

**COMPARISON OF 3D FACIAL ANCHOR POINT LOCALIZATION
METHODS**

**A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY**

BY

MUSTAFA YAGCIOGLU

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING**

MAY 2008

Approval of the thesis:

**COMPARISION OF 3D FACIAL ANCHOR POINT LOCALIZATION
METHODS**

submitted by **MUSTAFA YAĞCIOĞLU** in partial fulfillment of the requirements for
the degree of **Master of Science in Electrical and Electronics Engineering
Department, Middle East Technical University** by

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. İsmet Erkmén
Head of Department, **Electrical and Electronics Engineering**

Asst. Prof. Dr. İlkay Ulusoy
Supervisor, **Electrical and Electronics Engineering Dept., METU**

Examining Committee Members:

Prof. Dr. Uğur Halıcı
Electrical and Electronics Engineering Dept., METU

Asst. Prof. Dr. İlkay Ulusoy
Supervisor, Electrical and Electronics Engineering Dept., METU

Prof. Dr. Kemal Leblebicioğlu
Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. Şebnem Düzgün
Geodetic and Geographic Information Technologies Dept. , METU

M.Sc. Ayhan Yılmaz
Aydın Yazılım ve Elektronik San. A.Ş.

Date: 03 June 2008

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Mustafa Yağcıođlu

Signature:

ABSTRACT

COMPARISON OF 3D FACIAL ANCHOR POINT LOCALIZATION METHODS

YAĞCIOĞLU, Mustafa

M.Sc., Department of Electrical and Electronics Engineering

Supervisor: Asst. Prof. İlkay ULUSOY

May 2008, 52 pages

Human identification systems are commonly used for security issues. Most of them are based on ID card. However, using an ID card for identification may not be safe enough since people may not have any protection against the theft. Another solution to the identification problem is to use iris or fingerprints. However, systems based on the iris or fingerprints need close interaction to identification machine. Identifying someone from his photograph overcomes all these problems which can be called as face recognition.

Common face recognition systems are based on the 2D image recognition but success rates of these methods are strictly depending on the environment. Variations on brightness and pose, complex background are the main problems for 2D image recognition systems. At this point, three dimensional face recognition techniques gain importance. Although there are a lot of methods developed for 3D face recognition, many of them assume that face is not rotated and there is not any destructive (i.e. beard, moustache, hair, hat, and eyeglasses) on the face. However, identification needs to be done though these destructives. Basic step for the face recognition is the determination of the anchor points (i.e. nose tip, inner eye points). In this study, the goal is to implement previously proposed four face recognition methods based on anchor point detection; “Multimodal Facial Feature Extraction for Automatic 3D Face Recognition” ,

“Automatic Feature Extraction for Multiview 3D Face Recognition”, “Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression”, “3D face detection using curvature analysis”, to compare the success rates of them for rotated and destructed images and finally to propose improvements on these methods.

Keywords: 3D Face Recognition, Anchor Point Detection

ÖZ

**ÜÇ BOYUTLU YÜZ NİRENGİ NOKTALARI BULAN
METODLARIN KARŞILAŞTIRILMASI**

YAĞCIOĞLU, Mustafa

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Asst. Prof. İlkey ULUSOY

Mayıs 2008, 52 sayfa

Kontrollü insan giriş çıkışlarının otomatik olarak yapıldığı durumlarda en yaygın uygulama kimlik kartı kullanımıdır ancak bu tür sistemler hırsızlık ve kartın farklı kişilerce kullanılması gibi pek çok problemin oluşmasına çok açıktır. Son yıllarda biyometrik tabanlı sistemler hem kontrollü otomatik giriş çıkışlarda hem de güvenlik amacıyla yaygın olarak kullanılmaktadır. İris veya parmak izinden tanımayı sağlayan sistemler yakın temas gerektirdiği için genelde tercih edilmezler. Yüz tanıma tabanlı sistemler ise insanlar tarafından rahatsız edici bulunmamaktadır. Yaygın olan yüz tanıma sistemleri iki boyutlu fotoğraflardan tanıma amaçlıdır ve iki boyutlu sistemlerin başarıları buldukları ortama çok bağlıdır. Değişen ışıklandırma, karmaşık arka plan ve değişen poz tanımayı doğrudan ve çok ciddi şekilde etkileyebilen değişkenlerin başında gelir. Bu noktada üç boyutlu yüz tanıma sistemleri önem kazanır. Bu alanda pek çok yöntem geliştirilmiştir ancak pek çoğu veri alınırken insanın pozisyonunun üç boyutlu tarayıcıya göre sabit olduğunu ve yüzde yüz şeklini engelleyici (sakal, bıyık, saç, şapka, gözlük, sağa, sola, yukarı veya aşağı dönerek poz verme) bir durumun olmadığını varsayar. Oysa, bu tür engelleyici durumlar için de yüz tanıma yapılabilirdir. Bunun ilk aşaması ise yüzün veya yüzdeki bazı noktaların yerlerinin tarayıcı çıktısında belirlenmesidir. Bu çalışmada, yüz ve yüz noktalarının bulunması için literatürde var olan dört yöntem; “Multimodal Facial Feature Extraction for Automatic 3D Face Recognition” , “Automatic Feature Extraction for Multiview 3D Face Recognition”,

“Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression”, “3D face detection using curvature analysis”, deęişik pozlar için karşılaştırılmış, bu yöntemlerin karşılaştıkları problemler için çözümler önerilmiştir.

Anahtar Kelimeler: 3 Boyutlu Yüz Tanıma, Yüz Nirengi Noktaları

To my Wife and Son

ACKNOWLEDGEMENTS

I would specially like to thank to my supervisor Assoc. Prof. Dr. İlkey ULUSOY for her encouragements throughout the thesis. It was a real pleasure and chance to work with her.

I would also like to thank to Ayhan Yılmaz and Erdem Akagündüz for their devoted technical support.

Lastly, I would like to express my deepest appreciation to my son Mehmet and my wife Derya who always gave me her love and emotional support.

TABLE OF CONTENT

ABSTRACT	IV
ÖZ	VI
ACKNOWLEDGEMENTS	IX
TABLE OF CONTENT	X
LIST OF TABLES	XI
LIST OF FIGURES	XII
CHAPTERS	
1. INTRODUCTION	1
1.1. PROