or 3D space is proposed. To our best knowledge, the face recognition framework using the above method hasn't been found in others work and this dissertation is believed to be original.

The detailed original works done in this dissertation are summarized as follows:

- Feature vectors of both depth and intensity of subjects under consideration is extracted from 3D face data with the Principle Component Analysis (PCA).
- After the 3D face data is transformed to a point in the feature space spanned by the depth and intensity eigenvectors, the pose of the subject and illumination condition under which the image is taken are estimated and the information is subsequently used to compensate subject's pose and illumination condition.
- When a single 2D image of a person, whose images are not recorded *a priori*, is presented to the system, the face recognition system recreates the 3D face of that person using a non-linear parameter estimation technique with the aid of a small 3D model face database.
- Recognition of 2D faces is performed by first rotating each 3D gallery image to the same pose angle as the 2D probing image, and then compensating it with the same illumination condition. The rotated and light-compensated 3D

image is then projected onto the 2D plane. Recognition can be done through classification of the estimated 2D intensity feature vector.

- Face recognition in the 3D space is performed by recreating a 3D face whose 2D projection is closely matched with the probing 2D photo. Recognition can then be done through classification of the estimated 3D depth and intensity feature vectors.

1.4 Structure of the Dissertation

This dissertation consists of the following eight chapters.

Chapter 2 surveys the relevant literature in the field of face recognition. A brief account of the work completed so far is presented first, followed by a more comprehensive description of the recent advances in the field of automated face recognition.

Chapter 3 provides processing techniques. The detailed procedure used in this research will be explained, including taking 3D picture, storing 3D images into the database, the PCA method, extracting 3D depth and intensity information from 3D models based on PCA method, pose estimation algorithm, illumination compensation algorithm, and the optimization algorithm to create 3D face from 2D image based on the 3D depth and intensity vectors from the 3D database.

Chapter 4 proposes two different approaches, under the scenario of using 2D probing image and 3D gallery images. The first approach is to project the 3D gallery images to the 2D plane and the second one is to create 3D face model from 2D probing

images. Different applications for each methods are briefly reviewed. The proposed computing algorithms and the motivations behind the algorithms are then described.

Chapter 5 details the setup of our experiments, including the specification of a 3D face camera, the structure of a 3D face database, the format and usage of image files, as well as the explanation of our experiment procedure.

Chapter 6 presents experimental results using the above algorithms. Two sets of experiments are investigated. In the first experiment, each 3D gallery image is projected to the 2D plane and the recognition is performed in the 2D plane as well. In the second experiment, a 3D face model of a 2D probing image is generated and the recognition is performed in the 3D space.

Chapter 7 extends the insights gained from the above experiments. And a new method to handle 2D gallery images and 2D probing images is proposed, in which 3D face models are created for both gallery and probing images, and the recognition is performed in the 3D space.

Chapter 8 concludes the dissertation, by summarizing the pros and cons of our research and indicating the directions for the future work.

Chapter 2

REVIEW OF THE RELATED WORK

2.1 Introduction

The task of recognizing faces has attracted the attention of a great number of researchers. Several factors are the main motivation for these scientific interest and attention: [24]

1. The inherent challenge for the problem of face image processing and recognition poses.
2. The immense commercial significance that robust and reliable face recognition technology would entail
3. The widely spread application that are envisaged in physical and logical access control, security, man-machine interfaces and low bit rate communication.

The problem of automatic face recognition involves three key steps: face detection and rough normalization, feature extraction and accurate normalization, and face identification or verification.

Sometimes the task of face detection and feature extraction can be performed simultaneously. And sometimes the task can be very challenging depending on the nature

of the application, such as the face database size, facial expression, occlusion, background clutter and variability, noise, and computer speed.

In this chapter we survey the state of the art of face recognition in the engineering literature, specifically targeting those relevant to our approach. We start with 2D face recognition methods, and then we discuss about methods on 3D face recognition. We will see that significant progress has been achieved on various aspects of face recognition, and researchers are continually constructing fully automated systems that integrate all the techniques for various applications.

2.2 2D Face Recognition Methods

Face recognition has always been a very challenging problem, which, for many years, has attracted researchers from different backgrounds: psychology, pattern recognition, neural networks, computer vision, and computer graphics [25]. A great number of techniques for face recognition have been proposed and three general strategies for solving the problem have been identified in the literature, as pointed out by Brunelli and Poggio [26] and Robertson and Craw [27]:

1. *Geometric, Feature-based Approach*

Local features such as the eyes, nose, and mouth are first extracted in these methods. Then their locations and local statistics are fed into a structural classifier.

2. *Template Matching*

The whole face region is used in these methods as the raw input to a recognition system. The image is represented as single or multiple arrays of pixel values, and

then compared with templates in the training sets. The features of interest can be either manually picked or using some automatic approaches.

3. Hybrid Method

This is a combination of the above two methods.

2.2.1 Geometric, Feature-based Approach

2.2.1.1 Pure Geometry Approach

Many earlier methods in face recognition belong to this category. Kelly [28] proposed a method based on the width of the head, the distances between the eyes and from the eyes to the mouth, etc. Kanade [29] used the distances and angles between eye corners, mouth extrema, nostrils, and chin top. More recently, Cox et al. [30] employed a mixture-distance based approach by manually extracted distances.

2.2.1.2 Dynamic Link Architecture (DLA)

The graph matching system is based on the Dynamic Link Architecture (DLA) [31, 32]. One of the most successful examples of such systems is the Elastic Bunch Graph Matching (EBGM) system [33, 34]. Gabor wavelets are used as a building block for facial representation in the approach. A typical local feature representation consists of wavelet coefficients for different scales and rotations based on fixed wavelet. These locally estimated wavelet coefficients are robust to illumination change, translation, distortion, rotation, and scaling of facial images.

Instead of attaching only a single jet (wavelet coefficient) to each node in DLA, Elastic Bunch Graph Matching attaches a set of jets called the *bunch graph*

representation, which derived from a different face image each. The pose of the face is determined by using prior class information to handle pose variation [35], and jets transformations under pose variation are learned [36]. The EBGM approach has been widely applied to face detection and extraction, pose estimation, gender classification, sketch-image-based recognition, and general object recognition. The success of the EBGM system may be due to its resemblance to the human visual system [37].

2.2.1.3 Hidden Markov Model (HMM)

Without finding the exact locations of facial features, HMM based methods use strips of pixels that cover the forehead, eye, nose, mouth, and chin of a face [38, 39, 40]. Nefian and Hayes's approach showed better performance than Samaria's by using the KL projection coefficients instead of the strips of raw pixels.

2.2.1.4 Convolution Neural Network

Lawrence et al. [41] developed a system based on a convolutional neural network (CNN), using an unsupervised learning method based on a self-organizing map (SOM). The SOM provided a quantization of the image samples into a topological space for dimensionality reduction and invariance to minor changes in the image sample. The convolutional neural network provided partial invariance to translation, rotation, scale, and deformation, and it extracted successively larger features in a hierarchical set of layers. They used a database of 400 images of 40 individuals which contains quite a high degree of variability in expression, pose, and facial details, and reported 3.8% error rate.

2.2.2 Template Matching

2.2.2.1 Principal Component Analysis (PCA) and its Extensions

Due to its successfulness of low-dimensional reconstruction, it has been one of the major algorithms used in face recognition. The detailed PCA algorithm will be described in the next chapter, as it is also used in our research. Listed below are several applications of PCA.

2.2.2.2 Eigenfaces

Turk and Pentlad [6] made the first successful demonstration of automatic face recognition using eigenfaces. They used a nearest neighbor classifier to identify each probe image by representing it as a vector of weights. They demonstrated the method using a database of 2500 face image of 16 subjects. The standard eigenfaces approach was extended to a Bayesian methods later [42], which used a probabilistic distance metric instead of the simple Euclidean distance.

2.2.2.3 Fisherfaces/LDA

This is also a very successful approach, using linear/Fisher discriminant analysis (FLD/LDA) [7, 43, 44, 45], in which LDA training was carried out via scatter matrix analysis.

2.2.2.4 SVM Methods

Philips [46] developed this method using a support vector machine as the classifier.

2.2.2.5 Evolution Pursuit (EP) Method

It is a derivative of the popular PCA representation [47], and it is also based on artificial neural networks. Similar to projection pursuit methods, EP seeks to learn the optimal basis for both data compression and pattern classification.

2.2.2.6 Feature-line-based Methods

Li and Lu [48] developed this method by replacing the point-to-point distance with the distance between a point and the feature line linking two stored sample points.

2.2.2.7 Independent Component Analysis (ICA)

Based on the argument that much of the important face information is contained in high-order statistics, Bartlett et al. [49] proposed a method to use ICA to extract features. ICA is a generalization of PCA, which decorrelates the high-order moments of the data in addition to the second-order moments. It is widely argued in the literature that it is capable of providing a better recognition performance than PCA, therefore it is also applied in this research and the approach will be detailed in the next chapter.

2.2.2.8 Other Representations

Neural network methods have also been applied to face recognition to provide greater generalization through learning. Lin et al. [50] proposed a fully automatic face detection/recognition system based on Probabilistic Decision-Based Neural Network (PDBNN) method which consisted of three modules: a face detector, an eye localizer, and a face recognizer.

Etemad and Chellapa [51] used a pure LDA/FLD method on raw images using weighted metrics.

2.2.3 Hybrid Template Methods

2.2.3.1 Modular Eigenfaces Method

The concept of eigenfaces was extended to eigenfeatures, such as eigeneyes and eigenmouth. In [52], the use of hybrid features by combining eigenfaces and eigenmodules was explored; and recognition performance as a function of the number of eigenvectors was measured for eigenfaces only and for the combined representation. It was also reported that the eigenfeatures performed better than the eigenfaces for lower-order spaces. And only marginal improvement was obtained when the combined set was used.

2.2.3.2 Hybrid LFA

This is a hybrid representation based on PCA and local feature analysis (LFA) [53]. Its biological motivation comes from the fact that only a small fraction of receptors are active in the human retina. Researchers expected that the natural objects can be represented in a subspace of lower dimensionality by finding a suitable parameterization.

2.2.3.3 Shape-Normalized Method

Lanitis and Cootes [54] presented the flexible appearance model based method for automatic face recognition. To identify a face, both shape and gray-level information were modeled and used. The shape model was an ASM (Active Shape Model). Based on

the average shape of the shape model, a global shape-free gray level model was constructed using PCA. Local gray-level models were also built on the shape model points. For an input image, all three types of information were used to compute a Mahalanobis distance for classification.

2.3 3D Face Recognition

The general consensus for 2D face recognition technology is that it performs quite well in controlled conditions, where the subject is presented in a fronto-parallel pose, under good illumination, and with the same condition as acquiring the training (gallery) images. These list of problems motivated a radically change of approaches for face recognition by using 3D based properties of the face. Due to its robustness to pose and illumination variations, in recent years, 3D face recognition has emerged as a major research trend [55, 56, 57]. With the development of 3D acquisition system, a number of 3D approaches are proposed in the literature [58], ranging from subspace-based solutions which applies classical 2D methods for 3D face recognition by considering for the depth data [59, 60], to geometry feature-based techniques which represents facial surfaces by certain geometrical features set [61].

Earlier research in this field focused mainly on curvature analysis. Cartoux et al. [62] proposed an approach by segmenting a range image based on principal curvature and finding a plane of bilateral symmetry through the face. They used this plane to normalize for pose, and matched the profile from the plane of symmetry and matched the face surface. Their result showed 100% recognition rate for a small dataset.

Lee and Milios [63] segmented convex regions in the range image based on the sign of the mean and Gaussian curvatures, and created an Extended Gaussian Image (EGI) for each convex region. A match between a region in a *probing* image (image to be recognized) and in a *gallery* image (image enrolled in the system) was done by correlating EGIs. A graph matching algorithm was used to find the match of probe image to gallery image. This approach is robust with changes in facial expression since convex regions change shape less than other regions in response to changes in facial expression. However, EGIs are not sensitive to change in object size, hence it cannot distinguish the representation of two faces with similar shape but different size.

Gordon et al. adopted Gaussian and mean curvatures to characterize delicate features in 3D faces [64]. He used the nose region and ridge and valley lines from the segmentation to register the image to a standard pose. Then he computed the volume difference between registered probe and gallery surfaces. This approach can handle faces that are similar in shape but different in size. However, it requires some extension to deal with changes in facial expression.

Nagamine et al. [65] used an approach to find five feature points, and then those feature points were used to standardize face pose and to match the various curves through the face data. Experiments were performed for sixteen subjects, with ten images per subject. The best performance was reported by using vertical profile curves that pass through the central portion of the face.

The eigenface and hidden Markov model approaches were extended by applying 2D recognition to work with range images [65]. In the experiments reported in [65], the

authors used a dataset of 24 persons, with 10 images per person, and reported 100% recognition using an adaptation of the 2D eigenface algorithms.

Tanaka et al. [67, 68] treated 3D face recognition as a 3D shape recognition problem of free-form curved surfaces. Each face in both probe image and gallery database was represented as an Extended Gaussian Image (EGI), and recognition was achieved by evaluating the spherical correlation of EGIs. Their methods were based on curvature analysis and they made good use of 3D properties of the data. However, the algorithm requires a relatively high computational cost.

Hesher et al. [69] extended the PCA approach by using different numbers of eigenvectors and image sizes. Their experiment used 37 subjects and each has 6 different facial expressions. This method gives the probe image more chances to make a correct match.

Medioni et al. [70] performed 3D face recognition using iterative closest point (ICP) matching. This work used a stereo-based system, and reported better than 2% Equal Error Rate (EER).

Moreno et al. [71] proposed an approach of first performing a segmentation based on Gaussian curvature and then creating a feature vector based on the segmented regions. They used a dataset of 420 face meshes representing 60 different persons, with different expressions and poses for each person. They reported 78% rank-one recognition on the subset of fronto-parallel pose, and 93% overall rank-five recognition.

Lee et al. performed 3D face recognition by locating the nose tip, and then forming a feature vector based on contours along the face at a sequence of depth values

[72]. They used a small dataset of 35 persons and reported 94% correct recognition at rank five.

Both 3D and 2D images were used by Chang et al. [73]. They used a 3D image as the gallery image of a person; weighted sums of the three normal components of the 3D image were used to synthetic 2D images under different illumination conditions. Correlation was performed with a 2D probing image to achieve illumination compensation.

Lao et al. [74] presented an approach to use a sparse depth map constructed from stereo images. Iso-luminance contours were used for the stereo matching. Both 2D edges and iso-luminance contours were used in finding the irises. Using the iris locations, other feature points were found so that pose standardization can be done. A dataset of 10 persons with 4 images taken at 9 poses each were used and the recognition rates of 87% to 96% were reported.

Beumier and Achery [75] used a weighted sum of the 3D and 2D similarity measures to perform multi-modal recognition. Their experiments used a dataset of 26 persons in the gallery and 29 persons in the probe set, achieving recognition performance as high as 2% equal error rate (EER) for multi-modal recognition, compared to 4% for 3D alone and 8% for 2D alone.

Wang et al. [76] approached multi-modal face recognition by using Gabor filter responses in 2D and “point signatures” in 3D. The 2D and 3D features together formed a feature vector. Classification was done by support vector machines with a decision directed acyclic graph. Experiments were done using the dataset with 50 subjects of 6

images from each, with pose and expression variations. They reported over 90% recognition rates.

Bronstein et al used an isometric transformation approach to better handle the variation in face shape due to facial expression change [77]. They proposed a method for multi-modal 2D+3D recognition using eigen decomposition of flattened intensities and canonical images.

Recently due to the excellent performance of principal component analysis (PCA) in 3D face recognition [60], many appearance based methods have been adopted in this field. Mian et al. [78] proposed a fusion system to handle the expression problem. In their system, three kinds of methods, spherical face representation, SIFT based matching and a modified ICP were combined to achieve the final recognition. Their results showed the potential of appearance based methods to solve the expression problem in 3D face recognition. However, their experiments were under the controlled environment since they chose neutral faces as registered templates.

Chang et al. [79] performed a PCA-based method using 3D and 2D images from 200 persons. One experiment used a larger set of 676 probes taken in multiple acquisitions over a longer elapsed time, and another experiment used a single set of later images for each person as the probes. Results in both experiments were approximately 99% rank-one recognition for multi-modal 3D+2D, 94% for 3D alone and 89% for 2D alone. The combined result was obtained using a weighted sum of the distances from the individual 3D and 2D face spaces.

Romdhani et al. proposed to use a 3D morphable model for face recognition on 2D images [80]. With this method they tested on a dataset of 68 subjects, and their recognition rate reached 99.9% for neutral fronto-parallel pose images and 89% for profile images. Huang et al. added a component based approach to the morphable model [81] based on the approach of Heisele [82]. However, the recognition rate reported for all approaches of the morphable model was between 75% and 99%.

2.4 3D Face Models

Raw measurement data is captured from 3D acquisition system, and generally this raw data will be further pre-processed and the intermediate structured face model will be created:

2.4.1 Simple Models

Everingham and Zisserman [83] used both the coarse 3D geometry and multiple intensity maps to train the generative parts-based constellation models. The 3-D ellipsoid approximation of the person's head is created to generate candidate hypotheses in the image. Then the detected parts are then used to align the model across a wide range of pose and appearance.

2.4.2 Biomechanical Models

Approximating the structure and musculature of the human face has been widely used in computer animation [84], especially for simulation of facial movement during expression and speech. Using this method, 3D measurement of a person's face shape and

color appearance [85] are used to develop the model. Metaxas et al. [86] used anthropometric statistics of face shape to synthesize biomechanical models with natural variation in facial characteristics. This kind of model provides a basis for model-based 3D or 2D face recognition.

2.4.3 Morphable Models

Blanz and Vetter [18] estimated the 3D shape and intensity based on a single image, by fitting a statistical, morphable model of faces to the image. The 3D intensity data is used to create the model, using a laser-stripe scanner. To represent the statistical variation of shape and intensity of human heads, a probabilistic PCA-based model is used, and to handle the illumination variations, the Phong model [87] is applied. Then to fit the model to particular image data, the stochastic optimization of a multivariate cost function is used. To achieve an accurate fit of facial features, it first performed a holistically fitting and then eyes, mouth and nose independently. In the method of Yin and Yourst [88] shape is reconstructed by adapting a deformable elastic model to hyperresolution-enhanced input fronto-parallel pose and profile images. Similar approaches include the work of Grammalidis et al. [89] and Ansari and Abdel- Mottaleb [90].

2.5 Challenges to Face Recognition

Although recent advances in automatic face recognition have achieved a certain level of success, the performance of even a state-of-art system is still far from that of a human. Here are some aspects that need improvement:

- Develop a hybrid face recognition system using both holistic and local features to resemble the human perceptual system.
- Develop face detection/recognition techniques under large pose and illumination variations.
- Model face variation under aging or an outdoor environment.
- Enhance the robustness of the system so that it is less sensible to size variation and facial expression.

2.6 Types of 3D/2D Face Recognition

2.6.1 Using 3D Shape Only

Lee and Milius [63] did one of the first work to investigate the use of range images in face recognition. The most common method is to utilize the Correlation between Gaussian images of convex regions on the face image. Gordon [64] used another approach to investigate the 3D face recognition. This is a Curvature-based method in which descriptors are computed over various regions of the face and used as features. Tanaka et al. [68] presented another curvature-based method. Chua et al. [91]. used point signatures in their work, in which the algorithm can deal with the human facial expressions. Xu [92] presented a more recent approach of shape based recognition. A fishersurface based method with a combination of a variety of facial surface representations was proposed by Heseltine et al. [93]. Other methods and detailed performance description can be found in the survey of Bowyer et al. [60].

2.6.2 3D Shape Assisted 2D Recognition

As discussed in Chapter 1, Pose and illumination were two major problems in 2D face recognition. Due to the intrinsic 3D nature of this problem, approaches that trying to purely do 2D face recognition are bound to have limited performance.

An algorithm which takes a single image as input and then reconstructs 3D shape and illumination-free intensity was proposed by Blanz and Vetter [18]. They used Phong's illumination model to capture the illumination variance. The model allowed invariant description of the identity of faces by separating imaging parameters from personal parameters. The set of intensity and shape parameters which yielded the best fit are used as features. And the experimental results were evaluated by using the FERET and the CMUPIE databases.

A set of images of a convex Lambertian object obtained under arbitrary illumination is proved to be accurately approximated by a 9D linear space by Basri and Jacobs [94]. The Blanz and Vetter morphable model together with a spherical harmonic representation was used by Zhang and Samaras [95] for 2D recognition. Even when multiple illuminants are present this method is reported to perform well.

Yin and Yourst [88] used 2D data only for their 3D face recognition system. A dynamic mesh algorithm is applied to reconstruct 3D shape from fronto-parallel pose and profile images of the person. In their method, a curvature-based descriptor is first computed for each vertex of the mesh, then shape and intensity features are used for matching.

2.6.3 3D Shape and Intensity

It is common sense that use all available information will generate better recognition result in the decision making process. Generally, both intensity and shape information is fused at the feature level or the decision level. Lu [96] proposed a method to create robust similarity metric combining intensity and shape features. Morphing 2D face intensity onto a canonical shape computed from the range data is proposed by Bronstein et al. [77]. A detailed recognition comparison of multimodal approaches fusing 2D and 3D can be found in [58, 60].

2.7 Summary and Conclusions

In this chapter we have presented an extensive survey of the past research activities and results on face recognition. We have considered both the 2D and 3D recognition methods. We have categorized the methods for each type, and discussed their characteristics and some experimental results, together with the pros and cons. In addition, we have provided a summary of recent developments and challenging issues. We have also addressed the problems associated with illumination and pose variations in practical face recognition systems.

Due to the fact that most experiments are performed on different datasets, it is difficult to compare the results of various methods. One method with low error rate tested on noisy images with varying head poses is not necessarily better than another method with high recognition rate tested on neutral fronto-parallel pose images.

Some researchers combined two or more techniques and hence obtained a better performance than the individual techniques. Besides recognition rate, error rate and computational cost are also important. If the error rate of a method decreases significantly, even its recognition rate increases only slightly, it is still a preferred method. However, if the computational cost increases significantly, the method may not be suitable for real-time applications.

To conclude this chapter, one can see that the pioneer 3D face recognition approaches were usually tested with relatively small datasets. However, sizes of datasets are increasing recently due to the availability of better acquisition equipment. Most 3D face recognition methods treat the human face as a rigid object. Such method may not be capable of handling facial expressions. In the recent years, performances of many 2D face recognition algorithms have been improved in handling less perfect images, such as noisy images, half profile images, occlusion images, and images with different facial expressions. Currently a well-designed 2D face recognition method may still outperform a typical 3D face recognition method. However in the near future, we can expect that 3D face recognition methods will outperform their 2D counterparts. This is because 3D images hold more information of the face, such as the surface and depth data, which can be used for face recognition. Furthermore, 3D face recognition can be pose invariant. Therefore, 3D face recognition is still a challenging but very promising research area.

Chapter 3

PRELIMINARIES

3.1 Introduction

In recent years, research in the 3D face recognition field has focused on feature-based and template-based methods, as described in Chapter 2. Although it is a complicated task, it has distinguished advantages over 2D face recognition.

A novel approach of creating 3D face model from 2D photo is investigated in this dissertation. In order to do so, we first perform some preprocessing steps such as masking, rotating, scaling, translating and interpolating. After that we implement the pose and illumination compensation on 3D images. We also apply a feature extraction procedure, either the Principle Component Analysis (PCA) or Independent Component Analysis (ICA), to extract the depth and intensity features of 3D images. Finally, we estimate the 3D feature vector from the given 2D photos by using the optimization algorithm.

A human face may be modeled by a combination of faces. To illustrate the concept, let us think about a child who has some facial features resembling her father, mother, sister, brother, grandmother or grandfather (figure 3.1). Furthermore, certain features of a face may be similar to faces of other unrelated individuals.

In this research, it is assumed that a 3D face image is decomposed into a depth feature vector \mathbf{z} and intensity feature vector \mathbf{t} . A face model $(\mathbf{z}^f, \mathbf{t}^f)$ may then be represented as a linear combination of depth and intensity vectors shown in equation 3.1, where \mathbf{z}_i^f and \mathbf{t}_i^f , $i = 1, \dots, k$, are the i th depth and intensity vectors, a_i and b_i are the weights (parameters) representing the contribution of the i th vector, and k is the number of vectors. The task is now to select a_i and b_i to make a new face $(\mathbf{z}^f, \mathbf{t}^f)$ similar to the probing face.

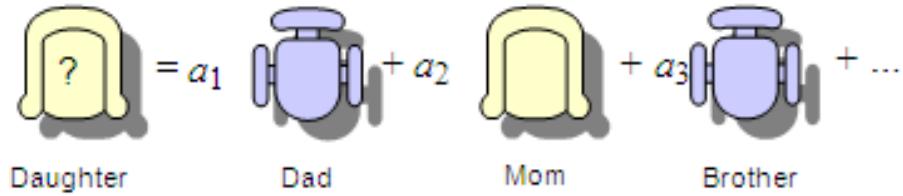


Figure 3.1: Conceptual Illustration of Linear Combination of Faces

$$\mathbf{z}^f = \sum_{i=1}^k a_i \mathbf{z}_i^f, \quad \mathbf{t}^f = \sum_{i=1}^k b_i \mathbf{t}_i^f \quad (3.1)$$

Since a main application of the proposed method is biometric authentication, more discussion on this subject is given next. An automatic authentication system is normally made up of an enrollment stage and a verification stage. In the enrollment stage, individuals who are supposed to have the authority of accessing are registered with their 3D photos. In the verification stage, only 2D face images are captured and compared with the 3D images stored in the gallery database for authentication. A 3D face model is created by selecting weights of 3D depth and intensity such that the projected 2D image of the 3D face is similar to the 2D probing image. Only the 3D weights are needed to be

saved in the database, then the general feature extraction procedure is applied to compare the probing image's 3D parameters with those saved in the database to determine the identity of the person. A detailed description of the enrollment and verification stages is given next.

3.2 Preprocessing Techniques

Preprocessing is conducted by performing the following five steps to convert a raw GTI image to a normalized image. The normalization schedule is:

1. Integer to float conversion - Converts 256 gray levels into floating point equivalents.
2. Geometric normalization – Lines up human chosen eye coordinates.
3. Masking – Crops the image using an elliptical mask and image borders such that only the face from forehead to chin and cheek to cheek is visible.
4. Histogram equalization – Equalizes the histogram of the unmasked part of the image.
5. Pixel normalization – scales the pixel values to have a mean of zero and a standard deviation of one.

In order to preprocess images, the exact coordinates of the eyes must be provided.

In this research, we manually identify eye coordinates.

3.2.1 Parameterized Model

The 3D image is a 3D array which represents samples of a certain physical quantity acquired over a rectangular 3D grid. A 3D image can be considered as a 3D

matrix $a([x][y][z][t])$, where x , y , and z denote slice (image), row and column as shown in figure 3.2. The variable t denotes the gray scale value at (x, y, z) . For each 3D image, one has a matrix of points as shown below:

$$\begin{bmatrix} x_1 & y_1 & z_1 & t_1 \\ x_2 & y_2 & z_2 & t_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_n & y_n & z_n & t_n \end{bmatrix} \quad (3.2)$$

In equation 3.2, z_i , $i = 1, 2, \dots, n$, is the depth corresponding to coordinates x_i and y_i ; and t_i , $i = 1, 2, \dots, n$, is the intensity corresponding to coordinates x_i , y_i and z_i . The data given in equation 3.2 uniquely represents a 3D image of a subject. An example of 3D images is given in figure 3.3.

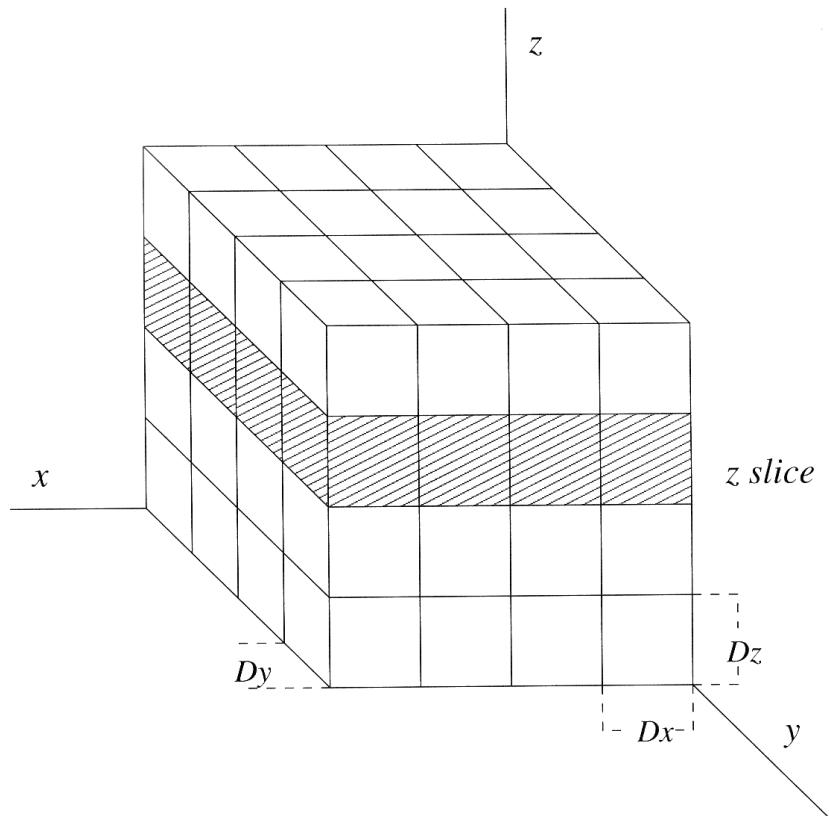


Figure 3.2: Cubic Representation of 3D Object

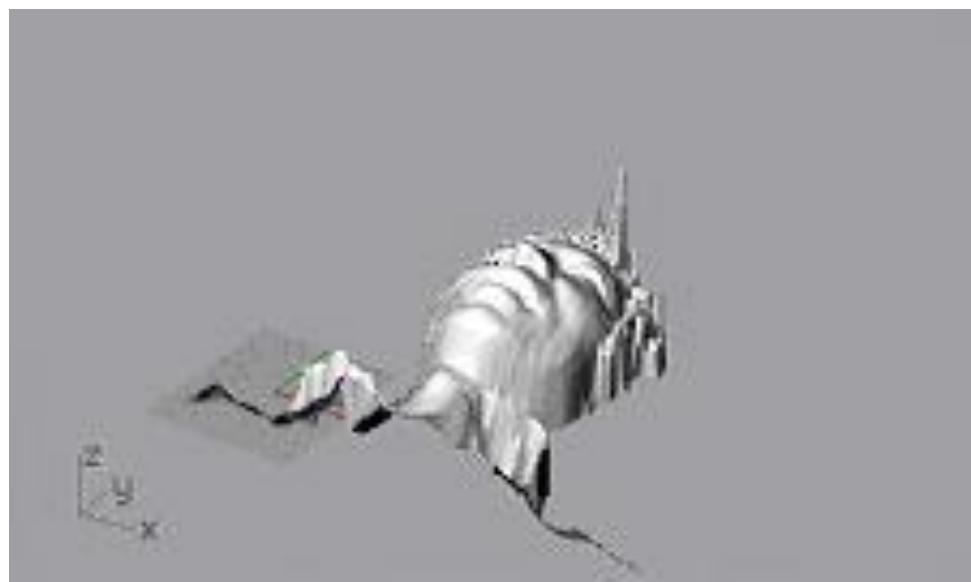


Figure 3.3: An Example of 3D Images

3.2.2 Masking

To eliminate the interference caused by the background, we crop the image using the elliptical mask, and only the face area that from forehead to chin and cheek to cheek is visible [97].

3.2.3 3D Rotation

Image rotation is often employed in visualization and this can be performed also on 3D face images. A 3D image can be observed from different viewpoints by keeping one view point fixed and rotating the shape. The calculation on the 3D shape for rotation requires all points in the original image be traversed and each point be mapped into the new position. To align the projected 2D image with the input image, we rotate the 3D face image on the x - y plane and then on the x - z plane. Let θ and γ be the rotation angles onto the x - y and x - z planes, respectively. Then, the rotation matrix R for this task is

$$R = \begin{bmatrix} \cos \theta \cos \gamma & -\sin \theta \cos \gamma & -\sin \gamma \\ \sin \theta & \cos \theta & 0 \\ \cos \theta \sin \gamma & -\sin \theta \sin \gamma & \cos \gamma \end{bmatrix} \quad (3.3)$$

3.2.4 3D Data Fitting

The raw 3D face data are represented in the form of scattered data point cloud. To reduce the data amount without losing the key face features, and also to reduce noises that may contained in the data to the maximum extend, the data should be fitted to be a standard form[98, 99, 100, 101].

3.2.4.1 3D Linear Interpolation

After a 3D rotation, the resultant x and y coordinates are not necessarily on the grid points. A simple linear interpolation step can be used to produce the x and y coordinates in integers. Assuming that we need to compute the integer z_D given x_D and y_D , the equation for interpolation is given in equation 3.4. For each output point, three neighboring points are employed, as shown in figure 3.4. In equation 3.4, x_n , y_n , and z_n are intermediate variables.

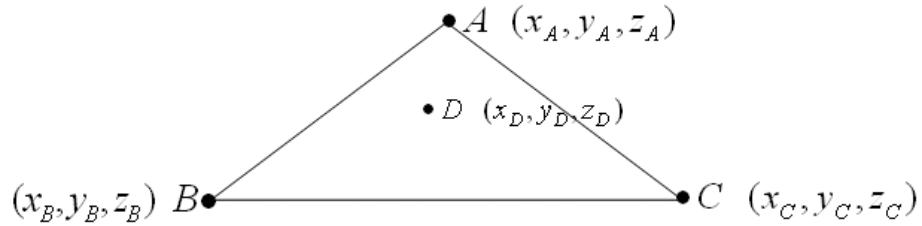


Figure 3.4: Triangulate Interpolation

$$\begin{aligned}
 x_n &= (y_A - y_B)(z_A - z_C) - (z_A - z_B)(y_A - y_C) \\
 y_n &= (z_A - z_B)(x_A - x_C) - (x_A - x_B)(z_A - z_C) \\
 z_n &= (x_A - x_B)(y_A - y_C) - (y_A - y_B)(x_A - x_C) \\
 z_D &= \frac{(x_A x_n + y_A y_n + z_A z_n - x_D x_n - y_D y_n)}{z_n}
 \end{aligned} \tag{3.4}$$

3.2.4.2 Multilevel B-Splines Approximation

Consider a set of scattered points $P = \{(x, y, z)\}$ in 3D space, let $\Omega = \{0 \leq x \leq m, 0 \leq y \leq n\}$ be a rectangular domain in the xy -plane, and (x, y) is a point in Ω . We formulate an $m \times n$ lattice to approximate the scattered data P . A uniform grid in Ω is spanned and the lattice

is defined by a control lattice Φ overlaid on domain Ω . Let \emptyset_{ij} be the value of the ij -th control point on lattice Φ , the approximation function of lattice Φ is defined:

$$f(x, y) = \sum_{k=0}^3 \sum_{l=0}^3 B_k(s) B_l(t) \emptyset_{(i+k)(j+l)} \quad (3.5)$$

$$\begin{cases} B_0(t) = (1-t)^3/6 \\ B_1(t) = (3t^3 - 6t^2 + 4)/6 \\ B_2(t) = (-3t^3 + 6t^2 + 3t + 1)/6 \\ B_3(t) = t^3/6 \end{cases} \quad (3.6)$$

Where $i = |x| - 1, j = |y| - 1, s = x - |x|, t = y - |y|$. B_k and B_l are the basis functions of the uniform cubic B-splines [5].

The problem now is to solve for the control points in Φ so that $f(x, y)$ can correspondingly best approximate the scattered data in P .

Let $f(x, y) = Z_c, w_{kl} = B_k(s) B_l(t), k = (i+1) - |x|, l = (j+1) - |y|$ and by equation 3.6 we can get

$$\emptyset_{kl} = \frac{w_{kl} Z_c}{\sum_{a=0}^3 \sum_{b=0}^3 w_{ab}^2} \quad (3.7)$$

To get the final value of \emptyset_{ij} approximating to the surface, we denote $e(\emptyset_{ij}) = \sum (w_{kl} \emptyset_{kl} - w_{kl} \emptyset_{ij})^2$ and let it be minimum by differentiating the $e(\emptyset_{ij})$ with respect to \emptyset_{ij} , we get

$$\emptyset_{ij} = \frac{\sum_c w_{kl}^2 \emptyset_{kl}}{\sum_c w_{kl}^2} \quad (3.8)$$

3.2.5 Depth Parameter Extraction

A useful and popular technique for data analysis and compression is Principle Component Analysis (PCA), although it is not very effective in discriminating data

patterns. Other methods such as LDA (Linear Discriminant Analysis) or ICA (Independent Component Analysis) may also be used to replace PCA for 3D face modeling without affecting the overall scheme. Two PCAs, depth and intensity, are employed here. Please also refer to Section 3.6 for the detail of Principal Component Analysis.

For the depth PCA, we need to use depth measurements from 3D image data. It has been mentioned that the matrix of points given in equation 3.2 is used for each 3D face model. In the depth PCA, we only need to use the data given in the following matrix:

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_n & y_n & z_n \end{bmatrix} \quad (3.9)$$

For the readability of the paper, the PCA algorithm is briefly summarized below. Let ξ_i be a set of depth values (z_1, z_2, \dots, z_k) of the i th subject. Then, $\xi_1, \xi_2, \dots, \xi_k$ are the depth vectors of the entire training set consisting k subjects. The average depth is given by:

$$\bar{\xi} = \frac{1}{k} \sum_{i=1}^k \xi_i \quad (3.10)$$

The difference from average depth for each subject is as follows:

$$\varphi_i = \xi_i - \bar{\xi} \quad i = 1, 2, \dots, k \quad (3.11)$$

The covariance matrix is defined as:

$$R_z = \frac{1}{k} \sum_{i=1}^k \varphi_i \varphi_i^T = DD^T \quad (3.12)$$

where $D = [\varphi_1 \ \varphi_2 \ \dots \ \varphi_k]$ and the size of D is $n \times k$. The R_z matrix is of size $n \times n$ which is large and leads to computationally complex calculations for eigenvalues and eigenvectors. An alternative approach is to use $\hat{R}_z = D^T D$.

The eigenvalues and eigenvectors of \hat{R}_z are determined by:

$$\hat{R}_z \mathbf{q}_{z,i} = \lambda_{z,i} \mathbf{q}_{z,i} \quad (3.13)$$

Where $\lambda_{z,i}$ and $\mathbf{q}_{z,i}$ are the i th eigenvalue and eigenvector, respectively. Let $Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_k]$. Note that the eigenvalues are placed in a descending order. The eigenvectors are subsequently normalized. And these are related to the eigenvectors of DD^T

$$E = DQ \quad (3.14)$$

The columns of E span the so-called depth eigenspace. The next step is to project all depth vectors to the depth eigenspace, by

$$\mathbf{z}_i^f = E^T \boldsymbol{\varphi}_i \quad (3.15)$$

and the size of \mathbf{z}_i^f depends on the number of eigenvectors selected. To recreate the depth map, we can calculate

$$\hat{\mathbf{z}}_i = (E\mathbf{z}_i^f) + \bar{\boldsymbol{\xi}} \quad (3.16)$$

where $\hat{\mathbf{z}}_i$ is the recreated depth vector for the i th subject.

3.2.6 Intensity Parameter Extraction

The procedure for the intensity PCA is quite similar to that of the depth PCA. For the intensity PCA, the data from the following matrix is employed:

$$\begin{bmatrix} x_1 & y_1 & t_1 \\ x_2 & y_2 & t_2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_n & y_n & t_n \end{bmatrix} \quad (3.17)$$

3.3 3D Face Recreation Method

In order to recreate a 3D face, we need to estimate coefficients a_i and b_i given in equation 3.1, which correspond to the depth and intensity features. For this purpose, an optimization procedure is employed. The optimization procedure includes the formulation of a cost function, initial conditions, an update rule, and terminal conditions. A quadratic cost function J given in equation 3.18 is constructed, in which g is the feature vector of an input 2D image, $\hat{g}(\rho)$ is the feature vector of the projected 2D image from the recreated 3D face using equation 3.1, and $\rho = [a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_k]^T$.

The optimization problem is to choose the parameter vector ρ such that the following cost J is minimized:

$$J = (g - \hat{g}(\rho))^T (g - \hat{g}(\rho)) \quad (3.18)$$

The proposed update rule is shown in equation 3.19 and 3.20,

$$\rho_{j+1} = \rho_j + w_j \Delta \rho_j \quad (3.19)$$

$$\Delta \rho_j = f(g - \hat{g}(\rho_j)) \quad (3.20)$$

where the subscript j indicates the j th iteration, ρ_j is the parameter vector, w_j is a number between 0 and 1, controlling the size of the parameters adjustment, $\Delta\rho_j$ is the parameter adjustment produced by equation 3.20, and $f(\cdot)$ is a function that calculates the parameter adjustment based on the errors between the given and the estimated feature vectors. For the nonlinear least squares problem, an often-used practice is to choose f as a linear mapping realized with a pseudo inverse.

To start the iteration, the average face of the models in the database is mapped to the eigenspace, from which an initial condition for the parameter vector ρ is obtained.

The algorithm can be terminated when either the cost function is less than a prescribed threshold value, or the number of iterations exceeds a given number.

3.4 Initial Pose Determination

Our starting point in this study is to assume that as the face changes its pose, the corresponding change in the 2D positions of the projected features can be approximated by an affine transformation. The estimation of the pose angle of the 2D probe photo uses the affine transformation between the feature positions in the 2D probe photo and their corresponding positions in the 3D synthetic face image in which the view of the face is fronto-parallel [102]. In our study, the 4 feature points we use are the center of each eye, the middle of the mouth, and the nose tip.

The three rotation angles, yaw, pitch, and roll, are described in Figure 3 below. In this study, we ignored pitch angle and focused on yaw and roll angles, as the pitch angle

is very small for the faces saved in our database. The framework, though, is still valid when the pitch angle cannot be ignored, in which case a more sophisticated algorithm [103] needs to be adopted for pose determination.

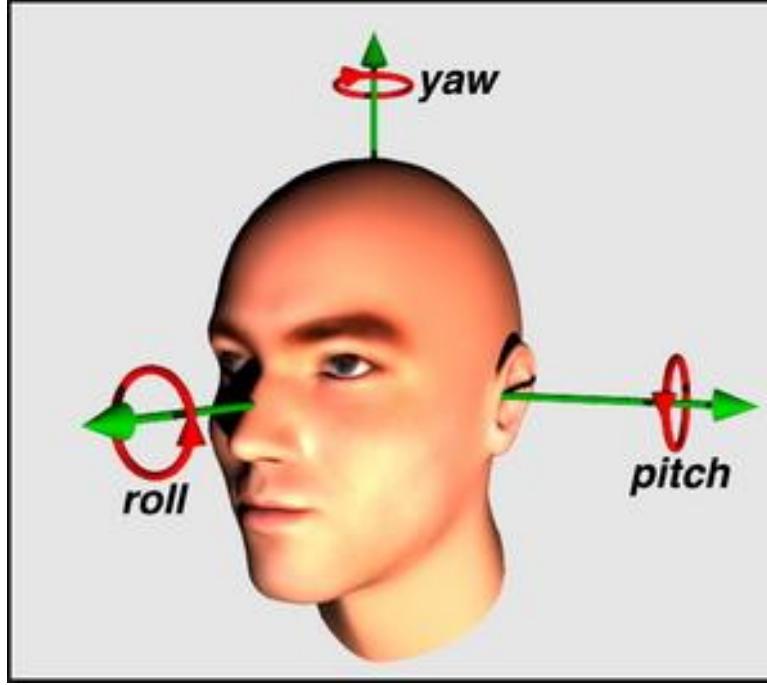


Figure3.5: The Yaw, Pitch, and Roll Angle

Let (u, v) be the 2D image coordinates and (x, y, z) be 3D coordinates. The resulting 3D synthetic image needs to be rotated so that its 2D projection matches the probing photo. Under the weak perspective transformation assumption, (u, v) can be related to (x, y, z) below:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{S}(s) \times \mathbf{R}(\phi, \theta, \psi) \times \mathbf{T}(x_0, y_0, z_0) \times \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.21)$$

where $\mathbf{S}(s)$ is the scaling matrix, $\mathbf{R}(\phi, \psi, \theta)$ is the rotation matrix in roll, pitch and yaw angles, and $\mathbf{T}(x_0, y_0, z_0)$ is the translation matrix [104].

To find the solution for the unknowns, Huttenlocher and Ullman proposed a closed form solution under the weak perspective transformation assumption [103], where 3 points in both the 2D image plane and the 3D space are used for face alignment. However, the drawback of this method is that the error will propagate, resulting large errors to the variables solved in the later steps.

We implement an alternative iterative method also under the weak perspective transformation assumption. To obtain an initial condition of the iterative algorithm, we assume the pitch angle is relatively small, and the roll angle can be roughly compensated in the 2D image plane. Therefore equation 3.21 is simplified to

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 0 & 0 & 0 \\ \sin\theta & 0 & -\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3.22)$$

The translation part $\mathbf{T}(x_0, y_0, z_0)$ can be removed:

$$u_1 - u_2 = s(\cos\theta(x_1 - x_2) + \sin\theta(z_1 - z_2)) \quad (3.23)$$

$$v_1 - v_2 = s(y_1 - y_2) \quad (3.24)$$

From equation 3.24, s can be solved, and from equation 3.23, the yaw angle θ can be solved. Then x_0 and y_0 can be determined. Note that under the weak perspective transformation, z_0 is not relevant. Once the initial condition is obtained, a nonlinear least

squares algorithm can then be applied to find all the unknown parameters more accurately.

The detailed pose determination procedure is described as follows:

- **Step 1**

For each 3D image, we assume that the (x, y, z) coordinates of 4 data points are given: Q_1 and Q_2 are the middle of each eye, Q_3 is the nose tip, and Q_4 is the middle of the lips. It is convenient to translate and rotate these points in the x - y and y - z planes since these points are in a 3D space. Similarly, for the 2D test image, we also assume that the (x, y) coordinates of the corresponding 4 data points, P_1 to P_4 , are given.

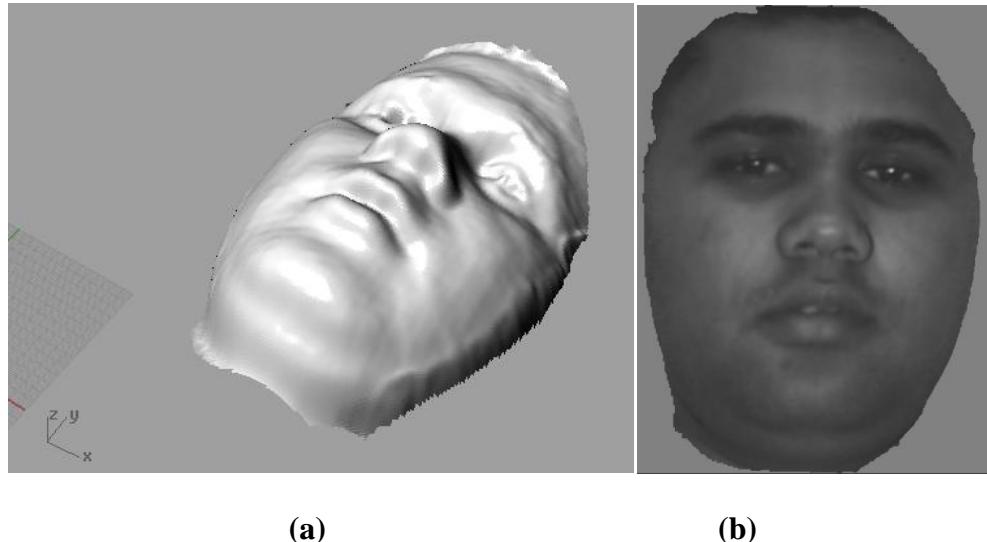


Figure 3.6: 3D Image of the Fronto-Parallel Pose. (a) The Shape Data in the x-y-z Coordinate.

(b) The Intensity Data of the Fronto-Parallel Pose.



Figure 3.7: The 2D Data Image with Slightly Side View

- **Step 2**

Normalize both 2D and 3D images to assure that the coordinate origins are the center of the two eyes.

$$\begin{aligned} P_{1x} + P_{2x} &= 0 \\ P_{1y} + P_{2y} &= 0 \\ Q_{1x} + Q_{2x} &= 0 \\ Q_{1y} + Q_{2y} &= 0 \end{aligned} \tag{3.25}$$

- **Step 3**

Rotate the 3D data point Q_i in the y - z plane as shown below, to match the pose of P_i in the 2D image.

$$\begin{aligned} Q_{ix}' &= l \times (Q_{ix} \cos(a) + Q_{iz} \sin(a)) \\ Q_{iy}' &= l \times Q_{iy} \end{aligned} \quad (3.26)$$

Note that there are also two unknown variables involved: an angle factor a and a scaling factor l .

- **Step 4**

Calculate the mean square error (MSE) in the x - y plane.

$$\begin{aligned} \text{MSE} &= \sqrt{\sum_{i=1}^4 ((P_{ix} - Q_{ix}')^2 + (P_{iy} - Q_{iy}')^2)} \\ &= \sqrt{2 \times (P_{1x} - Q_{1x}')^2 + (P_{3x} - Q_{3x}')^2 + (P_{3y} - Q_{3y}')^2 + (P_{4x} - Q_{4x}')^2 + (P_{4y} - Q_{4y}')^2} \end{aligned} \quad (3.27)$$

- **Step 5**

Differentiate the MSE of l and a to minimize the MSE with regard to parameters l and a as follows:

$$\frac{d(\text{MSE})}{dl} = 0 \quad (3.28)$$

Expend equation 3.28 we get the following equation 3.29:

$$\frac{d(2 \times (P_{1x} - Q_{1x}')^2 + (P_{3x} - Q_{3x}')^2 + (P_{3y} - Q_{3y}')^2 + (P_{4x} - Q_{4x}')^2 + (P_{4y} - Q_{4y}')^2)}{dl} = 0 \quad (3.29)$$

Further expend it we got equation 3.30:

$$\begin{aligned}
& 2 \times (P_{1x} - l \times (Q_{1x} \cos(a) + Q_{1z} \sin(a))) \times (Q_{1x} \cos(a) + Q_{1z} \sin(a)) \\
& - (P_{3x} - l \times (Q_{3x} \cos(a) + Q_{3z} \sin(a))) \times (Q_{3x} \cos(a) + Q_{3z} \sin(a)) \\
& - (P_{3y} - l \times Q_{3y}) \times Q_{3y} \\
& - (P_{4x} - l \times (Q_{4x} \cos(a) + Q_{4z} \sin(a))) \times (Q_{4x} \cos(a) + Q_{4z} \sin(a)) \\
& - (P_{4y} - l \times Q_{4y}) \times Q_{4y} = 0
\end{aligned} \tag{3.30}$$

The same procedure applied to a :

$$\frac{d(\text{MSE})}{da} = 0 \tag{3.31}$$

$$\frac{d(2 \times (P_{1x} - Q_{1x})^2 + (P_{3x} - Q_{3x})^2 + (P_{3y} - Q_{3y})^2 + (P_{4x} - Q_{4x})^2 + (P_{4y} - Q_{4y})^2)}{da} = 0 \tag{3.32}$$

$$\begin{aligned}
& -2 \times (P_{1x} - l \times (Q_{1x} \cos(a) + Q_{1z} \sin(a))) \times (-Q_{1x} \sin(a) + Q_{1z} \cos(a)) \\
& - (P_{3x} - l \times (Q_{3x} \cos(a) + Q_{3z} \sin(a))) \times (-Q_{3x} \sin(a) + Q_{3z} \cos(a)) \\
& - (P_{4x} - l \times (Q_{4x} \cos(a) + Q_{4z} \sin(a))) \times (-Q_{4x} \sin(a) + Q_{4z} \cos(a)) = 0
\end{aligned} \tag{3.33}$$

• Step 6

Eliminate a parameter l by combining equations 3.30 and 3.33. After some manipulations, the result is the third order trigonometric function as shown below in equation 3.34:

$$\begin{aligned}
& \left(\begin{array}{l} 2P_{1x} \times (Q_{1x} \cos(a) + Q_{1z} \sin(a)) \\ + P_{3x} \times (Q_{3x} \cos(a) + Q_{3z} \sin(a)) \\ + P_{4x} \times (Q_{4x} \cos(a) + Q_{4z} \sin(a)) \\ + P_{3y} \times Q_{3y} + P_{4y} \times Q_{4y} \end{array} \right) \times \left(\begin{array}{l} 2 \times (Q_{1x} \cos(a) + Q_{1z} \sin(a)) \\ \times (-Q_{1x} \sin(a) + Q_{1z} \cos(a)) \\ + (Q_{3x} \cos(a) + Q_{3z} \sin(a)) \\ \times (-Q_{3x} \sin(a) + Q_{3z} \cos(a)) \\ + (Q_{4x} \cos(a) + Q_{4z} \sin(a)) \\ \times (-Q_{4x} \sin(a) + Q_{4z} \cos(a)) \end{array} \right) \\
& = \\
& \left(\begin{array}{l} 2 \times (Q_{1x} \cos(a) + Q_{1z} \sin(a))^2 \\ + (Q_{3x} \cos(a) + Q_{3z} \sin(a))^2 \\ + (Q_{4x} \cos(a) + Q_{4z} \sin(a))^2 \\ + (Q_{3y})^2 + (Q_{4y})^2 \end{array} \right) \times \left(\begin{array}{l} 2 \times P_{1x} \times (-Q_{1x} \sin(a) + Q_{1z} \cos(a)) \\ + P_{3x} \times (-Q_{3x} \sin(a) + Q_{3z} \cos(a)) \\ + P_{4x} \times (-Q_{4x} \sin(a) + Q_{4z} \cos(a)) \end{array} \right)
\end{aligned} \tag{3.34}$$

• Step 7

Recover the angle factor a by using a standard iterative method from equation 3.34.

Figure 3.8 is the result of the above sample images.

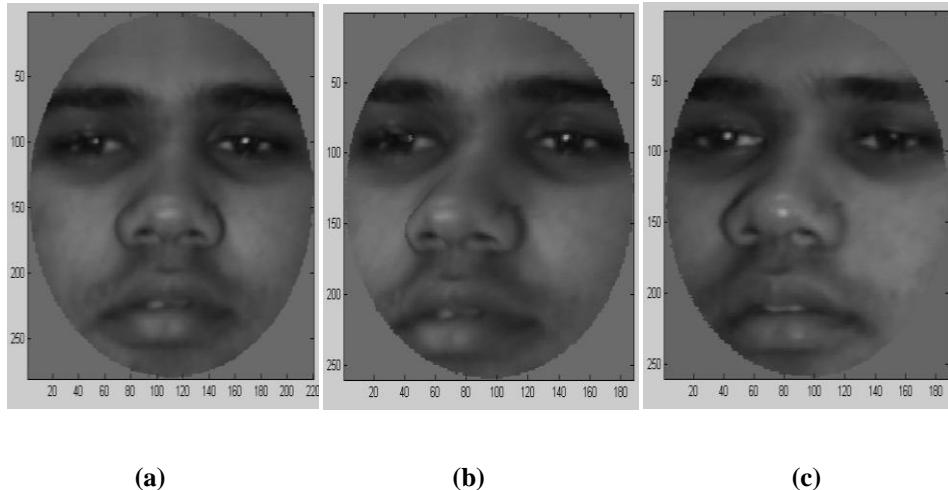


Figure 3.8: Sample Image Rotation. (a) The straight fronto-Parallel view directly got from the 3D-data image. (b) The same data rotate a degree as mentioned in the algorithm above. (c) The original 2D data image.

Comparing the (b) and (c) above, it's almost in the same position and angle, then we will use this data for further recognition.

3.5 Illumination Compensation

The Phong reflection model [105] is a shading model used heavily in 3D computer graphics for assigning shades to each individual pixel of an object. It was developed by Bui Tuong Phong in 1973. It considers the reflection from a surface to consist of three linearly combined components, ambient, diffused, and specular:

$$\text{Reflected light} = \text{ambient light} + \text{diffuse component} + \text{specular component}$$

(3.35)

The ambient term is a constant and simulates global or indirect illumination. This term is necessary because parts of a surface that cannot “see” the light source, but can be seen by the viewer, need to be lit. Otherwise they would be rendered as black. In reality such lighting comes from global or indirect illumination.

It is useful to consider what type of surface such a model simulates. Linear combination of a diffuse and specular component occurs in polished surfaces, and specular reflection results from the transparent layer and diffuse reflection from the underlying surface [106].

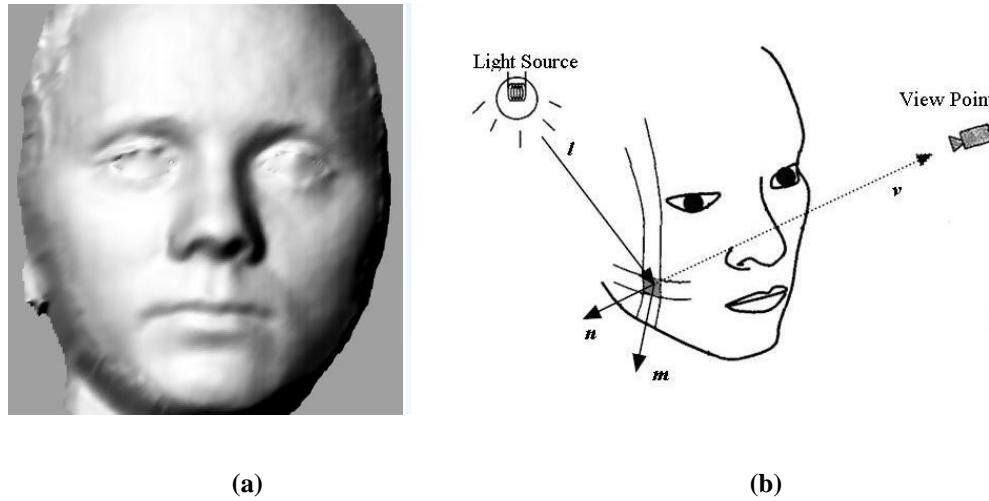


Figure 3.9: (a) 3D Face Shape with (b) Phong Reflectance Model

Suppose the illumination is in the direction \mathbf{l} , with irradiance $I(\lambda) = I(\lambda, -l)$. Let \mathbf{n} be the surface normal, and \mathbf{m} be the mirror reflection direction. Then the reflected radiance used by the Phong reflectance model at a surface point x_p , per unit area perpendicular to the viewing direction \mathbf{v} , is

$$R(\lambda, \mathbf{x}_p, \mathbf{v}) = k_a r(\lambda) + \sum_{\text{Lights}} (k_d r(\lambda) (\mathbf{n} \cdot \mathbf{l}) I(\lambda) + k_s S(\lambda) (\mathbf{m} \cdot \mathbf{v})^{k_e}) \quad (3.36)$$

where $r(\lambda)$ is the diffuse spectral reflectance distribution for the surface, k_a , k_d , k_s , are non-negative coefficients for the ambient, diffuse, and specular reflection terms, respectively; k_e is the spectral exponent, controlling the spread of the specular reflection (rougher surfaces modeled by smaller k_e); and $S(\lambda)$ is the spectral distribution of the specular reflection. It is just $I(\lambda)$ for painted or plastic surfaces. For metals it can be approximated by some linear combination of $I(\lambda)$ and $r(\lambda)$.

The basic way of rendering of a scene is on a polygon-by-polygon basis, where each polygon is rendered in turn, in isolation from all the rest. The order of rendering a scene places restrictions upon which hidden surface algorithms can be used, but is of itself independent of the method employed for hidden surface removal. Polygon-by-polygon rendering is simple to implement, and it requires little data active at any one time. Because of this, it places no upper limit on scene complexity. The common hidden surface removal algorithm that is compatible with this method is Z-buffer Catmull (1975).

The combination of the Z-buffer algorithm, the Phong model and interpolator represents one of the most popular rendering options [12]. Pixels in the interior of a polygon are shaded, using an incremental shading scheme, and their depth is evaluated by interpolation from the z values of the polygon vertices after a viewing transformation has been applied. This algorithm is equivalent, for each point (x, y) to a search through the associated z values of each interior polygon point, to find that point with the minimum z

value. This search is conveniently implemented by using a Z-buffer, which holds for a current point (x, y) the smallest z value so far encountered. During the processing of a polygon we decide to either write the intensity of a point (x, y) into the frame buffer or not, depending on whether the depth z , of the current point, is less than the depth so far encountered in the Z-buffer.

3.6 Principal Component Analysis

Eigenfaces [108, 109] have become one of the major contributors behind face representation and recognition, for its well known characteristic of reducing the significant statistical redundancies in natural images [110]. For face images that are normalized with respect to scale, translation, and rotation, the redundancy is even greater [53, 111]. PCA is one of the best global compact representations to decorrelate the outputs. Typically, sample vectors \mathbf{x} can be expressed as linear combinations of the orthogonal basis Φ_i : $\mathbf{x} = \sum_{i=1}^n a_i \Phi_i \approx \sum_{i=1}^m a_i \Phi_i$ (typically $m \ll n$) by solving the eigen problem

$$C\Phi = \Phi\Lambda \quad (3.37)$$

where C is the covariance matrix for input \mathbf{x} .

Such representation has one significant advantage of reducing sensitivity to noise. Good performance under blurring, partial occlusion and changes in background has been demonstrated in many eigenpicture-based systems as the result, since the PCA reconstructed images are much better than the original distorted images in terms of their global appearance.

Using extended training sets with additional mirror-imaged faces was shown to achieve lower approximation error [108]. With these sets, the eigenfaces are either symmetric or antisymmetric, with the most leading eigenfaces typically being symmetric.

Turk and Pentland made the first successful demonstration of machine face recognition [6] using eigenfaces. Given the eigenfaces, every face in the database was represented by a vector of weights; the weights are obtained by projecting the image into eigenface components by a simple inner product operation. Given a new test image for recognition, it is also represented by its vector of weights. The identification of the test image is done by locating the image in the database whose weights are the closest to the weights of the test image.

Moghaddam and Pentland extended [42] the standard eigenface method to a Bayesian approach by using a probabilistic measure of similarity. Generally speaking the Bayesian method needs to estimate probability distributions in a high dimensional space from very limited numbers of training samples per class. To avoid this problem, a similarity measure based on a Bayesian analysis of image differences are used to create a simpler two-class problem instead of the multiclass problem. Two mutually exclusive classes were defined: Ω_I , representing *intrapersonal* variations between multiple images of the same individual, and Ω_E , representing *extrapersonal* variations due to differences in identity. Assuming that both classes are Gaussian-distributed, likelihood functions $P(\Delta|\Omega_I)$ and $P(\Delta|\Omega_E)$ were estimated for a given intensity difference $\Delta = I_1 - I_2$. Given these likelihood functions and using the MAP rule, two face images are determined to belong to the same individual if $P(\Delta|\Omega_I) > P(\Delta|\Omega_E)$. The standard nearest-neighbor

eigenspace matching can be greatly improved by using this probabilistic matching technique. Phillips et al. reported the result using large face datasets including the FERET database [112]. Moghaddam and Pentland [42] proposed an efficient technique of probability density estimation by decomposing the input space into two mutually exclusive subspaces: the principal subspace F and its orthogonal subspace \hat{F} [113]. Only in the principal subspace Covariances are estimated for use in the Mahalanobis distance [114]. Experimental results have been reported using different subspace dimensionalities for Ω_I and Ω_E . For internal test, $M_I = 10$ and $M_E = 30$ were used, and for FERET test $M_I = M_E = 125$ were used. The extrapersonal eigenfaces appear more similar to the standard eigenfaces than the intrapersonal ones, and the intrapersonal eigenfaces represent subtle variations due mostly to expression and illumination, which are more critical for identification [42].

The detailed presentation of PCA is as follows:

Suppose we have a data set \mathbf{X} , where

$$\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_N]^T \quad (3.38)$$

and the mean of that data set is denoted by

$$\bar{\mathbf{X}} = \sum_{i=1}^N \mathbf{X}_i \quad (3.39)$$

and the covariance matrix of the same data set is

$$\text{cov}(\mathbf{X}) = \frac{\sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T}{(N-1)} \quad (3.40)$$

The covariance matrix is, by definition, always symmetric.

From a sample of vectors X_1, \dots, X_N , we can calculate the sample mean and the sample covariance matrix as the estimates of the mean and the covariance matrix.

From a symmetric matrix such as the covariance matrix, we can calculate an orthogonal basis by finding its eigenvalues and eigenvectors. The eigenvectors e_i and the corresponding eigenvalues λ are the solutions of the equation

$$Cov(\mathbf{X}) \cdot \mathbf{e}_i = \lambda_i \cdot \mathbf{e}_i, \text{ for } i = 1, \dots, N \quad (3.41)$$

For simplicity we assume that the λ_i are distinct. These values can be found, for example, by finding the solutions of the characteristic equation

$$|Cov(\mathbf{X}) - \lambda \mathbf{I}| = 0 \quad (3.42)$$

where the \mathbf{I} is the identity matrix having the same order as $Cov(x)$ and the $|.|$ denotes the determinant of the matrix. If the data vector has N components, the characteristic equation becomes of order N .

By ordering the eigenvectors in the order of descending eigenvalues, one can create an ordered orthogonal basis with the first eigenvector having the direction of largest variance of the data. In this way, we can find directions in which the data set has the most significant amounts of energy.

Suppose one has a data set of which the sample mean and the covariance matrix have been calculated. Let \mathbf{A} be a matrix consisting of eigenvectors of the covariance matrix as the row vectors.

By transforming a data vector \mathbf{x} , we get

$$\mathbf{Y} = \mathbf{A}(\mathbf{X} - \bar{\mathbf{X}}) \quad (3.43)$$

which is a point in the orthogonal coordinate system defined by the eigenvectors. Components of \mathbf{y} can be seen as the coordinates in the orthogonal base.

This means that we project the original data vector on the coordinate axes having the dimension K and transforming the vector back by a linear combination of the basis vectors. This minimizes the mean-square error between the data and this representation with given number of eigenvectors.

If the data is concentrated in a linear subspace, this provides a way to compress data without losing much information and simplifying the representation. By picking the eigenvectors having the largest eigenvalues we lose as little information as possible in the mean-square sense. One can choose a fixed number of eigenvectors and their respective eigenvalues and get a consistent representation, or abstraction of the data. This preserves a varying amount of energy of the original data. Alternatively, we can choose approximately the same amount of energy and a varying amount of eigenvectors and their respective eigenvalues. This would in turn give approximately consistent amount of information in the expense of varying representations with regard to the dimension of the subspace.

3.7 Linear Discriminant Analysis

It is also a very successful practice to use LDA/FLD (Fisher Linear Discriminant) in the face recognition systems [7, 43, 44, 51, 111]. Generally the training is done using scatter matrix analysis [114], and the within- and between-class scatter matrices S_w , S_b are computed as follows:

$$S_w = \sum_{i=1}^M P_r(\omega_i) C_i$$

$$S_b = \sum_{i=1}^M P_r(\omega_i) (\mathbf{m}_i - \mathbf{m}_0)(\mathbf{m}_i - \mathbf{m}_0)^T \quad (3.44)$$

where $P_r(\omega_i)$ is the prior class probability, and is usually replaced by $1/M$ in practice with the assumption of equal priors. *Within-class scatter matrix* S_ω shows the average scatter C_i of the sample vectors \mathbf{x} of different classes ω_i around their respective means m_i : $C_i = E[(x_\omega - m_i)(x_\omega - m_i)^T | \omega = \omega_i]$. Similarly, the *between-class Scatter Matrix* S_b , represents the scatter of the conditional mean vectors \mathbf{m}_i around the overall mean vector \mathbf{m}_0 . To quantify discriminatory power a commonly used measure is applied: $J(T) = |T^T S_b T| / |T^T S_\omega T|$. By solving the generalized eigenvalue problem, the optimal projection matrix W which maximizes $J(T)$ can be obtained:

$$S_b W = S_\omega W \Lambda_W \quad (3.45)$$

Discriminant analysis of eigenfeatures is applied in an image retrieval system to determine both the class (human face vs. nonface objects) and individuals within the face class [43]. Using tree-structure learning, the eigenspace and LDA projections are recursively applied to smaller and smaller sets of samples. Such recursive partitioning is carried out for every node until the samples assigned to the node belong to a single class. Experiments on this approach were reported with the result of 91% for 78 face images on images not in the training set; and 87% for 38 nonface images [43].

Belhumeur et al. [7] carried out a comparative performance analysis on four methods:

1. A correlation-based method,

2. A variant of the linear subspace method suggested by Shashua [115],
3. An eigenfaces method Turk and Pentland [6], and
4. Fisherface method which uses subspace projection [43].

The comparison results showed that the Fisherface method significantly outperformed the other three methods. However, no further experimental claim was made on larger databases.

A regularized subspace LDA system that unifies PCA and LDA was proposed by Zhao et al. [44, 111] to improve the performance of LDA based systems. Experimental results demonstrated good generalization ability. Later Penev and Sirovich [116] concluded that the global face subspace dimensionality is on the order of 400s for large databases of 5,000 images or more. To improve performance, Zhao used a weighted distance metric in the projection space z [111]. Finally, the LDA training was regularized by modifying the S_ω matrix to $S_{\omega+\delta I}$, where δ is a relatively small positive number. Doing this solves a numerical problem when S_ω is close to being singular. This regularization transforms the LDA problem into a standard PCA problem in the extreme case where only one sample per class is available, and in this case, S_b becomes the covariance matrix C . The performance report showed that based on a fronto-parallel pose gallery database of 738 images: 85.2% for all images and 95.1% for fronto-parallel pose.

3.8 Optimization Method

3.8.1 Non-Linear Least Square Problem

Least squares problems can be solved by general optimization methods, but we want to find special methods that are more efficient. In many cases they achieve better than linear convergence, sometimes even quadratic convergence, even though they do not need implementation of second derivatives.

Nonlinear least squares problems are generally described as follows: Give a function $\mathbf{f}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ with $m \geq n$. We want to find:

$$\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} \{\mathbf{F}(\mathbf{x})\} \quad (3.46)$$

where

$$\mathbf{F}(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m (f_i(\mathbf{x}))^2 = \frac{1}{2} \|\mathbf{f}(\mathbf{x})\|^2 = \frac{1}{2} \mathbf{f}(\mathbf{x})^\top \mathbf{f}(\mathbf{x}) \quad (3.47)$$

Nonlinear least squares problems can be solved by a number of optimization methods [117]. In this research, we use a hybrid method that presented by Madsen [118], which combines the Levenberg-Marquardt (L–M) algorithm with the Quasi–Newton algorithm.

3.8.2 Quasi-Newton Method

The basic Newton method is to find \mathbf{h}_n as the solutions to

$$\mathbf{F}''(\mathbf{x}) \mathbf{h}_n = -\mathbf{F}'(\mathbf{x}) \quad (3.48)$$

and compute the next iterate by

$$\mathbf{x}_{\text{new}} := \mathbf{x} + \mathbf{h}_n \quad (3.49)$$

The Quasi–Newton method is based on having an approximation \mathbf{B} to the Hessian $F''(x)$ at the current iterate \mathbf{x} , and the step \mathbf{h}_{qn} is found by solving

$$\mathbf{B}\mathbf{h}_{qn} = -F'(x) \quad (3.50)$$

This is an approximation to the Newton equation.

The approximation \mathbf{B} is updated by the BFGS strategy: Every \mathbf{B} in the series of approximation matrices is symmetric and positive definite. This ensures that \mathbf{h}_{qn} is “downhill”. We start with the symmetric, positive definite matrix $B_0 = I$, and the BFGS update consists of a rank 2 matrix to be added to the current \mathbf{B} . Madsen (1988) uses the following version, advocated by Al-Baali and Fletcher (1985),

$$\mathbf{h} := \mathbf{x}_{new} - \mathbf{x}; \quad \mathbf{y} := \mathbf{J}_{new}^T \mathbf{J}_{new} \mathbf{H} + (\mathbf{J}_{new} - \mathbf{J})^T f(\mathbf{x}_{new}) \quad (3.51)$$

If $\mathbf{h}^T \mathbf{y} > 0$

$$\mathbf{v} := \mathbf{B}\mathbf{h}; \quad \mathbf{B} := \mathbf{B} + \left(\frac{1}{\mathbf{h}^T \mathbf{y}} \mathbf{y}\right) \mathbf{y}^T - \left(\frac{1}{\mathbf{h}^T \mathbf{v}} \mathbf{v}\right) \mathbf{v}^T \quad (3.52)$$

With $\mathbf{J} = \mathbf{J}(\mathbf{x})$, $\mathbf{J}_{new} = \mathbf{J}(\mathbf{x}_{new})$. As mentioned, the current \mathbf{B} is positive definite, and it is changed only, if $\mathbf{h}^T \mathbf{y} > 0$. In this case it can be shown that also the new \mathbf{B} is positive definite.

3.8.3 Levenberg-Marquardt Method

Levenberg and later Marquardt suggested a damped Gauss-Newton method [117, 118]. The step \mathbf{h}_{lm} is defined by the followings:

$$(\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}) \mathbf{h}_{\text{im}} = -\mathbf{J}^T \mathbf{f}(\mathbf{x}) \quad (3.53)$$

where $\mathbf{J} \in \mathbf{R}^{m \times n}$ is the Jacobian, and μ is the non-negative damping parameter.

And compute the next iterate by

$$\mathbf{x}_{\text{new}} := \mathbf{x} + \mathbf{h}_{\text{im}} \quad (3.54)$$

The stopping criteria for the algorithm should reflect that at a global minimizer we have $\mathbf{F}'(\mathbf{x}^*) = \mathbf{g}(\mathbf{x}^*) = 0$, so we use

$$\|\mathbf{g}\|_{\infty} \leq \varepsilon_1 \quad (3.55)$$

where ε_1 is a small, positive arbitrarily chosen number. Another relevant criterion is to stop if the change in \mathbf{x} is small,

$$\|\mathbf{x}_{\text{new}} - \mathbf{x}\| \leq \varepsilon_2 (\|\mathbf{x}\| + \varepsilon_2) \quad (3.56)$$

This expression gives a gradual change from relative step size ε_2 where $\|\mathbf{x}\|$ is large to absolute step size ε_2^2 if \mathbf{x} is close to 0.

3.8.4 Hybrid Method

The Quasi–Newton method is not robust in the global stage of the iteration; it is not guaranteed to be descending. In 1988 Madsen presented a hybrid method which combines the L–M method (quadratic convergence if $\mathbf{F}(\mathbf{x}^*) = 0$, linear convergence otherwise) with a Quasi–Newton method. At the solution \mathbf{x}^* we have $\mathbf{F}'(\mathbf{x}^*) = 0$, and good final convergence is indicated by rapidly decreasing values of $\|\mathbf{F}(\mathbf{x}^*)\|$. If these norm values do not decrease rapidly enough, then we switch back to the L–M method,

and this gives super linear convergence, even if $F(x^*) \neq 0$. The iteration starts with a series of steps with the L-M method. If the performance indicates that $F(x^*)$ is significantly nonzero, then we switch to the Quasi-Newton method for better performance. It may happen that we get an indication that it is better to switch back to the L-M method, so there is also a mechanism for that.

The switch to the Quasi-Newton method is made if the condition

$$\|F'(x)\|_\infty < 0.02 * F(x) \quad (3.57)$$

is satisfied in three consecutive, successful iteration steps. This is interpreted as an indication that we are approaching x^* with $F'(x^*) = 0$ and $F(x^*)$ significantly nonzero, which can lead to slow, linear convergence.

The algorithm is summarized below. It calls the auxiliary functions *LMstep* and *QNstep*, implementing the two methods.

3.8.4.1 Hybrid Method Algorithm

$k := 0; x := x_0; \mu := \mu_0; B := I$

$\text{found} := (\|F'(x)\|_\infty \leq \varepsilon_1; \text{method} := \text{L - M}$

$\text{while } (\text{NOT found}) \text{ AND } (k < k_{\max})$

$k := k + 1$

case method of

L-M:

$[x_{\text{new}}; \text{found}; \text{better}; \text{method}; \dots] := \text{LMstep}(x; \dots)$

Q-N:

[xnew; found; better; method; ...]:= QNstep(x; B; ...)

Update B

if better

x := xnew

We have the following remarks:

1. Initialization. The stopping criteria are given by $\|F'(x)\|_{\infty} \leq \varepsilon_1$.
2. We also pass the current values of f and J etc, so that we do not have to recompute them for the same x .
3. Notice that both L-M and Quasi-Newton steps contribute information for the approximation of the Hessian matrix.

The two auxiliary functions are given below,

3.8.4.2 Levenberg–Marquardt Function

[xnew; found; better; method; ...]:= LMstep(x; ...)

xnew := x; method := L - M

Solve $(J(x)^T J(x) + \mu I)h_{lm} = -F'(x)$

if $\|h_{lm}\| \leq \varepsilon_2 (\|x\| + \varepsilon_2)$

found := true

else

$x_{new} := x + h_{lm}$

$\rho := (F(x) - F(x_{\text{new}})) / (L(0) - L(h_{\text{lm}}))$
 if $\rho > 0$
 better := true; found := ($\|F'(x_{\text{new}})\|_{\infty} \leq \varepsilon_1$)

if $\|F'(x_{\text{new}})\|_{\infty} < 0.02 * F(x_{\text{new}})$

count := count + 1

if count = 3

method := Q - N

else

count := 0

else

count := 0; better := false

3.8.4.3 Quasi–Newton Function

[xnew; found; better; method; ...]:= QNstep(x; B...)

$x_{\text{new}} := x$; method := Q - N; better := false

Solve $Bh_{qn} = -F'(x)$

if $\|h_{qn}\| \leq \varepsilon_2 (\|x\| + \varepsilon_2)$

found := true

else

if $\|h_{qn}\| > \Delta$

```


$$h_{qn} := (\Delta / \|h_{qn}\|) * h_{qn}$$


$$x_{new} := x + h_{qn};$$


$$\text{if } \|F'(x_{new})\|_\infty \leq \varepsilon_1$$


$$\quad \text{found} := \text{true}$$


$$\text{else}$$


$$\quad \text{better} := (F(x_{new}) < F(x)) \text{ OR } (F(x_{new}) \leq (1 + \delta)F(x))$$


$$\quad \text{AND } \|F'(x_{new})\|_\infty < \|F'(x)\|_\infty$$


$$\quad \text{if } \|F'(x_{new})\|_\infty \geq \|F'(x)\|_\infty$$


$$\quad \quad \text{method} := L - M$$


```

For the above functions *LMstep* and *QNstep*, please note that:

1. The gain ratio ρ is also used to update μ
2. Indication that it might be time to switch method. The parameter *count* is initialized to zero at the start of the algorithm.
3. We combine the Quasi–Newton method with a trust region approach, with a simple treatment of the case where the bound is active. At the switch from the

$L - M$ method Δ is initialized to $\max\{1.5\varepsilon_2(\|x\| + \varepsilon_2), \frac{1}{5}\|h_{lm}\|\}$.

4. In this part of the algorithm we focus on getting F' closer to zero, so we accept a slight increase in the value of F .
5. The gradients do not decrease fast enough.

3.8.5 Discrete Approximation

The methods discussed above assume that the vector function \mathbf{f} is differentiable, i.e. the Jacobian exists.

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_i}{\partial x_j} \end{bmatrix} \quad (3.58)$$

In my research project as well as many practical optimization problems it happens that we cannot give formulae for the elements in \mathbf{J} . The secant version is intended for problems of this type. The simplest remedy is to replace $\mathbf{J}(\mathbf{x})$ by a matrix \mathbf{B} obtained by *numerical differentiation*: The $(i, j)^{\text{th}}$ element is approximated by the finite difference approximation

$$\frac{\partial f_i}{\partial x_j}(x) \approx \frac{f_i(x + \delta e_j) - f_i(x)}{\delta} \equiv b_{ij} \quad (3.59)$$

where e_j is the unit vector in the j^{th} coordinate direction and δ is an appropriately small real number. With this strategy each iterate \mathbf{x} needs $n+1$ evaluations of \mathbf{f} , and since δ is probably much smaller than the distance $\|x - x^*\|$, we do not get much more information on the global behavior of \mathbf{f} than we would get from just evaluating $\mathbf{f}(\mathbf{x})$.

3.9 The Summary of the Training Process

In this project, we use 2D images to recognize person. There are two stages involved, the first stage is training, the second stage is recognition.

Assume that there is a set of 2D face images, $\{\Gamma_1, \Gamma_2, \dots, \Gamma_M\}$.

Each image has the size of $r \times c$.

- **Step 1a**

Normalize the input images.

Assume each image Γ_i is an $r \times c$ input image. Normalize it in terms of its intensity, using histogram equalization.

- **Step 1b**

Averaging the face images

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (3.60)$$

- **Step 2**

Calculate the difference images to remove average.

$$\Theta_i = \Gamma_i - \Psi, \text{ for } i = 1, 2, \dots, M \quad (3.61)$$

- **Step 3a**

Converting the $r \times c$ images into $p \times 1$ vectors where $p = r \times c$, using raster scan, which is a line-by-line sweep across the entire 2D matrix.

$$\Theta_{i(r \times c)} \rightarrow \mathbf{X}_{i(p \times 1)}, \text{ for } i = 1, 2, \dots, M \quad (3.62)$$

Where \mathbf{X}_i is a $p \times 1$ vector.

- **Step 3b.**

Stack M vectors into a matrix

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_M] \quad (3.63)$$

where \mathbf{X} is a $p \times M$ matrix.

- **Step 4.**

Compute the covariance matrix \mathbf{R} from \mathbf{X} .

$$\mathbf{R} = \frac{1}{M} \mathbf{X}^T \mathbf{X} \quad (3.64)$$

where \mathbf{R} is a $M \times M$ matrix.

- **Step 5.**

Find the eigenvalues and eigenvectors of the covariance matrix \mathbf{R} .

$$\mathbf{R} \cdot \mathbf{q}_i = \lambda_i \cdot \mathbf{q}_i, \text{ for } i = 1, 2, \dots, M \quad (3.65)$$

And we ordered the eigenvalues in the order of descending eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{M-1} \geq \lambda_M = 0 \quad (3.66)$$

- **Step 6.**

Create the eigenface matrix \mathbf{E}_s

$$\mathbf{E}_s = \mathbf{X} \mathbf{Q}_s \quad (3.67)$$

where \mathbf{Q}_s is the eigenvector matrix

$$\mathbf{Q}_s = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{M-1}] \quad (3.68)$$

\mathbf{E}_s is a $p \times M - 1$ matrix.

When we choose less number of eigenvectors in \mathbf{Q}_s , \mathbf{Q}_s still hold significant eigenvectors and \mathbf{E}_s holds significant eigenfaces.

- **Step 7.**

Normalization of eigenfaces

$$\mathbf{E}_n = \mathbf{E}_s \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & & & 0 \\ & \frac{1}{\sqrt{\lambda_2}} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sqrt{\lambda_M}} \end{bmatrix} \quad (3.69)$$

- **Step 8.**

Project the training difference image into eigenspace, which is spammed by

$$\hat{\mathbf{X}} = \mathbf{E}_n^T \mathbf{X} \quad (3.70)$$

where $\hat{\mathbf{X}}$ is a $M' \times M$ matrix.

Each $\hat{\mathbf{X}}_i$ is a point in the M' -th order eigenspace. It is the abstractive (parametric) representation of a person in that eigenspace.

3.10 The Summary of the Recognize Process

The task in this state is that given an image, try to find out if it is a face image, and if it is, then determine who this person is.

- **Step 1.**

Normalize the input image.

Assume image \mathbf{A} is an $r \times c$ input image. Normalize it in terms of its intensity, using histogram equalization.

- **Step 2.**

Compute the difference image to remove average.

$$\Theta_Y = \mathbf{A} - \Psi \quad (3.71)$$

- **Step 3.**

Converting the $r \times c$ image into $p \times 1$ vectors where $p = r \times c$, using raster scan, which is a line-by-line sweep across the entire 2D matrix.

$$\Theta_{Y(r \times c)} \rightarrow U_{(p \times 1)} \quad (3.72)$$

- **Step 4.**

Compute U 's eigenspace representation.

$$\widehat{U} = E_n^T U \quad (3.73)$$

where \widehat{U} is a $M' \times 1$ vector.

- **Step 5.**

Compare \widehat{U} with \widehat{x}_i by computing the vector distance.

$$e_i = \|\widehat{U} - \widehat{X}_i\|, \text{ for } i = 1, 2, \dots, M \quad (3.74)$$

There are 3 cases of the distance e_i :

- **Case 1.**

e_i is greater than the face threshold. -- This means the input image is not a face.

- **Case 2.**

e_i is smaller than the face threshold, but still greater than some reorganization threshold. -- This means the input image is a face image, but not in the database so that we can't recognize the person.

- **Case 3.**

e_i is smaller than the reorganization threshold. -- This means we recognize the image. Then the one with smallest e_i is the recognized person.

3.11 Summary

In this chapter the preliminary materials essential to our research are briefly reviewed. The implementation of these algorithms provides us the way to handle 3D images and to achieve the challenging objective of recognizing 3D faces.

Chapter 4

THE PROPOSED 2D/3D RECOGNITION APPROACHES

4.1 Introduction

In this chapter, we focus on the 2D/3D face recognition problem, which is, the probing face image is two-dimensional (2D) while the gallery face database is three-dimensional (3D).

Nowadays, it's not too difficult to obtain 3D models during the enrollment stage. However, it could become problematic to use 3D equipment for verification, as the method could be time consuming, computationally expensive, and impractical in implementation, especially for the purpose of identity verification.

Our research proposes two approaches for the 2D/3D face recognition:

- **3D Shape Assisted 2D Recognition**

Rotating the 3D gallery faces to the same pose angle as the 2D probing face, and applying a similar illumination condition. Then the 3D gallery faces are projected to the 2D image plane. The final recognition procedure is done in the 2D image plane. In the later discussion, we also call this method as the “downgrade method”, since the 3D image is downgraded to the 2D image.

- **2D Assisted 3D Recognition**

Creating a 3D face model by using a single 2D face image. When a 2D probing face image is presented to the system, a 3D face image, which starts with an average 3D face image derived from the 3D face database, is projected onto the 2D image plane, after rotation, translation, scaling, and interpolation. An optimization algorithm is then applied to minimize an error index, which is a function of the difference between the projected and the input images, by selecting the 3D depth and intensity parameters. Once the algorithm converges, the resulting 3D depth and intensity parameters can be employed to construct a 3D face model of the subject photographed in the 2D image. We also call this method as the “upgrade method”, since the 2D image is upgraded to the 3D image.

4.2 System Setup

The 3D camera system, developed and patented by Genex Technologies, Inc., consists of a 3D facecam camera that takes 3D pictures, a digital camera that takes 2D photos, a 3D computer platform with integrated frame grabber board, and a video card and capture software that controls the operation of the 3D camera and allows viewing, editing, and saving of 3D pictures.

The 3D camera system is a 3D surface profile measurement system capable of acquiring full frame dynamic 3D images of objects with complex surface geometry at a high speed. The (x, y, z) coordinates for all visible points on the object surface are provided by a single 3D image. The three-sensor system captures over 300,000 data points of geometric and intensity information. It enables capture of intensity information

in full color, or as black-and-white intensity. A snapshot of the camera system at the FAU DSP Laboratory is given in figure 4.1.



Figure 4.1: A Snapshot of the 3D Camera System in the FAU DSP Lab

With the camera system described above, a 3D face model database was created. About 50 persons with different age, sex, and race were modeled. Each subject took nine 3D pictures of various poses under different lighting conditions. In addition, 2D digital pictures of each subject were also taken. These 3D and 2D pictures were the basis for our experiments.

The 3D face image includes both 3D shape information and 2D intensity information. These data are processed separately, and are both used in the research.

4.3 3D Shape Assisted 2D Recognition

4.3.1 Solution Method

Facial biometrics can utilize 3D reconstruction of faces in two contexts:

- **Enrollment**

Offline gallery database capture of individuals for use in training and as exemplars for verification and recognition.

- **Identification**

Online probing face capture for identity recognition or verification.

For this approach, in the enrolment stage the 3D images are taken to form a gallery database; in the identification stage, only 2D images are needed for probing.

To create a robustness system to recognize any subject given a 2D probe image, it is necessary to compensate for the variation of pose and illumination differences prior to face recognition. Given the 2D probe image, the 2D/3D pose determination algorithm is first applied to estimate the pose of the probe image, and then the 3D gallery images are rotated to the same pose. Illumination compensation is then applied to the 3D gallery images to simulate the same light environment of the 2D probe image. They are then projected to the 2D image plane, and the 2D PCA algorithm (or other feature extraction procedure) is performed to transform the 2D projected image to another parameter space, on which each projected gallery images is compared with the 2D probe image.

The proposed enrolment method includes the pre-processing and 3D model creation. The identification method is thus based on the following three building blocks: pose determination, illumination compensation, and optimization, which will be discussed in the next section. The overall face recognition diagram can be simplified and described as is shown in figure 4.2.

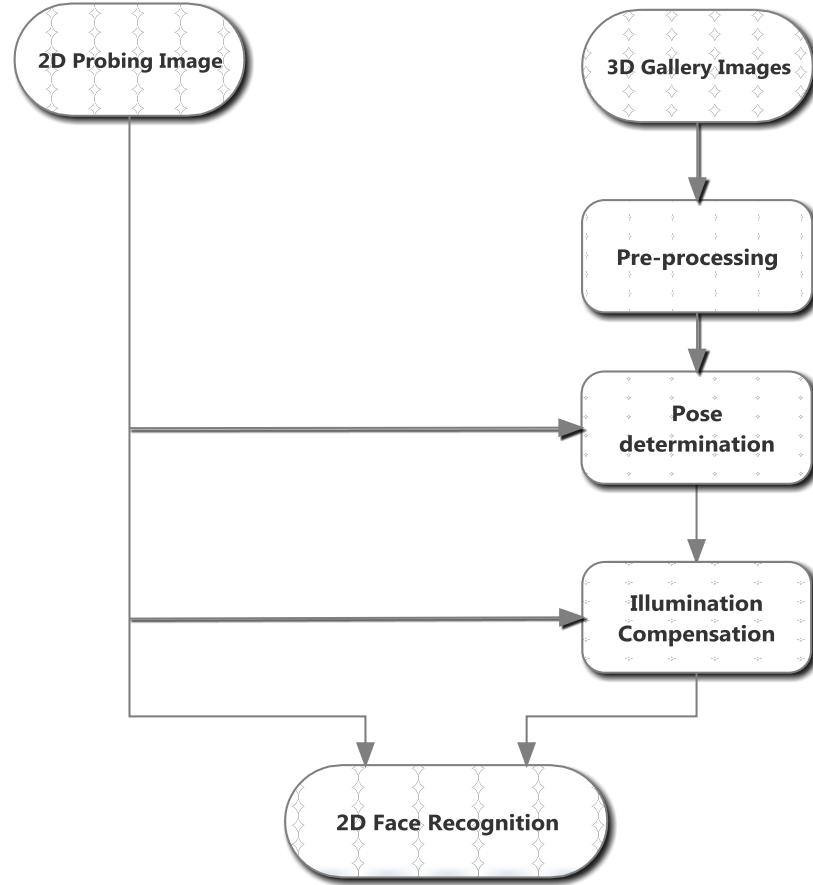


Figure 4.2: Face Recognition Diagram

In the following experiments, with the camera system described above in Section 4.2, a 3D face model database was created. About 46 persons with different ages, sexes, and races were modeled. For each subject, nine 3D pictures were taken from various poses under different lighting conditions. Meanwhile, 2D digital pictures of each subject were also taken. These 3D and 2D pictures were the basis for our experiments.

In the first experiment, we implemented the scheme presented in Chapter 3 without pose and illumination compensations. In the second experiment, the 3D gallery images were rotated to the same pose as the 2D probe image. And in the last experiment,

both pose and illumination compensations were applied to examine the robustness of the proposed method.

4.3.2 Pre-Processing

The detailed pre-processing procedure is described in Chapter 3. Figure 4.3(a) shows an example of the depth image, and figure 4.3(b) shows the corresponding intensity image.

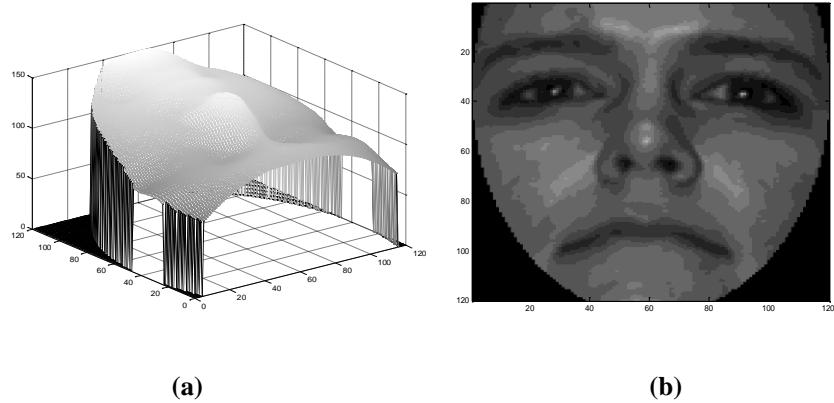


Figure 4.3: (a) Example of 2D Depth Image, and (b) Example of 2D Intensity Image

4.3.3 Pose Acquisition

Figure 4.4(a) is the 2D fronto-parallel projection of the 3D intensity image, figure 1(b) is the test 2D image. We applied the pose determination algorithm outlined in Section 3.1 to calculate the pose of the 2D probe image. We then rotated the 3D gallery image to the computed pose and projected it to the 2D image plane. The resulting 2D image is shown figure 4.4(c).

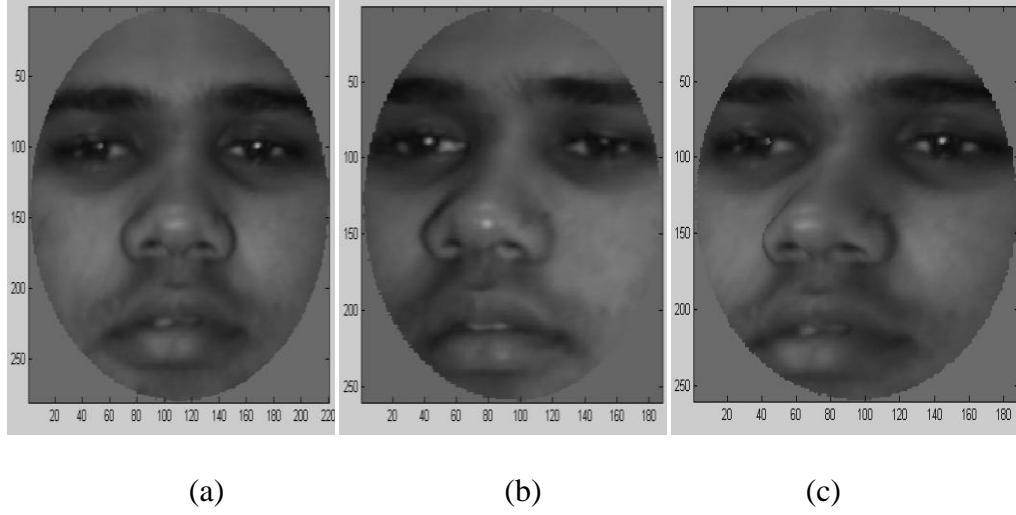


Figure 4.4: Sample Image Rotation. (a) The Frontal View from the 3D Gallery Image, (b) The Original 2D Probe Image, and (c) The Projection of the 3D Gallery Image after Pose Compensation

Comparing Figure 4.4(b) and 4.4(c), one observes that the two have almost identical viewing angles. This will contribute to a better performance for 2D face recognition.

4.3.4 Illumination Compensation

The Phong reflection model and the Z buffer rendering method described in Section 3.2 were used in this experiment. Due to the nature of human faces, based on the experimental studies, the coefficients were chosen as follows: the ambient reflection coefficient was set to 0.1, the diffuse reflection coefficient 0.1, the specular reflection coefficient 1.0, the spectral exponent 5.0, and the specular color reflectance 0.5.

For the number of light sources, the experimental investigation revealed that the result improved when adding new light sources from 1 to 3, and slowed dramatically the improvement when more light sources were added. The intuitive explanation for this

result is that 3 light sources are sufficient to describe any illumination condition for the face – one from above, one from left and one from right. Therefore three light sources were chosen in our experiments.

Figure 4.5(a) shows a 3D face image without any illumination compensation, and 4.5(b) shows that compensated utilizing 3 lights with the Phong reflection model.

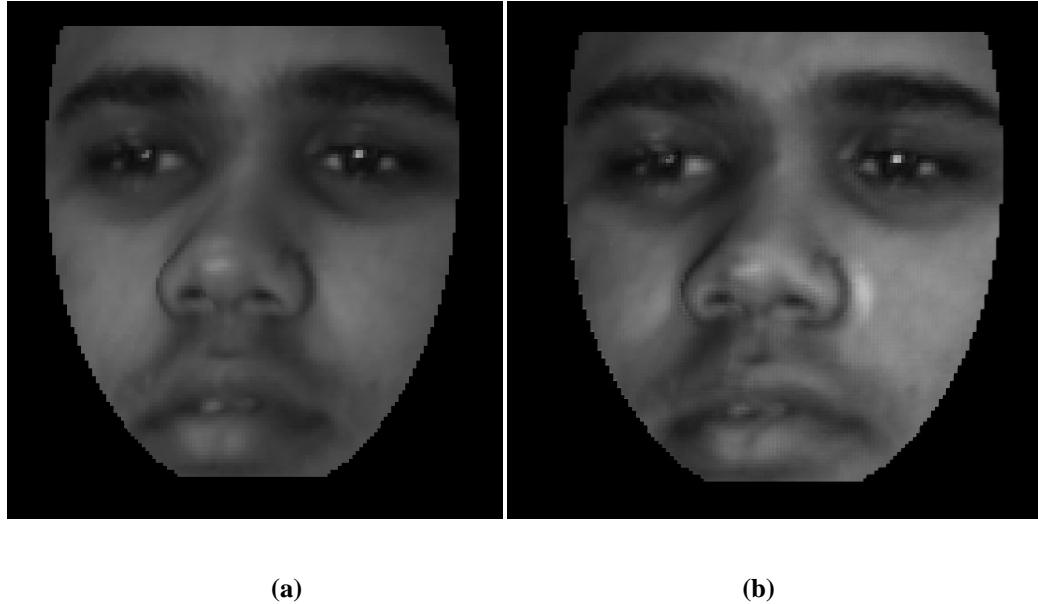


Figure 4.5: Illumination Compensation Example. (a) No Compensation, and (b) Compensation with 3 Light Sources

4.3.5 Optimization Method

Nonlinear least squares problems are generally described as follows: Give a function $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ with $m \geq n$. We want to find:

$$\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} \{F(\mathbf{x})\} \quad (4.1)$$

where

$$\mathbf{F}(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^m (f_i(\mathbf{x}))^2 = \frac{1}{2} \|\mathbf{f}(\mathbf{x})\|^2 = \frac{1}{2} \mathbf{f}(\mathbf{x})^\top \mathbf{f}(\mathbf{x}) \quad (4.2)$$

Nonlinear least squares problems can be solved by a number of optimization methods [117]. In this research, we use a hybrid method that presented by Madsen [118], which combines the Levenberg-Marquardt (L–M) algorithm with the Quasi–Newton algorithm. The general L–M method gives quadratic convergence if $\mathbf{F}(\mathbf{x}^*) = 0$, however, it gives linear convergence otherwise. And the general Quasi–Newton method, gives super linear convergence even if $\mathbf{F}(\mathbf{x}^*) \neq 0$. The iteration starts with a series of steps with the L–M method. If the performance indicates that $\mathbf{F}(\mathbf{x}^*)$ is significantly nonzero, as described at the above, this can lead to slow, linear convergence, in which case we switch to the Quasi–Newton method for a better performance. This switch is made if the condition [118]

$$\|\mathbf{F}'(\mathbf{x})\|_\infty < 0.02 \times \mathbf{F}(\mathbf{x}) \quad (4.3)$$

is satisfied in three consecutive, successful iteration steps. This is interpreted as an indication that we are approaching \mathbf{x}^* with $\mathbf{F}'(\mathbf{x}^*) = 0$ and $\mathbf{F}(\mathbf{x}^*)$ significantly nonzero.

Both the Quasi–Newton method and the L–M method are briefly summarized in Chapter 3.

4.3.6 2D PCA for Dimensional Reduction and Parameter Extraction

Two PCAs, depth and intensity, are employed here.

For the depth PCA, we need to use depth measurements from 3D image data. It has been mentioned that for 3D images, one has a matrix of points given below in equation 4.4.

$$\begin{bmatrix} x_1 & y_1 & z_1 & t_1 \\ x_2 & y_2 & z_2 & t_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_n & y_n & z_n & t_n \end{bmatrix} \quad (4.4)$$

In (4), z_i , $i = 1, 2, \dots, n$, is the depth corresponding to coordinates x_i and y_i and t_i , $i = 1, 2, \dots, n$, is the intensity corresponding to coordinate x_i, y_i and z_i . The data given in equation 4.4 represents uniquely a 3D image of a subject. In the depth PCA, we only need to use the data given in the following matrix:

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_n & y_n & z_n \end{bmatrix} \quad (4.5)$$

Let ξ_i be a set of depth values (z_1, z_2, \dots, z_n) of the i th subject. Then, $\xi_1, \xi_2, \dots, \xi_k$ are the depth vectors of the entire training set consisting k subjects. The average depth is given by:

$$\bar{\xi} = \frac{1}{k} \sum_{i=1}^k \xi_i \quad (4.6)$$

The difference from average depth for each subject is as follows:

$$\varphi_i = \xi_i - \bar{\xi} \quad i = 1, 2, \dots, k \quad (4.7)$$

The covariance matrix is defined as:

$$R_z = \frac{1}{k} \sum_{i=1}^k \boldsymbol{\varphi}_i \boldsymbol{\varphi}_i^T = DD^T \quad (4.8)$$

where $D = [\varphi_1 \ \varphi_2 \ \dots \ \varphi_k]$ and the size of D is $n \times k$. The R_z matrix is of size $n \times n$ which is large and leads to computationally complex calculations for eigenvalues and eigenvectors. An alternative approach is to use $\hat{R}_z = D^T D$.

The eigenvalues and eigenvectors of \hat{R}_z are determined by:

$$\hat{R}_z \mathbf{q}_{z,i} = \lambda_{z,i} \mathbf{q}_{z,i} \quad (4.9)$$

Where $\lambda_{z,i}$ and $\mathbf{q}_{z,i}$ are the i th eigenvalue and eigenvector, respectively. Let $Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_k]$. Note that the eigenvalues are placed in a descending order. The eigenvectors are subsequently normalized. And these are related to the eigenvectors of DD^T

$$E = DQ \quad (4.10)$$

The columns of E span the so-called depth eigenspace. The next step is to project all depth vectors to the depth eigenspace, by

$$\mathbf{z}_i^f = E^T \boldsymbol{\varphi}_i \quad (4.11)$$

and the size of \mathbf{z}_i^f depends on the number of eigenvectors selected. To recreate the depth map, we can calculate

$$\hat{\mathbf{z}}_i = (E\mathbf{z}_i^f) + \bar{\xi} \quad (4.12)$$

where $\hat{\mathbf{z}}_i$ is the recreated depth vector for the i th subject.

The procedure for the intensity PCA is quite similar to that of the depth PCA. For the intensity PCA, the data from the following matrix is employed:

$$\begin{bmatrix} x_1 & y_1 & t_1 \\ x_2 & y_2 & t_2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_n & y_n & t_n \end{bmatrix} \quad (4.13)$$

4.4 2D Assisted 3D Recognition

4.4.1 Solution Strategy

As discussed in the previous sections, the automatic authentication system is normally made up of an enrollment stage and a verification stage. In the enrollment stage, individuals who are supposed to have the authority of accessing are registered with their 3D photos. In the verification stage, only 2D face images are captured and compared with the 3D images stored in the gallery database for authentication. A 3D face model is created by selecting weights of 3D depth and intensity such that the projected 2D image of the 3D face is similar to the 2D probing image. Only the 3D weights are needed to be saved in the database, then the general feature extraction procedure is applied to compare the probing image's 3D parameters with those saved in the database to determine the identity of the person.

4.4.2 Enrollment Stage

- . In this modeling stage, a number of 3D face images are taken and their features are extracted and stored in the database. In the rest of the enrollment stage, a 3D

photo of a subject is obtained. And the 3D weights (or 3D parameters) are then saved in the database.

4.4.3 Verification Stage

The proposed scheme for the verification stage is shown in figure 4.6. The input to the procedure is a 2D probing face photo. Initially, a 3D neutral face is created from the average 3D face derived from the gallery database. A pose angle determination procedure is applied to determine the viewing angle of the 2D probing face approximately, and the initial 3D face is then rotated to align with the probing face. Afterwards, an illumination compensation algorithm is then applied to the 3D face so that its 2D projection has the similar intensity profile as the 2D probing image. The 3D synthetic face image is then projected to the 2D image plane to compare with the probing photo in the 2D parametric space. The discrepancy between the two images will lead to another iteration in which a new set of 3D parameters are generated in an effort to reduce the discrepancy. The iterative process continues until the algorithm converges or the number of iterations has exceeded a preset value. These sets of parameters are then used to represent the 3D face model which resembles the subject in the 2D probing photo.

As has been mentioned, the set of 3D face parameter vectors created in the enrollment stage and stored in a database is utilized in the verification stage. The verification stage process is to create a 3D face, project the face onto the 2D image plane, and then compare it with the 2D input image of a subject in question in the 2D parameter space. The iterative process is similar to the one described in the enrollment stage. As the iterative process proceeds, if the parameters of the projected and the input images are

close enough, then the corresponding 3D depth and intensity parameter vectors are used to compare with those stored in the database and an authentication decision can be reached based on a set of preset decision rules.

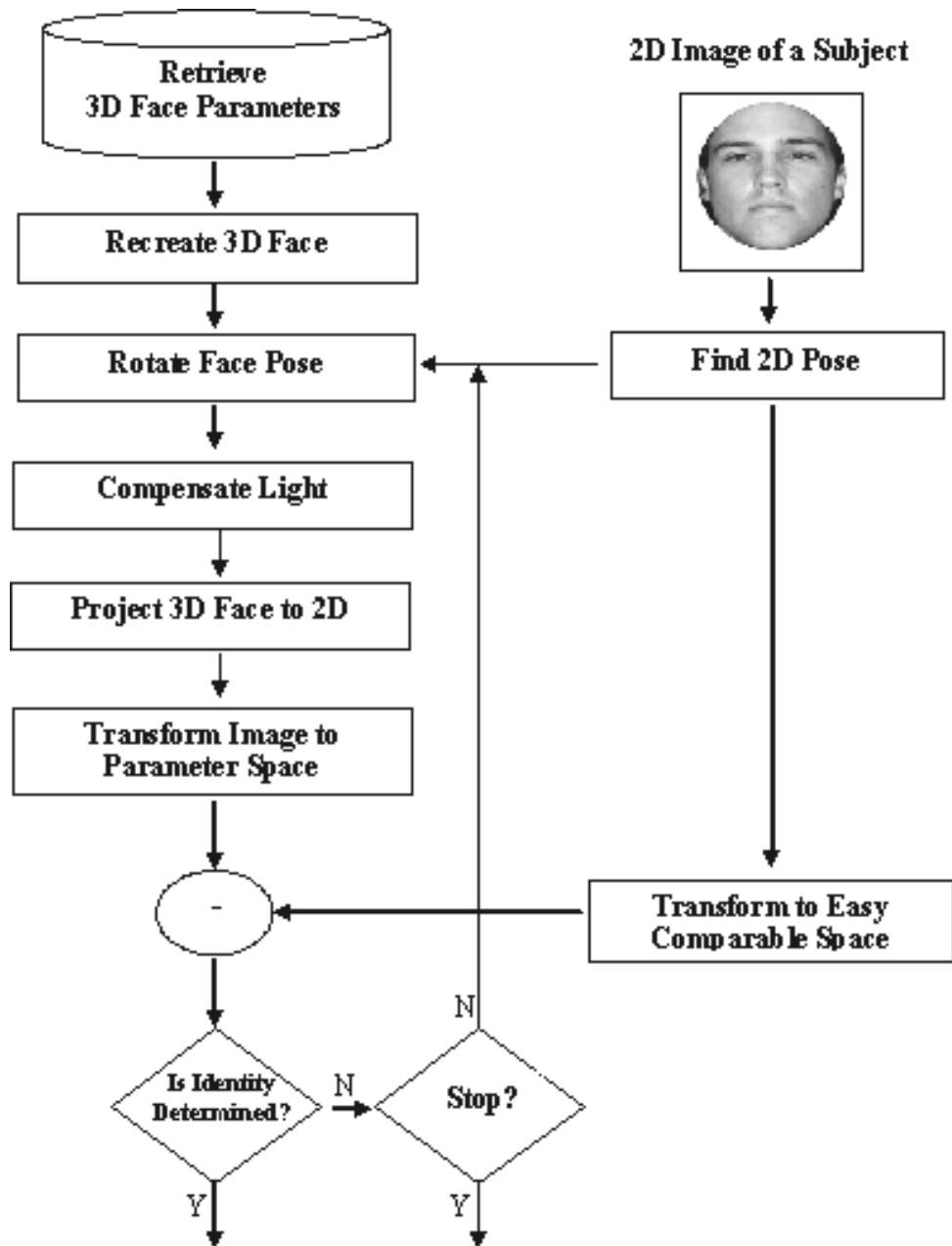


Figure 4.6: Verification Stage

4.4.4 3D Face Recreation

In this section, it is assumed that the 3D face image of the person to be enrolled is not available in the 3D model database. It is further assumed that up to r number of photos of the subject to be enrolled is presented to the system. In this case, all the intensity images including those given for the verification of the subject will be used in 3D face recreation.

$$\mathbf{z}^f = \sum_{i=1}^k a_i \mathbf{z}_i^f, \quad \mathbf{t}^f = \sum_{i=1}^{k+r} b_i \mathbf{t}_i^f \quad (4.14)$$

The diagram given in figure 4.7 is modified slightly for creating a 3D face given a number of 2D face images. The technical details given in above still apply here.

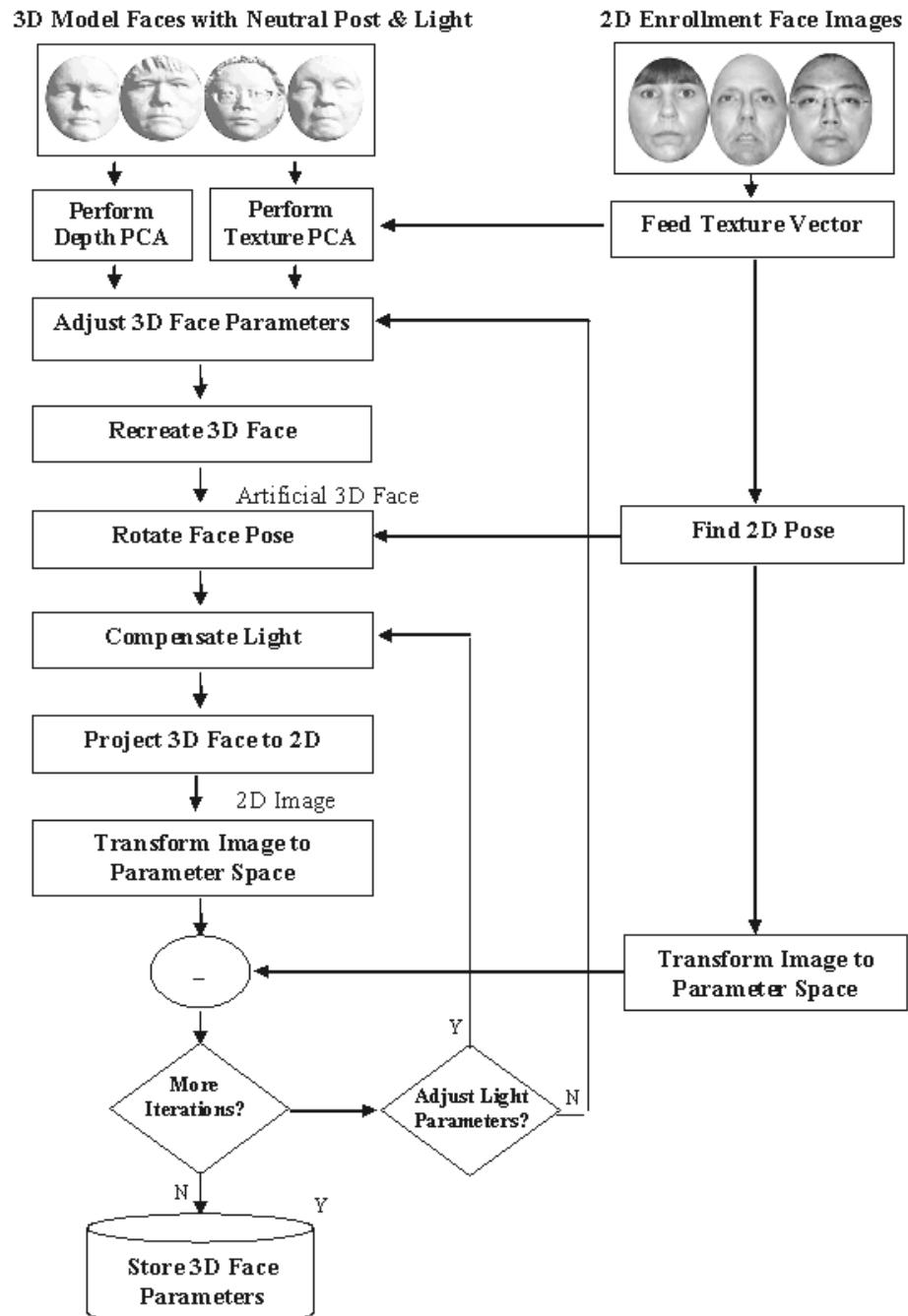


Figure 4.7: 3D Face Recreation

4.5 Summary

In this chapter we discussed two methods for 2D/3D face recognition. The first method uses the 3D shape information to project the gallery images to the 2D plane and then applies 2D recognition algorithm; and the second method creates 3D face models on 2D probing images and then applies 3D recognition algorithms. The experiment results are discussed in the next chapter.

Chapter 5

EXPERIMENT CONFIGURATIONS

5.1 Introduction

In the previous chapters, some solution methods and detailed implementation strategies for 2D/3D face recognition were proposed. To verify the effectiveness of the proposed methods, recognition experiments have been conducted. In this chapter, the 3D camera system used in our experimental studies will be presented first. To demonstrate our experiment process, the individual tests we ran and the result we collected will also be shown.

5.2 Experiment Setup

As briefly introduced in Chapter 4, the 3D camera system we used in our experiments are developed and patented by Genex Technologies, Inc. We setup the test following the requirement from Genex:

Environment:

- Ambient light should be between 50 lux and 300 lux on the target (normal average lighting), with consistent and uniform ($\pm 5\%$) levels around the immediate subject capture area. Most normal office lighting will not be a problem.

- Light conditions around the immediate subject capture area should remain the same during all times of day. A closed room can ensure this, as can heavy blinds.
- We should also pay attention to other factors such as light bulbs, and ensure that a consistent number and brightness of bulbs are present in the room.
- The camera can perform well in many environments, and it has settings allowing adjustment of the projector's brightness, but it is important to maintain the above conditions to ensure consistency.

Camera Backdrop:

- A medium-toned, flat color cloth or other non-reflective backdrop should be used. Suggested colors include: brown, tan, grey and avoid overly light or dark background.
- The backdrop should be uniform and flat.
- The backdrop should be at least as wide and tall as the camera's capture area (20" Wide x 16" Tall).

Camera Position:

- Place system 115 cm from the backdrop. Measure this distance from the front of the camera.
- Mark the floor where the system is placed, to ensure that it can be returned in case of accidental moving.

Camera Configuration:

- Adjust the camera to the same height and slightly (define) below the subject's face.
- If positioned properly, you can see the ears equally in the center sensor and the camera has a perspective on the bottom of the nose (where it touches the face).

Each subject has to sign an agreement form so that we can use their images or 3D images in our database for research purpose. The brief procedure to take 3D pictures of subjects is given in figure 5.1.

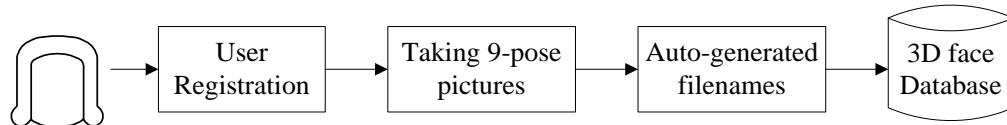


Figure 5.1: Database Construction Flowchart

With the procedure in figure 5.1, a 3D face model database was created. It currently consists of 3D face images of 50 subjects. For each subject, nine 3D pictures of various poses from different angles and with varying lighting conditions are taken. Future work will include expanding the database with more subjects with different ethnic backgrounds.

The 3D picture taking procedure is illustrated in figure 5.2. There are two sets of cameras, the 2D intensity camera is just a commercial digital camera which captures the 2D face photo, and the 3D shape cameras capture the shape information in a points cloud format, with .stl file extension.

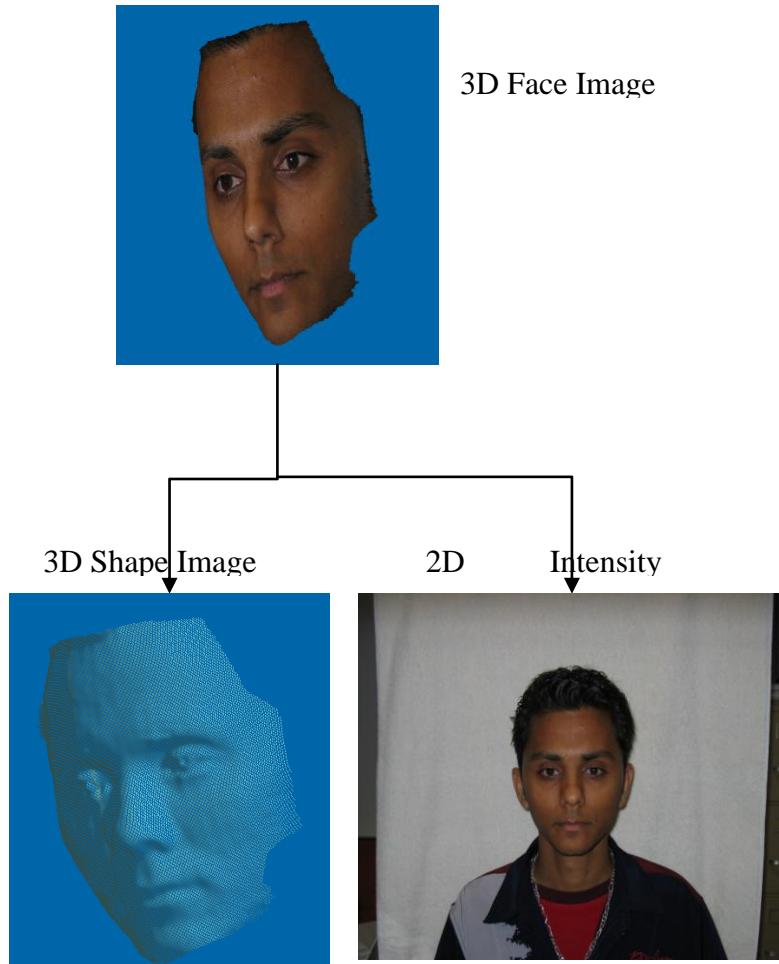


Figure 5.2: Raw 3D data

5.3 Raw 3D Shape Data

The original raw 3D shape data, in the .stl format, are first rotated and transformed to the same alignment, to make sure the corner of each eye lies on the same x-y plane. The data is then scaled so that the distance between the corners of two eyes is the same for all the images. The unused area, such as the hair and leave the data in the ellipse shape, is then cropped. In figure 5.3, it shows a sample raw 3D shape image, from the top view, front view and the right view, as long as the prospective view on the top left corner. And figure 5.4 shows the same data after the pre-processing.

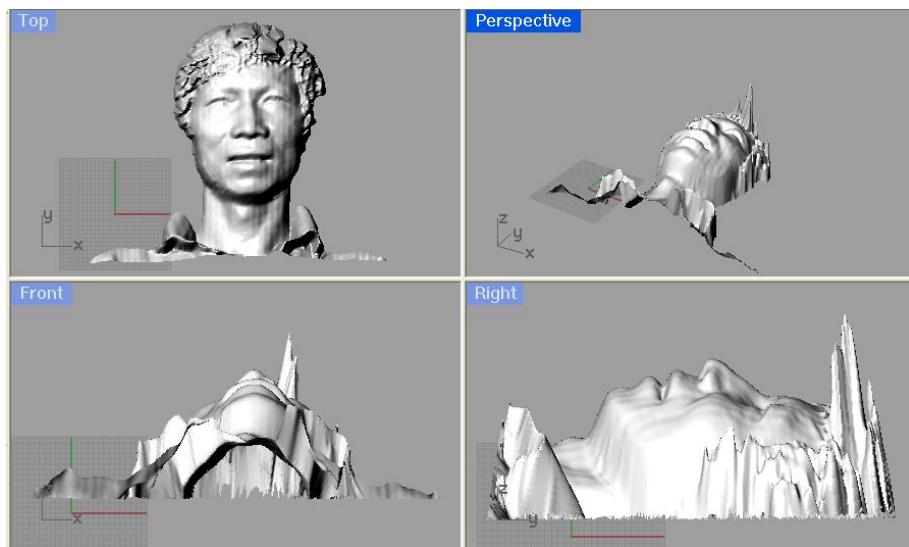


Figure 5.3: Original 3D Shape Data

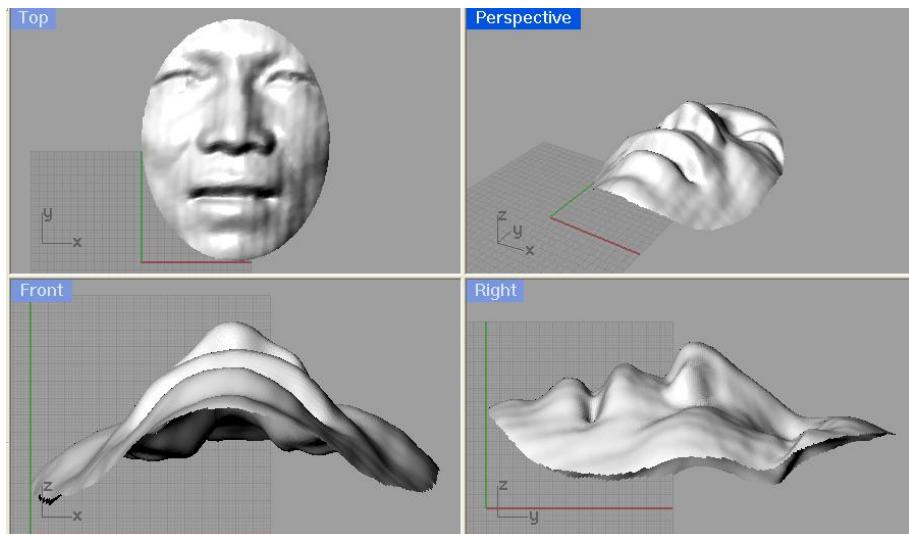


Figure 5.4: Data after Transformation, Rotation, Scaling, and Crop

At the next step, the Finite Element Methods are applied on shape data to create the data grid of the point cloud. NURBS (non-uniform rational B-spline) interpolation is

applied to mesh the 3D surface. Figure 5.5 shows the same sample data after the re-generation using NURBS.



Figure 5.5: The 3D Shape Image Re-generated by the NURBS.

Figure 5.6 to 5.8 shows another example of the test 3D data. And in figure 5.8 it demonstrates the re-generated data in the point clouds format.

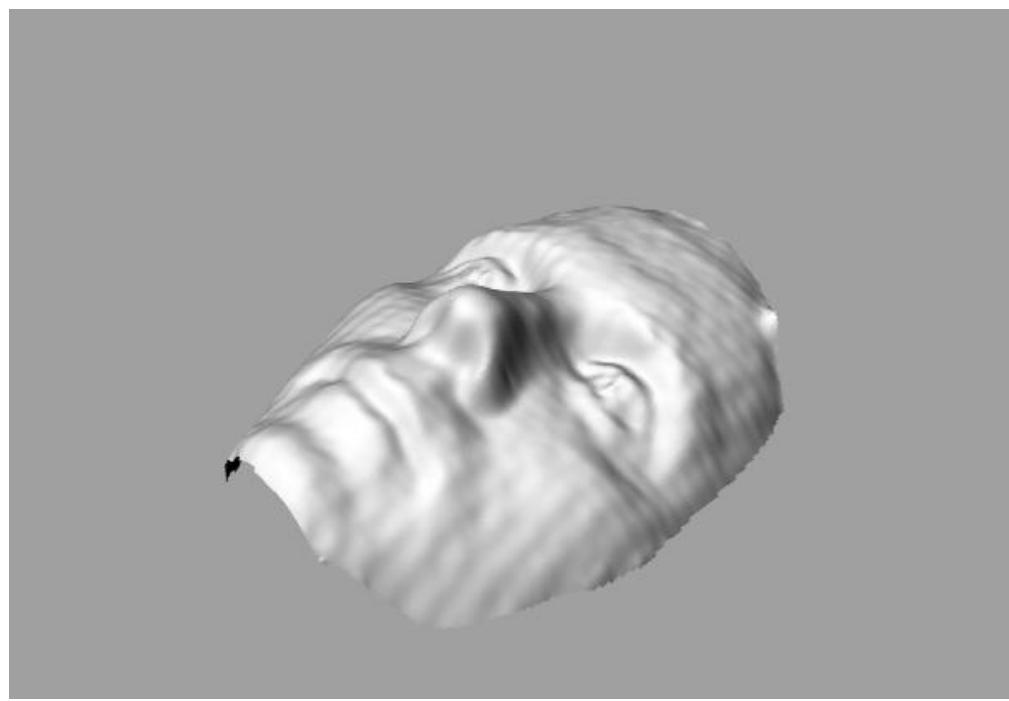


Figure 5.6: Original 3D Shape Image (amit-20041028-1.stl)

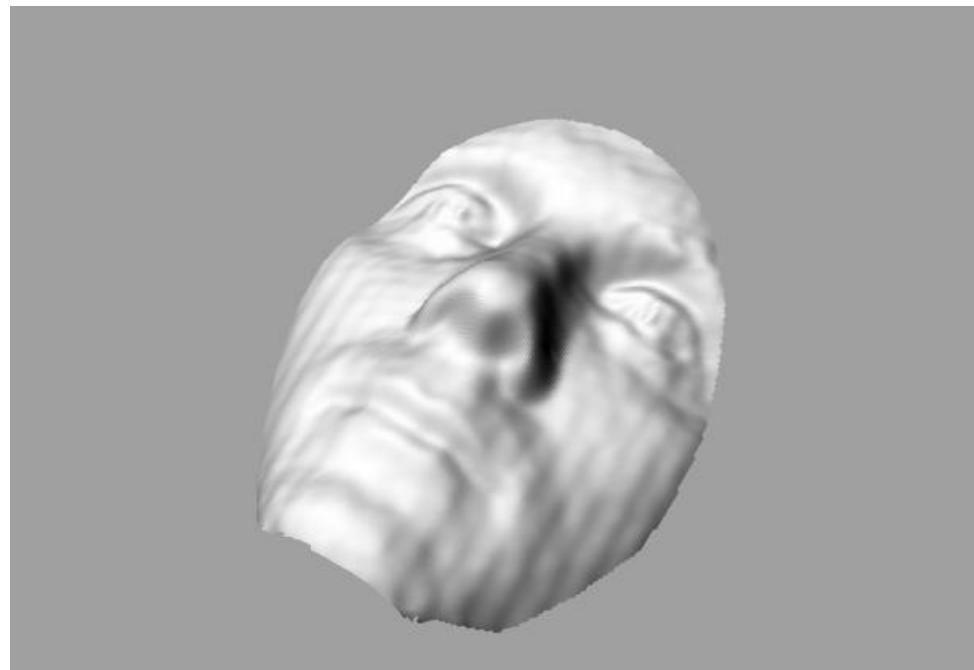


Figure 5.7: After the Pre-processing (amit-20041028-1-output.stl)

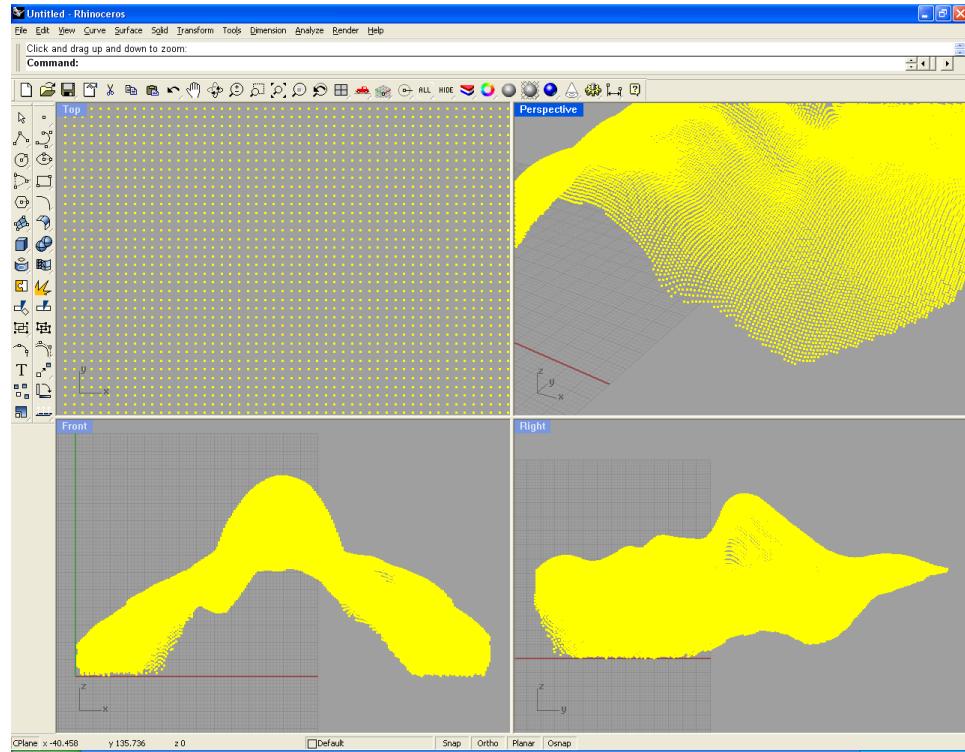


Figure 5.8: The Generated Point Clouds (temp.txt)

5.4 Raw 2D Intensity Data

The same pre-processing methods are applied to the raw intensity data as well. It transforms and rotates the image, and save the 2D intensity information. In Figure 5.9, it shows the raw 2D intensity data as a normal face photo. It then is converted to the grayscale image in Figure 5.10. Figure 5.11 shows the modified image after the scaling and alignment.



Figure 5.9: Original High Definition Intensity Image

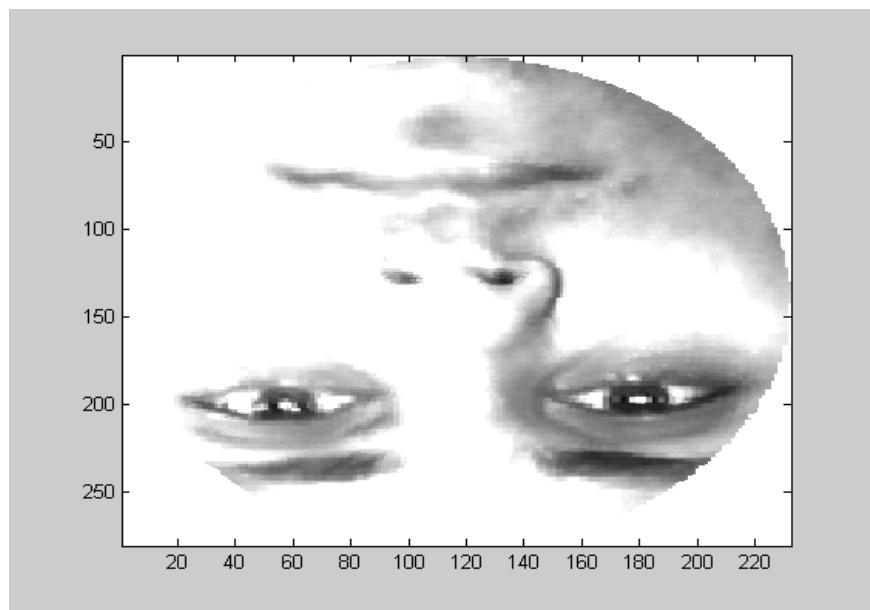


Figure 5.10: Original Grayscale 2D Intensity Image



Figure 5.11: Generated Grayscale Image

5.5 Model Fitting

As has been mentioned in Chapter 4, in the 3D face database we constructed, there were pictures of 50 subjects, and we took pictures of each subject from 9 poses. The individual subject's 2D photos were used as the input images for 3D face image recreation. The initial parameter vector was obtained using an average face at the front view shown in figure 5.12. Partial results of recreated 3D faces with the proposed scheme are shown in figure 5.13. In the figure, the first row depicts the input images of the subjects, and the remaining rows, the recreated 3D images from various viewing angles.

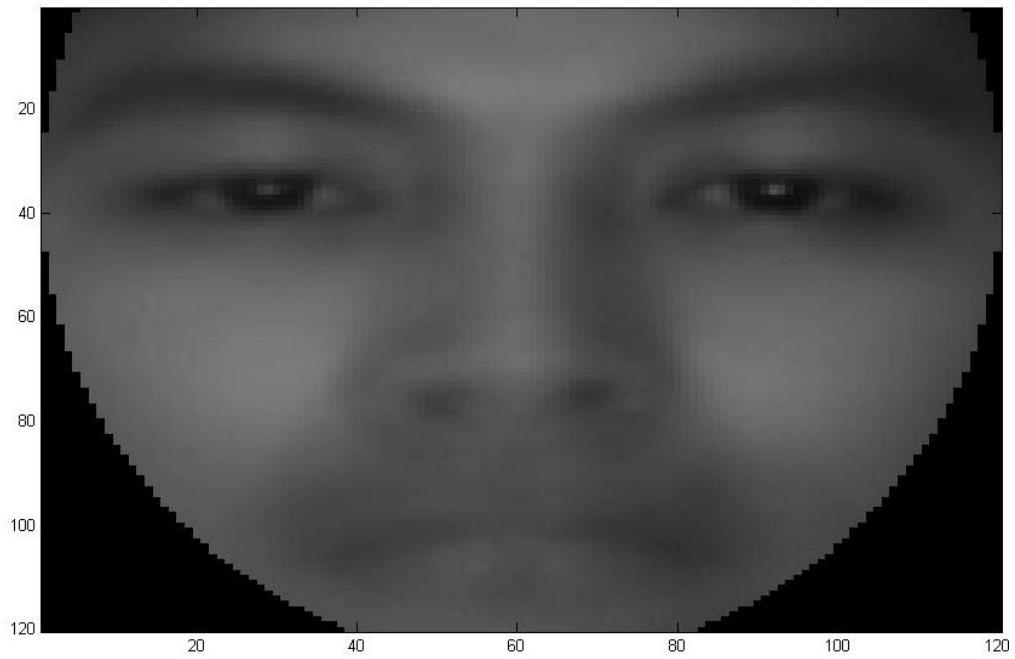


Figure 5.12: Average Face



Figure 5.13: Recreated 3D Faces from 2D Images

Shown in figure 5.14 are the recreated 3D images at the first three iterations.

Large changes can be easily detected after only a small number of iterations. After this

small number of iterations, subsequent changes are usually too small to be observed by the human eye.



Figure 5.14: First Three Iterations of Recreated 3D Faces

5.6 Summary

It is very important to setup the system properly and create the original 3D face database correctly. These factors will affect the later face recognition experiments. From the system configuration point of view, one needs to first understand the detailed specification of the 3D face image taking procedure. Then one needs to analyze the raw data that the system captured, and use third party applications to further analyze and parse the raw data. In our studies, we also wrote our own programs to pre-process the data using C programming language. To create a 3D face database, digital image processing techniques were applied. These procedures and processes are guidelines for our 3D face recognition goal. In the next chapter, the experimental results with the proposed methods will be presented.

Chapter 6

2D/3D FACE RECOGNITION EXPERIEMNT RESULTS

6.1 Introduction

In this chapter the experimental results with our approach of using 2D probing images and 3D face models will be shown. As mentioned in Chapter 5, two different solution methods are used upon the same data. In the first method, the 3D gallery images are projected to the 2D plane and the recognition is done in 2D plane; In the second method, 3D face model is created for each probing 2D photo and then the recognition is done in the 3D domain.

6.2 3D Shape Assisted 2D Recognition

6.2.1 Introduction

In this approach, 3D gallery images are projected to the 2D plane. The step by step experiments start with simple 2D PCA without any compensation, and then we applied the pose angle compensation, and finally the pose and illumination compensation both. The detailed information is described below.

For the experiments in this section, we use the pictures of 46 subjects from our 3D gallery face database, there were, and each subject had 9 3D pictures from various

viewing angles, expression and illumination conditions. Each subject's 2D pose angle view pictures were also taken and used as the probe image. Three experiments were conducted: In the first experiment, each 3D gallery image was simply projected to the 2D image plane and the 2D PCA recognition algorithm was performed to see how well the algorithm could recognize the 2D pose angle view probe image. In the second experiment, we first applied the 3D-2D pose determination algorithm described in Chapter 3, to estimate the pose angle of the probe image, and then all the 3D gallery images are rotated to the same pose angle followed by projecting to the 2D image plane. And in the third experiment, we further compensate the illumination of the gallery images by applying the algorithm described before.

6.2.2 Experiment with Simple 2D PCA Face Recognition

Figure 6.1 shows an example of 2D probing images in a semi-profile view and the projection of its corresponding 3D gallery image in the fronto-parallel view. In this example we showed the picture of the same person, but the variation of pose angle makes them looks quite different for computers.

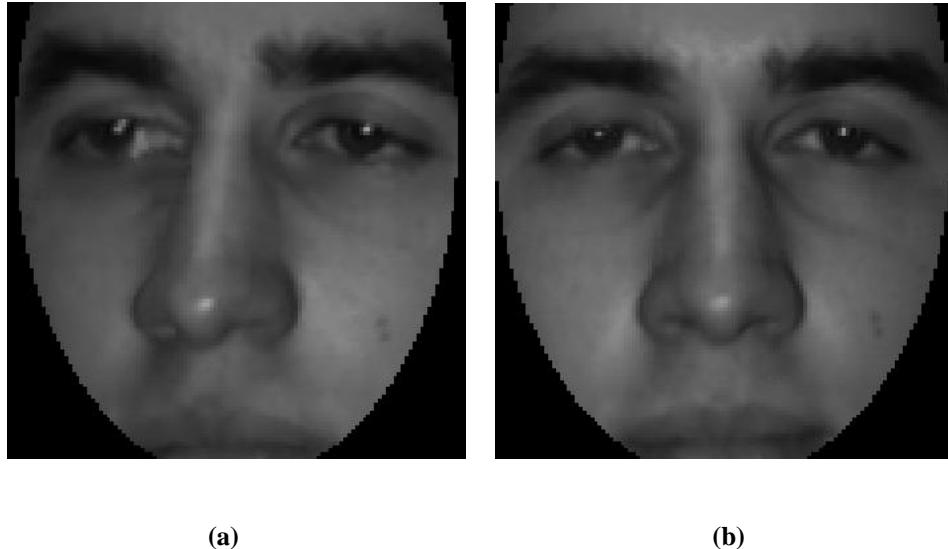


Figure 6.1: Example of 2D Probing Image
(a) A 2D Probe Image
(b) Projection of a Gallery Image from the Fronto-parallel View

For each 2D probing image, the PCA algorithm is applied on each of the 3D gallery images. And the test is repeated also for each of the probe images.

To judge the performance of the face recognition algorithm, an error is defined as the Euclidean norm of the (PCA feature) vector difference between a probe image and a projected gallery image. Figure 6.2 shows the errors generated by testing all probe images. Note that the PCA feature vector of each probe image is compared with those of all 46 gallery images. For the recognized faces, the mean vector distance is 16.19 and the standard deviation is 3.45; for the other faces, the vector distance mean is 27.27 and the standard deviation is 4.13. And the recognition rate is 91.30% since 4 out of 46 images are not correctly identified.

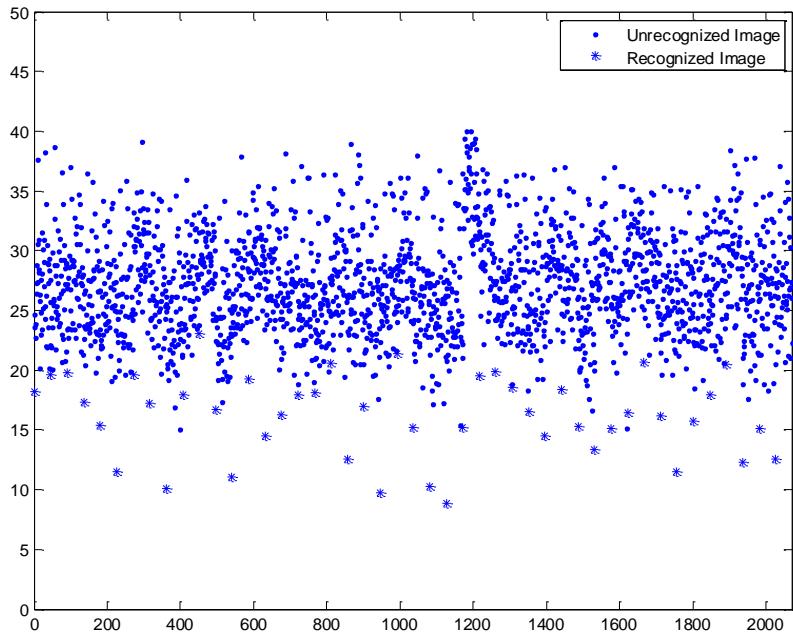


Figure 6.2: Error Norms between Each of the Probe and Gallery Images: The horizontal axis denotes the image index, and the vertical axis denotes the error norm in terms of the PCA feature vectors

6.2.3 Experiment with Pose Compensation

Figure 6.3(a) shows another example of the 2D probing images in a semi-profile view, figure 6.3(b) shows the projection of its corresponding 3D gallery image in the fronto-parallel view, and figure 6.3(c) is the 2D projection of the same 3D gallery image after pose adjustment.

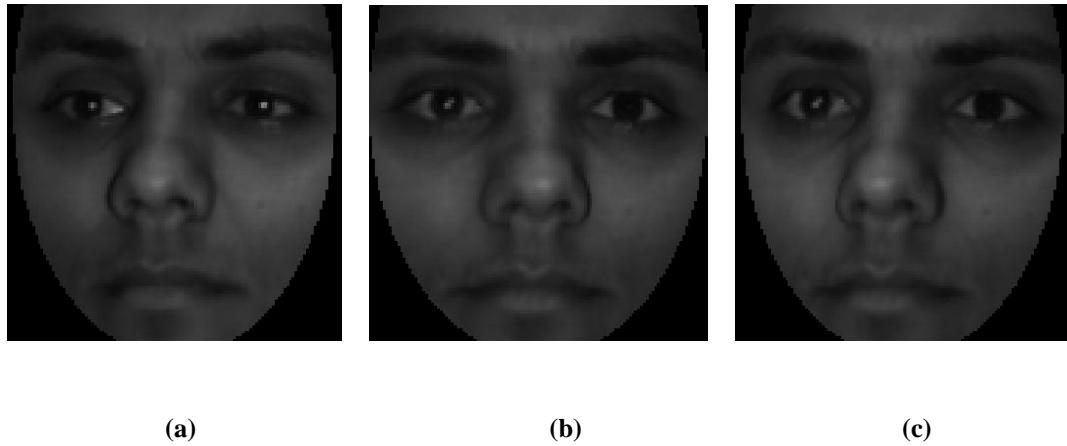


Figure 6.3: Example of the 2D Probing Image in Semi-profile View (a) The Probe 2D Image (b) the 2D Projection of its 3D Gallery Image from the Fronto-parallel View (c) the 2D Projection of the 3D Gallery Image after Rotation to the Same Pose

The same experiment method is used as the above, figure 6.4 shows again the errors feature vectors between the probe images and the projected gallery images. For the recognized faces, the mean vector distance is 14.10 and the standard deviation is 2.93; for the other faces, the vector distance mean is 28.01 and the standard deviation is 4.16. And the recognition rate is 97.83% since one out of 46 images is not correctly identified. Comparing to the first experiment described in 6.2.1, one observes that the difference of the average error distances between the recognized face and the unrecognized ones is larger, and the standard deviation of errors within each group is smaller as well. This means that within both the recognized and unrecognized face groups, the members are closer, but the two groups are more separated, which can lead to a better recognition rate.

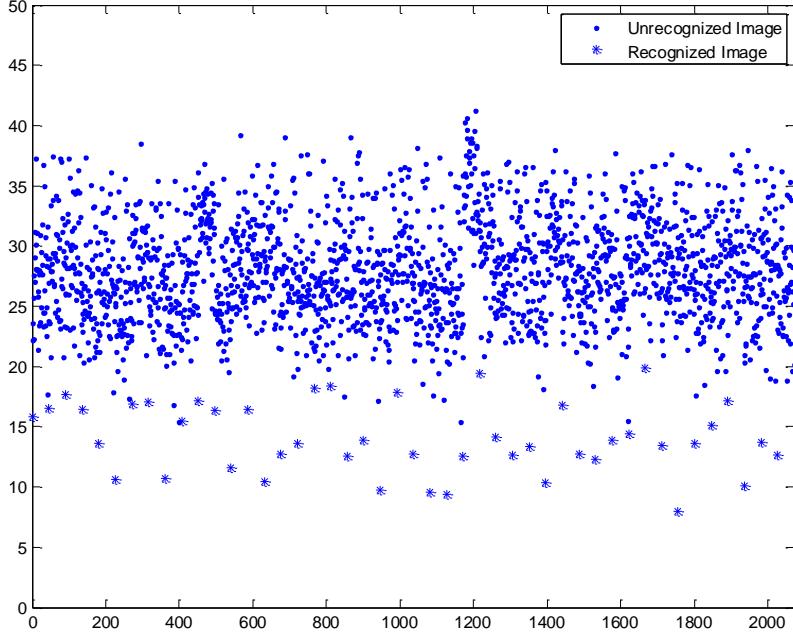


Figure 6.4: Error Norms between Each of the Probe and Gallery Images. The horizontal axis denotes the image index, the vertical axis denotes the error norm in terms of the PCA feature vectors

6.2.4 Experiment with Both Pose and Illumination Compensation

In this experiment, we first positioned each 3D gallery image to the same pose as the 2D probe image, and then applied the Phong reflection model with the optimization algorithm outlined in previous chapter for light compensation. After this procedure, the projected gallery images are used for face recognition given any probe image.

Figure 6.5 shows an example of illumination compensation, after the pose adjustment so that both the 2D probing image and the projected 3D gallery image have the same pose angle. And figure 6.6 shows again the errors feature vectors between the probe images and the projected gallery images. For the recognized faces, the mean vector distance is 7.45 and the standard deviation is 2.99; for the other faces, the vector distance mean is 24.79 and the standard deviation is 4.06. And the recognition rate is 100% since

all of 46 images are correctly identified. The average error distance for the recognized face is significantly smaller than those in the previous two experiments, which leads to the most accurate recognition result.

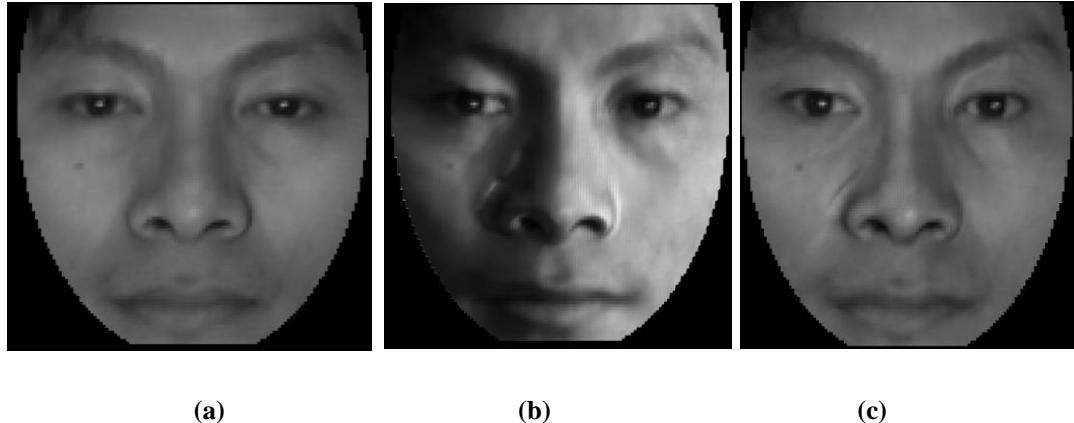


Figure 6.5: 101. (a) Probing Image, (b) No Compensation, and (c) Compensation with 3 Light Sources

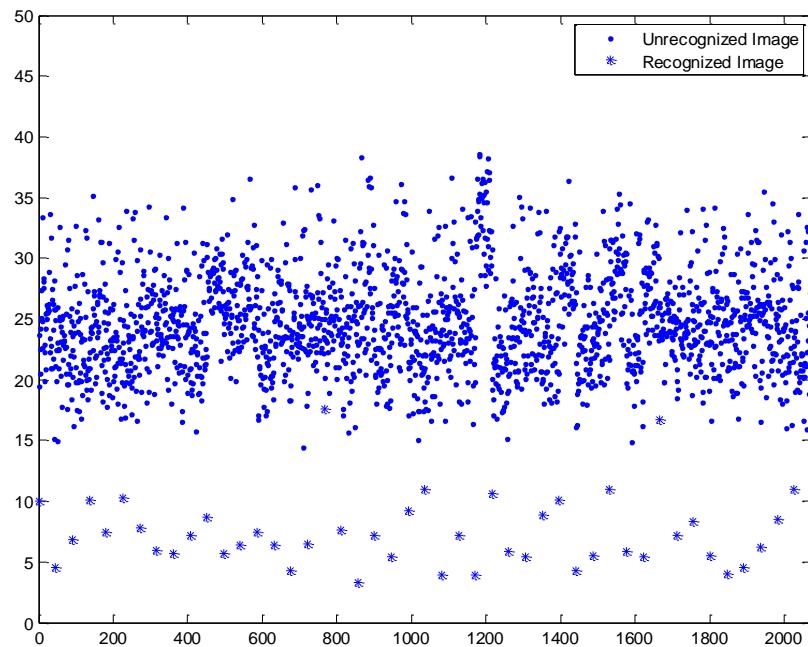


Figure 6.6: Error Norms between Each of the Probe and Gallery Images: The horizontal axis denotes the image index, and the vertical axis denotes the error norm in terms of the PCA feature vectors

6.2.5 Summary

The results from these three experiments are compared in table 6.1. The False Positive Rate here describes the error of matching a probe image to a wrong gallery image. The result shows a significant improvement by applying pose angle compensation and then applying the illumination compensation.

	Recognition Rate	False Positive Rate
Experiment 1	91.30%	8.70%
Experiment 2	97.83%	2.13%
Experiment 3	100%	0%

Table 6.1: 3D Shape Assisted 2D Recognition Rate Comparison

6.3 2D Assisted 3D Recognition

6.3.1 Introduction

In this section, we starts with creating the 3D face model using only one 2D probing image, then we demonstrate the test result using multiple 2D face images. This method may be optimal since in real life usage, it is possible to get multiple 2D probing images, and the more images we acquire, the better the 3D face model will be created. After that we demonstrate the experiment using noisy probing images. Finally two different recognition algorithms, PCA versus FLD, are utilized in our experiments.

6.3.2 Using Single 2D Image

Partial results of recreated 3D faces with the proposed scheme are shown in figure 6.7. In the figure, the first row depicts the input images of the subjects, the second row the recreated 3D images aligned with the input photos, and the third row the norm of the cost functions against the number of iterations.

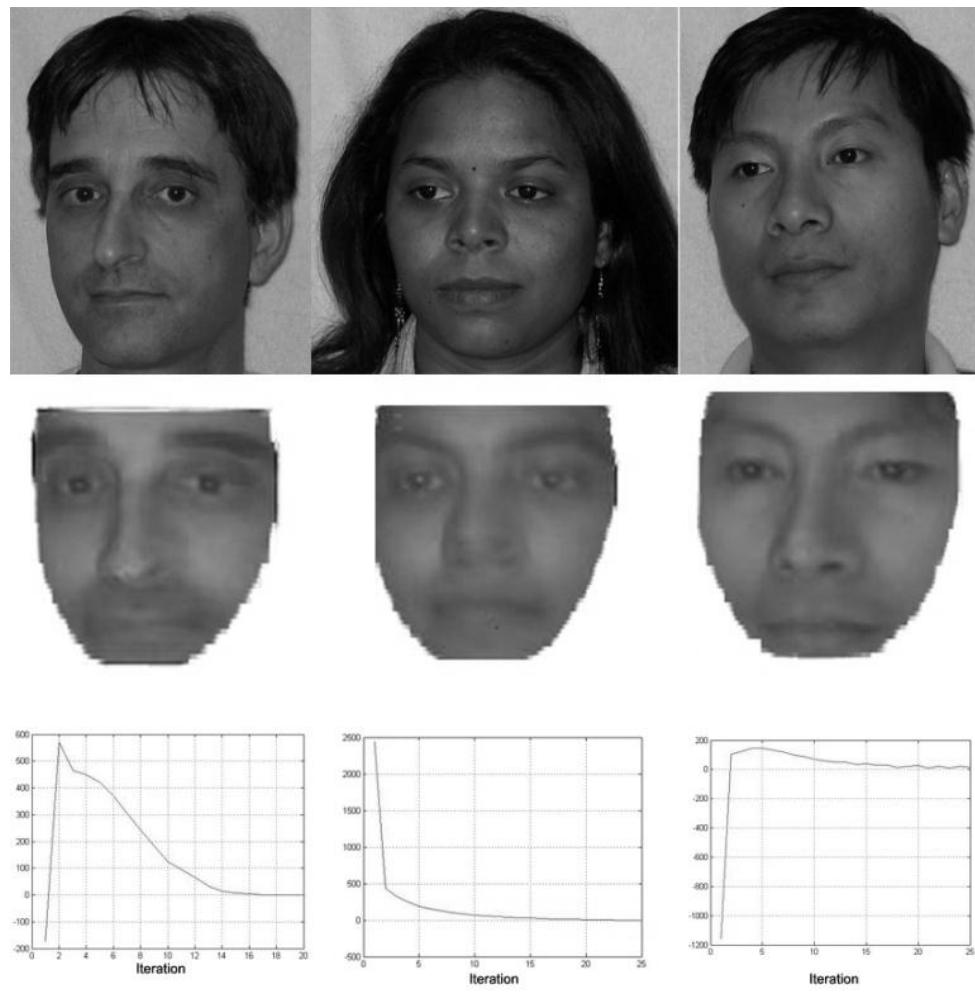


Figure 6.7: Recreated 3D Faces from 2D Images

6.3.3 Using Multiple 2D Photos

Shown in figure 6.8 is an example of the results obtained by applying the method in Chapter 4. The second row shows the recreated 3D faces of the subject whose photo is given in the first row. From the left to the right, the number of enrollment photos changes from one to three. It is evident that with 3 input photos of different poses of a subject, our method is able to create a 3D face image of the subject who is not in the 3D database.



Figure 6.8: Recreated 3D Faces from Different Number of Input Photos

6.3.4 Noisy Input Images

The third experiment involved noisy images in an effort to verify how well the 3D face recreation algorithm works under non-ideal situations. Two types of noisy images were created. In the first case, a Gaussian blur equation was applied to mimic the

situation that the camera was not focused properly, and in the second case, objects were moved to create a blurred image. Shown in figure 6.9 are the recreated 3D images together with the input images blurred from “an unfocused camera.” Similarly, shown in figure 6.10 are the recreated 3D images together with the images blurred due to “camera motion.” The results illustrate that the proposed scheme works reasonably well under these not so perfect situations.

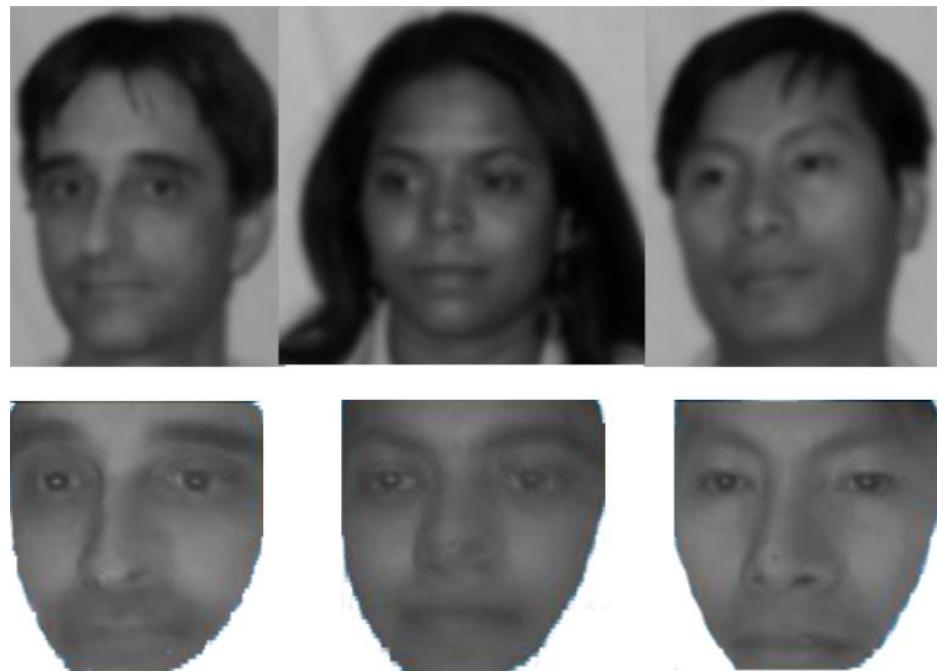


Figure 6.9: Recreated 3D Faces from “Unfocused Images”



Figure 6.10: Recreated 3D Faces from Images under “Camera Motion”

6.3.5 Experiment Using PCA for Feature Extraction

Figure 6.11 shows an example of the test result. Figures 6.11(a) and 6.11(c) are the same 3D synthetic image generated to match figure 6.11(d), the 2D probing photo. And in the recognition part, figure 6.11(a) is used to compare with 6.11(b), the 3D gallery images stored in the database.

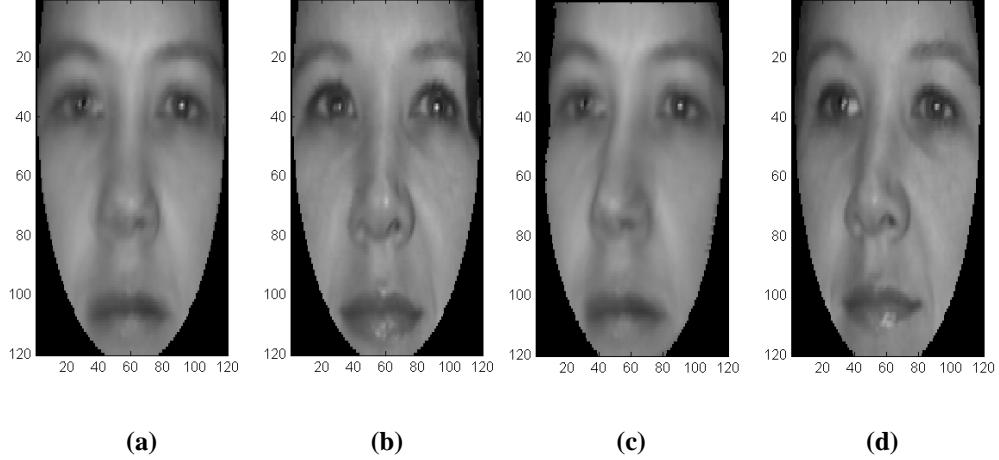


Figure 6.11: Example of 3D PCA Test Result (a) The 3D Synthetic Image in Front View, (b) The 3D Gallery Image in Front View, (c) The Same Synthetic Image in Semi-profile View, (d) The 2D Probing Image in Semi-profile View.

In this experiment, we first positioned each 3D gallery image to the same pose as the 2D probing photo, and then applied the Phong reflection model with the optimization algorithm outlined in Chapter 4 for light compensation. This procedure was iteratively solved until the algorithm converges. The depth and intensity parameters of the resulting 3D synthetic image were then used for face recognition against those of each 3D gallery image.

Figure 6.12 shows the errors vectors between the feature vectors of the probing photos and the projected synthetic images. For the recognized faces, the mean distance was $2.0939e+006$ and the standard deviation was $2.6251e+006$; for other faces, the mean distance was $1.4479e+007$ and the standard deviation was $1.5372e+007$. And the recognition rate was 91.3% since 4 out of 46 images was not correctly identified.

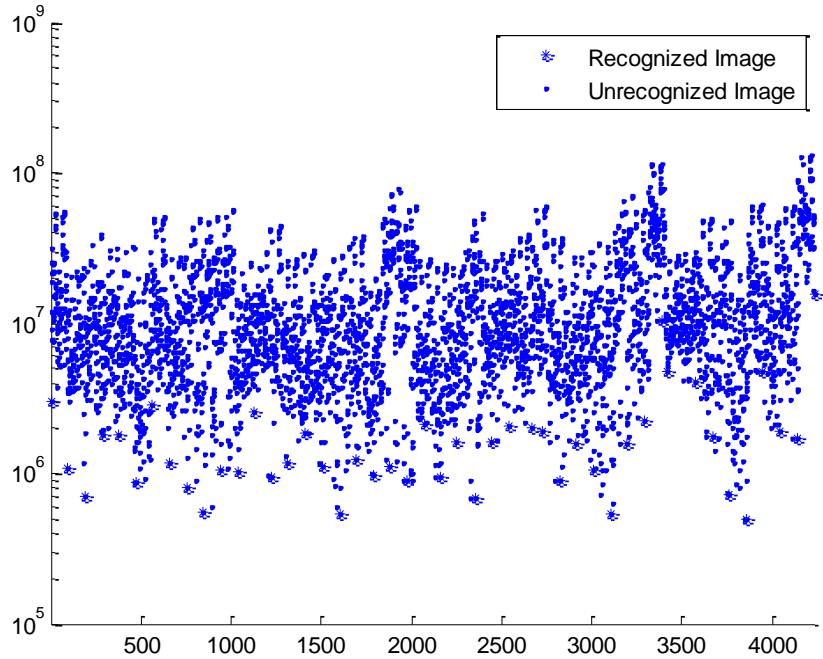


Figure 6.12: Error Norms between Each of the Probe and Gallery Images: The horizontal axis denotes the image index, and the vertical axis denotes the base-10 logarithm of error norm in terms of the PCA feature vectors

6.3.6 Experiment Using FLD for Feature Extraction

In this experiment, we implement the FLD method under the same conditions. Figure 6.13 shows again the error vectors between the feature vectors of the probe photos and the projected gallery images. For the recognized faces, the mean distance was $1.052\text{e+}06$ and the standard deviation was $1.278\text{e+}006$; for the other faces, the mean distance was $1.081\text{e+}007$ and the standard deviation was $1.178\text{e+}007$. And the recognition rate was 91.3% as well.

In comparison to the result from the previous PCA experiment, one observes that

with FLD, the difference of the average errors between the feature vectors of the recognized face and the unrecognized ones was larger, and the standard deviation of errors within each group was smaller as well. This means that within both the accepted and rejected face groups, the members were clustered tighter, and the two groups were more separated, which led to a better recognition.

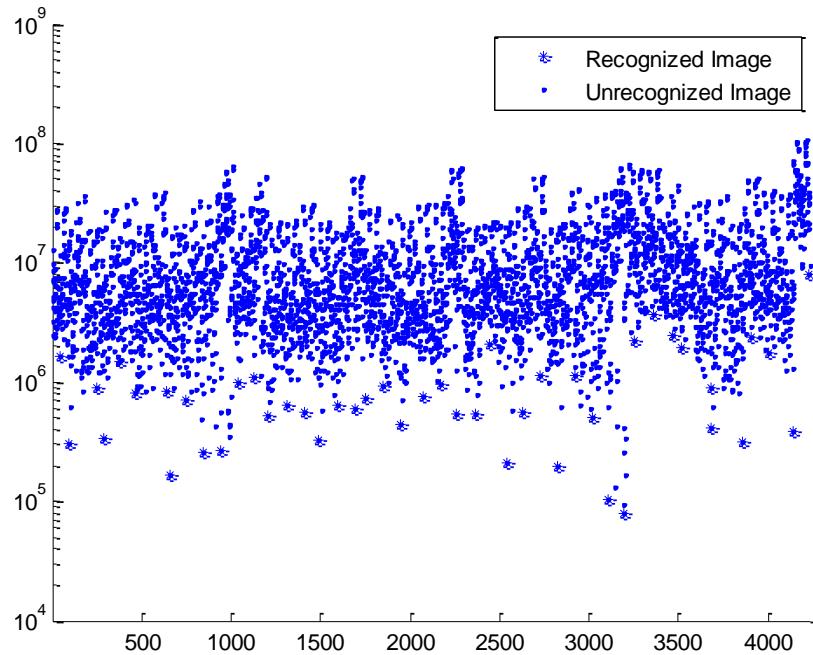


Figure 6.13: Error Norms between Each of the Probe and Gallery Images: The horizontal axis denotes the image index, and the vertical axis denotes the base-10 logarithm of error norm in terms of the FLD feature vectors

6.3.7 Summary

The above test results show the similar recognition rate using either PCA or FLD. We also compared the result of intensity parameters only with combined intensity and depth parameters, with the weight of 0.1 for depth vectors. The results from this

comparison are displayed in table 6.2. The False Positive Rate here describes the rate of matching a probe photo to a wrong gallery image.

	Recognition Rate	False Positive Rate
Intensity Parameters Only	89.10%	10.90%
Intensity and Depth	91.30%	8.70%

Table 6.2: 2D Assisted 3D Recognition Rate Comparison

6.4 Summary

Based on the results from the comparative studies demonstrated above, both face recognition approaches have very satisfactory recognition rates. The results show also the advantage and robustness of adding pose and/or illumination compensation to the 3D face model creation process. Later with the second approach, it is further demonstrated better 3D face models can be obtained from more 2D photos, which shows the flexibility and addictiveness of this new method.

Chapter 7

ENHANCED RECOGNITION METHOD USING 3D MODEL

7.1 Introduction

In the previous chapter, we discussed the condition that the probing face image is in 2D and the gallery images in the recognition database are in 3D. In this chapter, we enhanced our solution methods to handle both the 2D probing image and 2D gallery images. For the experiments below, we use the face data taken from AT&T Laboratories database of face, which includes ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). These images are first rotated in the 2D plane to make both eyes on the same horizontal line, then the image is cropped using an ellipse shaped mask. A preview image of the Database of Faces is shown in figure 7.1.

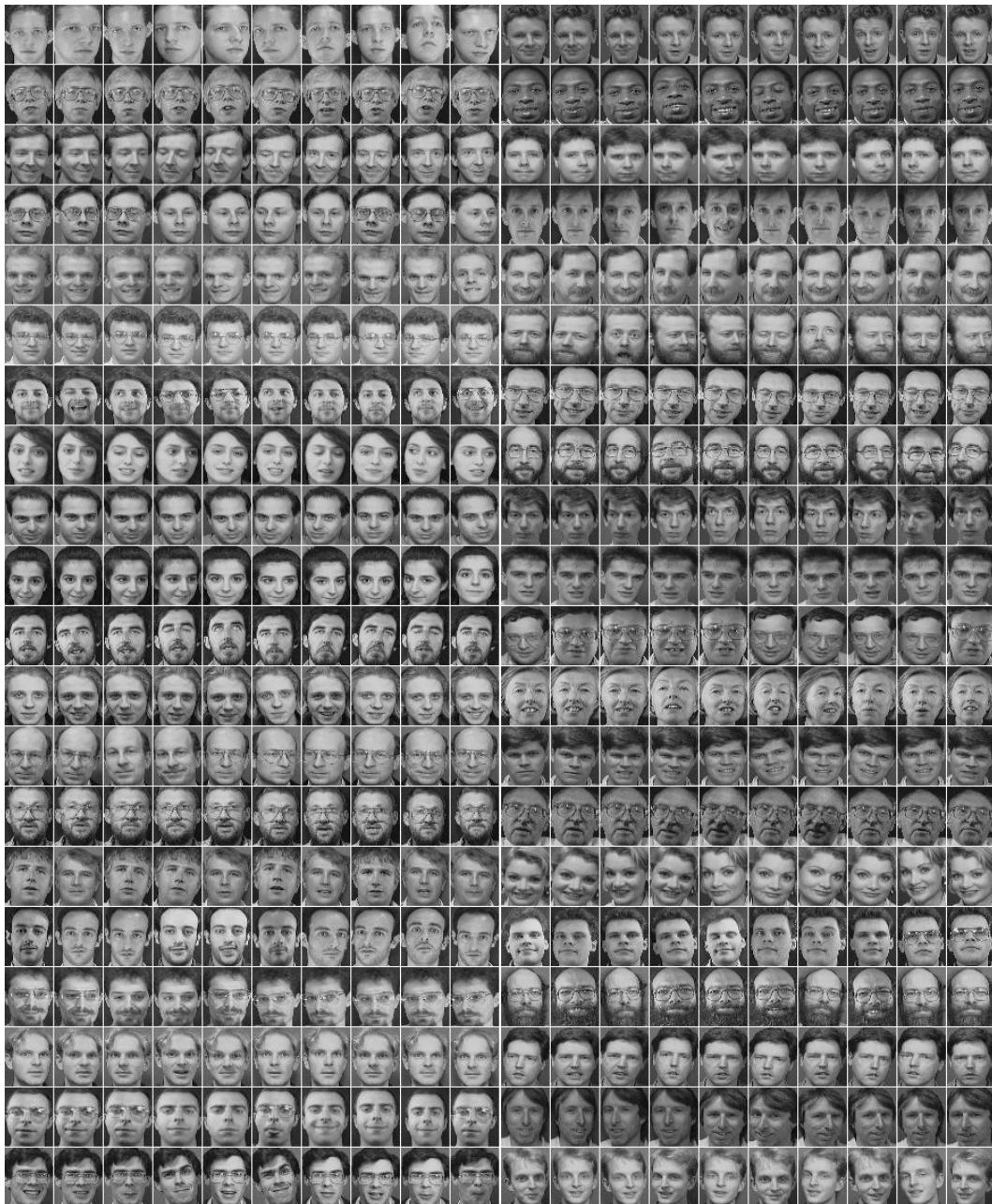


Figure 7.1: Preview of AT&T Database of Faces

In the first experiment, we simply implemented the 2D PCA on only the intensity data for the AT&T database. In the second experiment, we first calculated the 3D depth and intensity parameters, then five images of each subject were used for gallery images and five for probing, giving a total of 200 gallery images and 200 probing images. 3D

PCA is performed to recognize the probing images, and for each probing image, the results of the 200 identification test are reported as an error rate. Each error rate is calculated as the proportion of the images which are misclassified.

7.2 Solution Strategy

7.2.1 Modified Enrollment Stage

This approach is a variation of the outlined scheme mentioned in Chapter 4. In this modified enrollment stage, a 2D photo of a subject is obtained. A 3D face is created by selecting weights of the 3D models such that the projected 2D image of the 3D face is similar to the subject's 2D photo. The 3D weights (or 3D parameters) are then saved in the database. A detailed description on the enrollment stage is given next.

The proposed scheme for the enrollment stage is shown in figure 7.2. The process starts with the creation of a database for storing 3D PCA parameters of 3D model faces. Here 3D PCA means principal component analysis is performed on both depth and intensity feature, respectively. After this, the vectors \mathbf{z}_i^f and \mathbf{t}_i^f in equation 7.1 are obtained for the i th model.

$$\mathbf{z}^f = \sum_{i=1}^k a_i \mathbf{z}_i^f, \quad \mathbf{t}^f = \sum_{i=1}^k b_i \mathbf{t}_i^f \quad (7.1)$$

The next step involves an iteration process. An initial face is created by a given set of initial depth and intensity parameters. The created face is then rotated, scaled, and illuminated with light. It is then projected to the image plane, and a 2D PCA is performed to transform the 2D projected image to another parameter space, on which the projected

image is compared with the input image. The discrepancy between the two images will lead to another iteration in which a new set of parameters (more precisely, a new set of parameter adjustments) are generated in an effort to reduce the discrepancy. The iterative process continues until certain terminal conditions are met. The output of this algorithm is a set of 3D depth and intensity parameter vectors ($\mathbf{z}_i^f, \mathbf{t}_i^f$), $i = 1, 2, \dots, m$, where m is the number of enrolled subjects. This set of parameter vectors can be used to represent or recreate 3D faces which resemble the subjects given in 2D images. These parameter vectors can be used in the verification stage for human authentication.

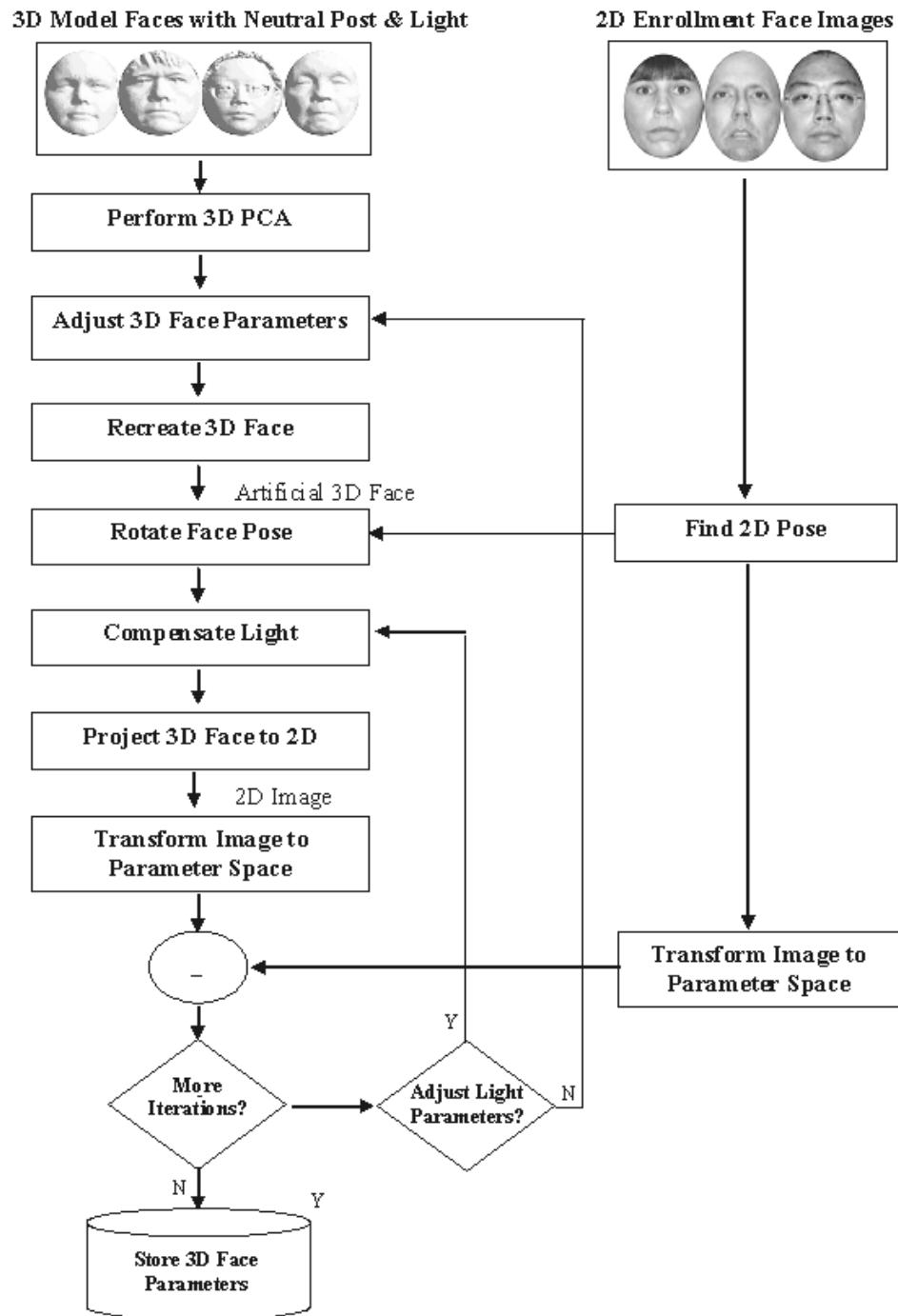


Figure 7.2: Enrollment Stage

7.2.2 Verification Stage

In the verification stage, a procedure similar to that found in the enrollment stage is applied to obtain the 3D parameters and compare them with those saved in the database to determine the identity of the person.

7.3 Experiment with Simple 2D PCA Face Recognition

In this experiment, the simple 2D PCA is performed on the AT&T face database. Figure 7.3(a) shows an example of a 2D probing image, and figure 7.3(b) and (c) are 2 gallery images. In this example, figure 7.3(b) is from the same person and figure 7.3(c) is from a different person.



Figure 7.3: Example of Probing and Gallery Images

For each 2D probing image, the 2D PCA algorithm is applied on it and each of the 200 gallery images. And the test is repeated for each of the 200 probing images.

To judge the performance of the face recognition algorithm, an error is defined as the Euclidean norm of the PCA feature vector difference between a probe image and a

projected gallery image. Figure 7.4 shows the errors generated by testing all 200 probing images. Note that the PCA feature vector of each probing image is compared with those of all 200 gallery images. For the recognized faces, the mean vector distance is 1516.5 and the standard deviation is 539.7; for the other faces, the vector distance mean is 4418.9 and the standard deviation is 1209.3. And the error rate is 89% since 22 out of 200 images are not correctly identified.

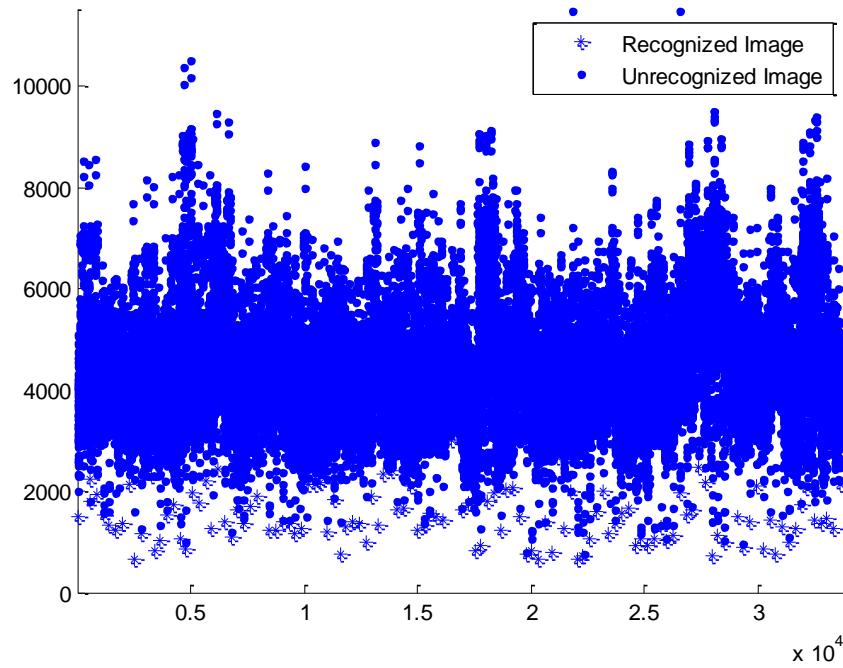


Figure 7.4: Error Norms between Each of the Probe and Gallery Images:

The horizontal axis denotes the image index, and the vertical axis denotes the error norm in terms of the PCA feature vectors.

7.4 Experiment with 3D Modeling and Recognition

We applied the 3D face modeling process on the 2D AT&T face database to gather the 3D depth and intensity parameters, partial results of recreated 3D faces are

shown in figure 7.5 below. In the figure, the first row depicts the input images of the subjects, which is the 2D face photo. The second row is the recreated 3D images aligned with the input photos.

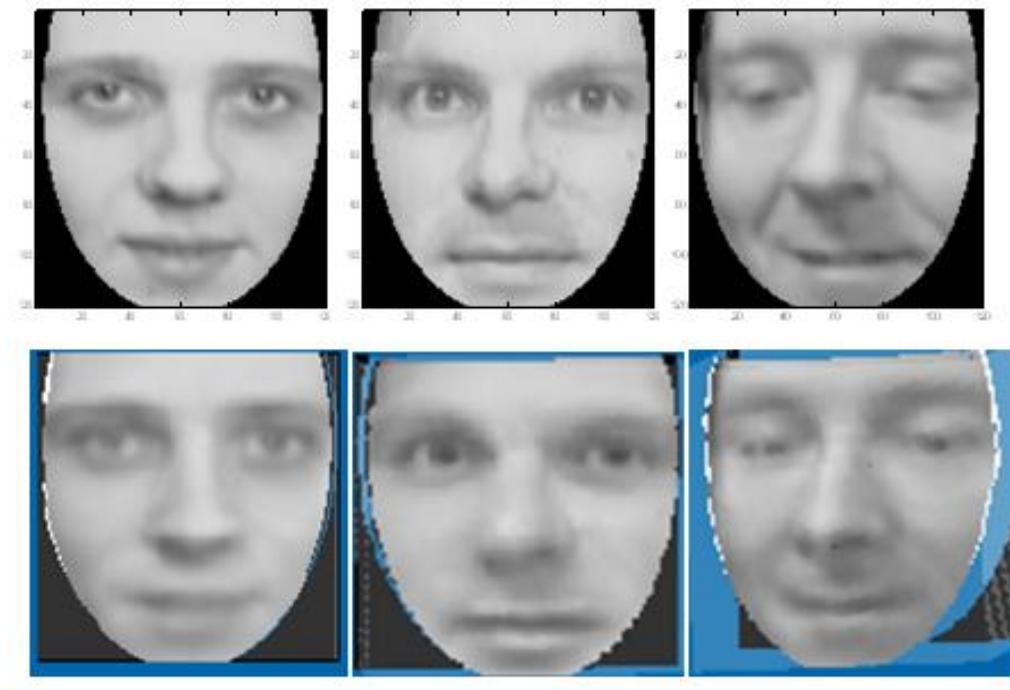


Figure 7.5: Recreated 3D Faces from 2D Images

The same experiment method is used as above -- a total of 200 gallery images are used in each test, and the test is repeated for each of the 200 probe images. Figure 7.6 shows again the errors feature vectors between the probing images and the gallery images. For the recognized faces, the mean vector distance is 1629.6 and the standard deviation is 569.4; for the other faces, the vector distance mean is 4423.9 and the standard deviation is 1109.0. And the error rate is 90.5% since 19 out of 200 images is not correctly identified.

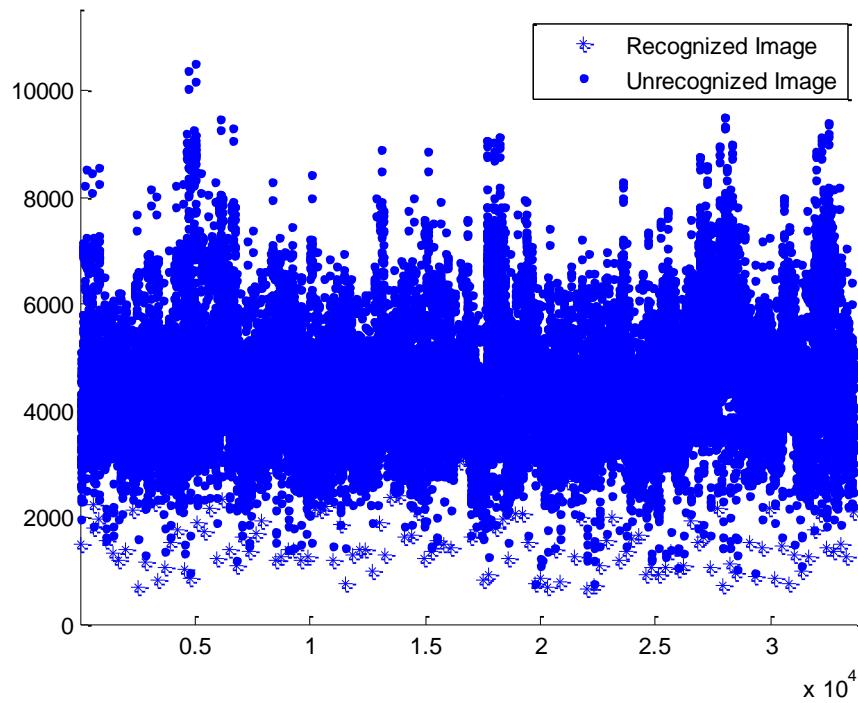


Figure 7.6: Error Norms between Each of the Probe and Gallery Images: The horizontal axis denotes the image index, and the vertical axis denotes the error norm in terms of the PCA feature vectors

Comparing to the previous experiment, one observes that the difference of the average error distances between the recognized face and the unrecognized ones is larger, and the standard deviation of errors within each group is smaller as well. This means that within both the recognized and unrecognized face groups, the members are closer, but the two groups are more separated, which can lead to a better recognition rate. The comparison results using either the proposed algorithm or the normal 2D recognition method are also compared in table 7.1. The error rate again is defined as the proportion of the images which are misclassified.

	Recognition Rate	False Positive Rate
2D PCA	89.00%	11.00%
3D Modeling and 3D PCA	90.50%	9.50%

Table 7.1: 3D Recognition Rate Comparison

7.5 Summary

In this chapter, a novel approach for more commonly available 2D probing image/2D gallery images face recognition is described, based on our approaches described in previous chapter. By creating 3D face models for the 2D face images, the 3D face recognition algorithm can then be applied, and it provide robustness against face angle and illumination condition changes. To test the performance of this new method, the widely used AT&T 2D Database of Face is utilized and our test results show an excellent recognition rate achieved by the proposed method.

Chapter 8

CONCLUSIONS

A novel approach for modeling human faces is proposed in this dissertation for the purpose of recognition. The potential advantages of the new approach are analyzed in this chapter, along with a discussion of the direction for future work.

8.1 Summary of the Research

Two methods for face recognition using 2D probe photos along with a 3D gallery face database have been presented in this dissertation. These approaches gave satisfactory recognition rates of up to around 98%.

In the first method, the 3D gallery photos from the database are projected to the 2D plane and then compared with the 2D probing faces using 2D recognition algorithms. The key of this method is to compensate pose and illumination before the 3D images are used for face recognition, which is feasible when gallery images are three-dimensional.

In the second method, 3D face model is first created using the 2D probing images. A major difference between this method and those in the literature is that in the new method, the discrepancy between the feature vectors of a 2D probing photo and a projected gallery image is used to guide an optimization algorithm, which seeks an optimal solution in a 3D face parametric space.

These methods have applications in biometric authentication and recognition. They are practical as it does not use 3D probe images, and 2D photos are easy to collect. Experimental studies have revealed that the proposed methods can achieve a superior performance over methods without pose and/or light compensation, and also the better recognition rates compared with the existing 2D recognition techniques. These methods work reasonably well, even with noisy, impaired images.

Based on the fundamentals of the above two methods, we further presented a method of using 2D probing images and 2D gallery images. The above 3D face model is created on both the probing images and the gallery images, and both the depth and the intensity data are used in the 3D recognition algorithm. This method inherent the benefits of robustness against pose and illumination variations, and it adds also some flexibility of using multiple 2D probing images or using noisy images.

8.2 Future Works

High speed internet is becoming widespread, which makes the multimedia applications, such as people exchanging video and audio data files, more commonly available. Nowadays more desktop computers, laptop computer, and tablet computers equipped with video cameras as default factory configurations, this makes the face recognition in this context a valuable tool. New applications will be possible, for example, logging onto the computer with authentication means based on face recognition, indexing the social network video data/contact through face recognition. The fast expansion of the home entertainment industry will also bring more novel applications such as search specific actor/actress movies via face recognition, etc.

In light of these scenarios, the ideas introduced in this dissertation can be used to develop a large scale system capable of recognizing faces in real-time.

The algorithms developed in our research provide only the framework of 3D face recognition using 2D photos, and more complicated algorithms can be applied to achieve better recognition results:

1. For the classification algorithm, the regularized least square criteria can be replaced by the Support Vector Machine (SVM).
2. For the feature extraction algorithms, in this dissertation we use Principle Component Analysis (PCA) and Linear Fisher Discriminant (FLD), and they can be replaced by Independent Component Analysis (ICA).

Future studies also include performing more comprehensive experimentations in order to explore the strengths and weaknesses of the proposed approach, and adding face expression compensation to the overall scheme.

BIBLIOGRAPHY

- [1] P. Benson and D. Perrett. *Perception and recognition of photographic quality facial caricatures: implications for the recognition of natural images*, In V. Bruce, editor, Face Recognition, pages 105-135, Laurence Erlbaum Associates, 1991.
- [2] R. Diamond and S. Carey. Why faces are and area not special: an effect of expertise. *Journal of Experimental Psychology: General*, 115:107-117, 1986.
- [3] R. C. Fowler. FINGERPRINT: an old touchstone decriminalized. *IEEE Spectrum*, page 26, February 1994.
- [4] J. E. Siedlarz. IRIS: more detailed than a fingerprint. *IEEE Spectrum*, page 27, February 1994.
- [5] R. Mandelbaum. SPEECH: just say the word. *IEEE Spectrum*, page 30, February 1994.
- [6] M. Turk and A. Pentland, Eigenfaces for recognition, *J. of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
- [7] V. Belhumeur, J. Hespanda, and D. Kiregeman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, *IEEE Trans. PAMI*, Vol. 19, No. 7, pp. 711-720, July 1997.

- [8] Pong C. Yunen, and J.H. Lai, Face representation using independent component analysis, *Pattern Recognition*, Vol. 35, pp. 1247-1257, 2002.
- [9] A. Tefas, C. Kotropoulos, and L. Pitas, Using Support Vector Machines to Enhance the Performance of Elastic Graph Matching for Frontal Face Authentication, *IEEE Trans. PAMI*, Vol. 23, No. 7, pp.735-746, 2001.
- [10] Z. Pan, A. G. Rust, and H. Bolouri, Image Redundancy for Neural Network Classification Using Discrete Cosine Transforms, *Neural Networks*, 2000. IJCNN 2000, *Proceedings of the IEEE-INNS-ENNS International Joint Conference*, Volume: 3 , 2000.
- [11] Z. Riaz, A. Gilgiti and S. M. Mirza, Face recognition: a review and comparison of HMM, PCA, ICA and neural networks, *ETech 2004*, pp.41-46, July 2004.
- [12] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002: Evaluation Report. *Technical Report IR 6965*, <http://www.itl.nist.gov/iad/894.03/face/face.html>, National Institute of Standards and Technology, 2004.
- [13] K. Messer, J. Kittler, M. Sadeghi, M. Hamouz, A. Kostin, S. Marcel, S. Bengio, F. Cardinaux, C. Sanderson, N. Poh, Y. Rodriguez, K. Kryszczuk, J. Czyz, L. Vandendorpe, J. Ng, H. Cheung, and B. Tang. Face authentication competition on the Banca database . D. Zhang and A. K. Jain, editors, *Proc First International Conference on Biometric Authentication*, pages 8–15. Springer, July 2004.
- [14] D. M. Blackburn, M. Bone, and P. Phillips, *Evaluation report. Facial Recognition Vendor Test 2000*, Feb. 2001.

- [15] T. S. Huang and L. A. Tang, *3D Face Modeling and Its Applications*, *Int'l J. Pattern Recognition and Artificial Intelligence*, vol. 10, no. 5, pp. 491-519, 1996.
- [16] D.G. Lowe, Fitting Parameterized Three-Dimensional Models to Images, *IEEE Trans. Pattern Analysis & Machine Intelligence*, vol. 13, no. 5, pp. 441-450, May 1991.
- [17] W. Zhao and R. Chellappa, SFS Based View Synthesis for Robust Face Recognition, *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, pp. 285-292, 2000.
- [18] V. Blanz and T. Vetter, Face Recognition Based on Fitting a 3D Morphable Model, *IEEE Trans. Pattern Analysis & Machine Intelligence*, vol. 25, no. 9, pp. 1063-1074, Sep. 2003.
- [19] G. D. Hager and P. N. Belhumeur. Efficient Region Tracking with Parametric Models of Geometry and Illumination, *IEEE Trans. PAMI*, vol. 20, no. 10, pp. 1025-1039, 1998.
- [20] P. W. Hallinan. A Low-Dimensional Representation of Human Faces for Arbitrary Lighting Conditions, *Proc. CVPR*, pp. 995-999, 1994.
- [21] J. Lee, B. Moghaddam, H. Pfister, and R. Machiraju, A Bilinear Illumination Model for Robust Face Recognition, *IEEE International Conference on Computer Vision (ICCV)*, ISSN: 1550-5499, vol. 2, pp. 1177-1184, Oct. 2005.
- [22] S. Malassiotis, and M. G. Strintzis, Robust Face Recognition using 2D and 3D Data: Pose and Illumination Compensation, *Pattern Recognition*, vol. 38, no. 12 , pp. 2537-2548, Dec. 2005.

- [23] Y. Wang and C.-S. Chua, Robust Face Recognition from 2D and 3D Images Using Structural Hausdorff Distance, *Image and Vision Computing*, vol 24, no. 2, pp. 176-185, Feb. 2006.
- [24] M. Mayo and E. Zhang, 3D Face Recognition Using Multiview Keypoint Matching, *Advanced Video and Signal Based Surveillance, AVSS '09. Sixth IEEE International Conference on*, 2-4 Sept. 2009 page(s): 290 – 295. 2009.
- [25] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, Face recognition: A literature survey, *ACM Computing Surveys (CSUR)*, v.35 n.4, p.399-458, December 2003.
- [26] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Trans on Pattern Analysis and Machine Intelligence*. 15(10):1042-1052. 1993.
- [27] G. Robertson and I. Craw. Testing face recognition systems. *British Machine Vision Conference*. Volume 1, page 25-34. MBVA Press, 1993.
- [28] M. D. Kelly, Visual identification of people by computer. *Tech. rep.* AI-130, Stanford AI Project, Stanford, CA. 1970.
- [29] T. Kanade, Computer recognition of human faces. Birkhauser, Basel, Switzerland, and Stuttgart, Germany. 1973.
- [30] I. J. Cox, J. Ghosn, and P. N. Yianilos. Feature-based face recognition using mixturedistance. *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*. 209–216. 1996.

- [31] J. Buhmann, M. J. Lades, and C. Malsburg. Size and distortion invariant object recognition by hierarchical graph matching. *Proceedings, International Joint Conference on Neural Networks*. 411–416. 1990.
- [32] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, C. Malsburg, R. Wurtz, and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. Comput.* 42, 300–311. 1993.
- [33] K. Okada, J. Steffans, T. Maurer, H. Hong, E. Elagin, H. Neven, and C. Malsburg. The Bochum/USC Face Recognition System and how it fared in the FERET Phase III Test. *Face Recognition: From Theory to Applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. S. Huang, Eds. Springer-Verlag, Berlin, Germany, 186–205. 1998.
- [34] L. Wiskott, J. Fellous, and C. Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. Patt. Anal. Mach. Intell.* 19, 775–779. 1997.
- [35] N. Kruger, M. Potzschi, C. Determination of face position and pose with a learned representation based on labeled graphs. *Image Vis. Comput.* 15, 665–673. 1997.
- [36] T. Maurer and C. Singleview based recognition of faces rotated in depth. In *Proceedings, InternationalWorkshop on Automatic Face and Gesture Recognition*. 176–181. 1996.
- [37] I. Biederman, and P. Kalocsal. Neural and psychophysical analysis of object and face recognition. In *Face Recognition: From Theory to Applications*, H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. S. Huang, Eds. Springer-Verlag, Berlin, Germany, 3–25. 1998.

- [38] A. V. Nefian and M. H. Hidden Markov models for face recognition. In *Proceedings, International Conference on Acoustics, Speech and Signal Processing*. 2721–2724. 1998.
- [39] F. Samaria. Face recognition using hidden markov models. Ph.D. dissertation. University of Cambridge, Cambridge, U.K. 1994.
- [40] F. Samaria and S. HMM based architecture for face identification. *Image Vis. Comput.* 12, 537–583. 1994.
- [41] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back. Face recognition: A convolutional neural-network approach. *IEEE Trans. Neural Netw.* 8, 98–113. 1997.
- [42] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Trans. Patt. Anal. Mach. Intell.* 19, 696–710. 1997.
- [43] D. L. Swets and J. Weng. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Patt. Anal. Mach. Intell.* 18, 831–836. 1996.
- [44] W. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant analysis of principal components for face recognition. *Proceedings, International Conference on Automatic Face and Gesture Recognition*. 336–341. 1998.
- [45] R. A. Fisher. The statistical utilization of multiple measurements. *Ann. Eugen.* 8, 376–386. 1938.
- [46] P. J. Phillips. Support vector machines applied to face recognition. *Adv. Neural Inform. Process. Syst.* 11, 803–809. 1998.

- [47] C. Liu and H. Andwechsler. Evolutionary pursuit and its application to face recognition. *IEEE Trans. Patt. Anal. Mach. Intell.* 22, 570–582. 2000.
- [48] S. Z. Li and J. Lu. Face recognition using the nearest feature line method. *IEEE Trans. Neural Netw.* 10, 439–443. 1999.
- [49] M. S. Bartlett, H. M. Lades, and T. Sejnowski. Independent component representation for face recognition. In *Proceedings, SPIE Symposium on Electronic Imaging: Science and Technology*. 528–539. 1998.
- [50] S. H. Lin, S. Y. Kung, and L. J. Lin. Face recognition/detection by probabilistic decisionbased neural network. *IEEE Trans. Neural Netw.* 8, 114–132. 1997.
- [51] K. Etemad and R. Chellappa. Discriminant analysis for recognition of human face images. *J. Opt. Soc. Am. A* 14, 1724–1733. 1997.
- [52] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*. 1994.
- [53] P. Penev and J. Atick. Local feature analysis: A general statistical theory for objecct representation. *Netw.: Computat. Neural Syst.* 7, 477–500. 1996.
- [54] A. Lanitis, C. J. Taylor, and T. F. Cootes. Automatic face identification system using flexible appearance models. *Image Vis. Comput.* 13, 393–401. 1995.
- [55] B. Amor, M. Ardabilian, and L. Chen. Toward a region-based 3D face recognition approach, *2008 IEEE International Conference on Multimedia & Expo (ICME 2008)*, June 23-26 2008, Hannover, Germany.

- [56] C. Zhong, Z.A. Sun, T.N. Tan, and Z.F. He. Robust 3D face recognition in uncontrolled environments, *CVPR08(1-8)*. 2008.
- [57] A. Scheenstra, A. Ruifrok, and R. C. Veltkamp, A Survey of 3D Face Recognition Methods. *AVBPA 2005*: 891-899. 2005
- [58] K. W. Bowyer, K. I. Chang, and P. J. Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition. *Computer Vision and Image Understanding*, 101:1–15, 2006.
- [59] C. Xu, Y. Wang, T. Tan, and L. Quan, Depth vs. intensity: Which is more important for face recognition? *Proc. 17th International Conference on Pattern Recognition*, Cambridge, UK, 2004.
- [60] K. I. Chang, K. W. Bowyer, and P. J. Flynn, An evaluation of multimodal 2d+3d face biometrics, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 619–624, 2005.
- [61] X. Lu and A. K. Jain, Integrating range and texture information for 3d face recognition. *Proc. 7th IEEEWorkshop on Applications of Computer Vision*, pp. 156–163. 2005.
- [62] J. Y. Cartoux, J. T. LaPreste, and M. Richetin. Face authentication or recognition by profile extraction from range images. *Proc. of the Workshop on Interp. of 3D Scenes*, pages 194–199, November 1989.
- [63] J. C. Lee and E. Milios. Matching range images of human faces. *Int'l Conf. on Comp. Vision*, pages 722–726, 1990.

- [64] G. Gordon. Face recognition based on depth maps and surface curvature. *Geometric Methods in Computer Vision, SPIE* 1570:1–12, July 1991.
- [65] T. Nagamine, T. Uemura, and I. Masuda. 3D facial image analysis for human identification. *International Conference on Pattern Recognition (ICPR 1992)*, pages 324–327, 1992.
- [66] B. Achermann, X. Jiang, and H. Bunke. Face recognition using range images. *International Conference on Virtual Systems and MultiMedia*, pages 129–136, 1997.
- [67] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation principal directions for curved object recognition. *International Conference on Pattern Recognition*, pages 638 – 642, 1996.
- [68] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation. *Third International Conference on Automated Face and Gesture Recognition*, pages 372–377, 1998.
- [69] C. Hesher, A. Srivastava, and G. Erlebacher. A novel technique for face recognition using range images. *Seventh Int'l Symp. on Signal Processing and Its Applications*, 2003.
- [70] G. Medioni and R. Waupotitsch. Face recognition and modeling in 3D. *IEEE Int'l Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2003)*, pages 232—233, October 2003.

- [71] A. B. Moreno, A\'ngel Sa\'nchez, J. F. V\'elez, and F. J. D\'iaz. Face recognition using 3D surface-extracted descriptors. *Irish Machine Vision and Image Processing Conference (IMVIP 2003)*, September 2003.
- [72] Y. Lee, K. Park, J. Shim, and T. Yi. 3D face recognition using statistical multiple features for the local depth information. *16th International Conference on Vision Interface*, available at www.visioninterface.org/vi2003, June 2003.
- [73] S. Chang, M. Rioux, and J. Domey. Face recognition with range images and intensity images. *Optical Engineering*, 36 (4):1106–1112, April 1997.
- [74] S. Lao, Y. Sumi, M. Kawade, and F. Tomita. 3D template matching for pose invariant face recognition using 3D facial model built with isoluminance line based stereo vision. *Int'l Conf. on Patt. Rec. (ICPR 2000)*, pages II:911–916, 2000.
- [75] C. Beumier and M. Achery. Face verification from 3D and grey level cues. *Patt. Rec. Letters*, 22:1321–1329, 2001.
- [76] Y. Wang, C. Chua, and Y. Ho. Facial feature detection and face recognition from 2D and 3D images. *Pattern Recognition Letters*, 23:1191–1202, 2002.
- [77] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Expression-invariant 3D face recognition. *Audio- and Video-Based Person Authentication (AVBPA 2003)*, LNCS 2688, J. Kittler and M.S. Nixon, eds.:62–70, 2003.
- [78] A. S.Mian, M. Bennamoun, and R. Owens. An efficient multimodal 2d-3d hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1927–1943, 2007.

- [79] K. Chang, K. Bowyer, and P. Flynn. Face recognition using 2D and 3D facial data. *2003 Multimodal User Authentication Workshop*, pages 25–32, December 2003.
- [80] S. Romdhani and T. Vetter, Efficient, Robust and Accurate Fitting of a 3D Morphable Model. *Proceedings of the European Conference on Computer Vision*. 2003.
- [81] J. Huang, B. Heisele, and V. Blanz, Component-Based Face Recognition with 3D Morphable Models. *Proceedings of the Audio- and Video-Based Biometric Person Authentication*, Vol 2688. 27–34. 2003.
- [82] B. Heisel and T. Koshizen, Components for Face Recognition. *Proceedings of the Audio- and Video-Based Biometric Person Authentication*, Vol 2688. 153–159. 2003.
- [83] M. Everingham and A. Zisserman. Automated Visual Identification of Characters in Situation Comedies. *Proc .of 17th International Conference on Pattern Recognition, (ICPR '04)*, pp. 983–986, 2004.
- [84] F. Parke and K. Waters. Computer Facial Animation. A.K. Peters, 1996.
- [85] Y. Lee, D. Terzopoulos, and K. Waters. Realistic Modeling for Facial Animation. *Proceedings of ACM SIGGRAPH*, pages 55–62, 1995.
- [86] D. DeCarlo, D. Metaxas, and M. Stone. An Anthropometric Face Model using Variational Techniques. *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 67–74, 1998.

- [87] B. T. Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975.
- [88] L. Yin and M. T. Yourst. 3D face recognition based on highresolution 3D face modeling from frontal and profile views. *WBMA '03: Proceedings of the 2003 ACMSIGMM workshop on Biometrics methods and applications*, pages 1–8. ACM Press, 2003.
- [89] N. Grammalidis, N. Sarris, C. Varzokas, and M. Strintzis. Generation of 3-D head models from multiple images using ellipsoid approximation for the rear part. *International Conference on Image Processing*, 1:284–287, 2000.
- [90] A.-N. Ansari and M. Abdel-Mottaleb. 3D face modeling using two orthogonal views and a generic face model . *ICME 03, Proceedings of International Conference on Multimedia and Expo*, 3:289–92, 2003.
- [91] C.-S. Chua, F. Han, and Y.-K. Ho. 3D human face recognition using point signature. *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 233–238, 2000.
- [92] C. Xu, Y. Wang, T. Tan, and L. Quan. Three-dimensional face recognition using geometric model . *Proc. SPIE Vol. 5404*, pp. 304–315, 2004.
- [93] T. Heseltine, N. Pears, and J. Austin. Three-Dimensional Face Recognition Using Surface Space Combinations. *Proc. of British Machine Vision Conference (BMVC'04)*, pages 527–536, 2004.
- [94] R. Basri and D. W. Jacobs. Lambertian Reflectance and Linear Subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2):218–233, 2003.

- [95] L. Zhang and D. Samaras. Pose Invariant Face Recognition under Arbitrary Unknown Lighting using Spherical Harmonics. *Proc. Biometric Authentication Workshop 2004, (in conjunction with ECCV2004)*, pp. 10–23, 2004.
- [96] X. Lu, D. Colbry, and A. K. Jain. Three-Dimensional Model Based Face Recognition . *Proc. of 17th International Conference on Pattern Recognition, (ICPR '04, Volume 1:362–366*, 2004.
- [97] R. Beveridge, et. al., The CSU Face Identification Evaluation System Users Guide: Version 5.0, 2003.
- [98] Z. Wu, Research on 3D Facial Surface Construction and Feature Extraction Based on B-splines, M.S. thesis, Xiamen University, Xiamen, 2008.
- [99] S. Lee and G. Wolberg, Scattered Data Interpolation with Multilevel B-Splines, *IEEE Transactions on Visualization and Computer Graphics*, Vol 3, No. 3, July, 1997.
- [100] K. S. Arun, T. S. Huang, S.D.Blostein, Least Square Fitting of Two 3D Point Sets, *IEEE Transactions on Pattern Analysis and Machine Intelligence(PMAI)*, pp. 698-700. 1987.
- [101] M. Ma and J. P. Kruth, Parametrization of randomly measured points for least squares fitting of B-spline curves and surfaces, *Computer-Aided Design*, Vol. 27, No. 9, pp. 663-675. 1995.
- [102] P. Yao, G. Evans, and A. Calway, Using affine Correspondence to Estimate 3-D Facial Pose, 2001. *Proceedings. International Conference on Image Processing*, vol. 3, page(s):919 – 922, Oct. 2001.

- [103] S. Ullman and D. P. Hutternlocher, Recognizing Solid Objects by Alignment with an Image, *International Journal of Computer Vision*, 5(2):195-212, 1990.
- [104] K.S. Fu, R.C. Gonzalez, C.S.G. Lee, *Robotics, Control, Sensing, Vision, and Intelligence*, pp23-25, McGraw-Hill, Inc., 1987.
- [105] S. Ullman and D. P. Hutternlocher, Recognizing Solid Objects by Alignment with an Image, *International Journal of Computer Vision*, 5(2):195-212, 1990.
- [106] A. Watt, *3D Computer Graphics, third ed*, Addison-Wesley, 2000.
- [107] R. Ishiyama and S. Sakamoto, Geodesic Illumination Basis: Compensating for Illumination Variations in any Pose for Face Recognition, *Proceedings. 16th International Conference on Pattern Recognition*, vol. 4, page(s):297 - 301, 2002.
- [108] M. Wand, M. Fischer, I. Peter, F. Meyer auf der Heide, and W. Straßer, The Randomized z-Buffer Algorithm: Interactive Rendering of Highly Complex Scenes, *SIGGRAPH 2001 Conference Proceedings*, 2001.
- [109] M. Kirby and L. Sirovich, Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Trans. Patt. Anal. Mach. Intell.* 12. 1990.
- [110] L. Sirovich and M. Kirby, Low-dimensional procedure for the characterization of human face. *J. Opt. Soc. Am.* 4, 519–524. 1987.
- [111] D. L. Ruderman, The statistics of natural images. *Netw.: Comput. Neural Syst.* 5, 598–605.1994.
- [112] W. Zhao, R. Chellappa, and P. J. Phillips, Subspace linear discriminant analysis for face recognition. *Tech. rep. CAR-TR-914*, Center for Automation Research, University of Maryland, College Park, MD. 1999.

- [113] P. J. Phillips, H. Moon, S. Rizvi, and P. Rauss, The FERET evaluation methodology for facerecognition algorithms. *IEEE Trans. Patt. Anal. Mach. Intell.* 2000.
- [114] K. Sung and T. Poggio, Example-based learning for view-based human face detection. *IEEE Trans. Patt. Anal. Mach. Intell.* 20, 39–51. 1997.
- [115] K. Fukunaga, Statistical Pattern Recognition, Academic Press, New York, NY. 1989.
- [116] A. Shashua. Geometry and photometry in 3D visual recognition. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, MA. 1994.
- [117] P. Penev and L. sirovich, the global dimensionality of face space. *Proceedings, International Conference on Automatic Face and Gesture Recognition*. 2000.
- [118] M. Al-Baali and R. Fletcher, Variational Methods for Non-Linear Least-Squares, *J. Opl. res. soc.* 36, no. 5, pp 405–421, 1985.
- [119] K. Madsen, A Combined Gauss-Newton and Quasi-Newton Method for Non-Linear Least Squares, *Institute for Numerical Analysis (now part of IMM). DTU Report NI-88-10*, 1998.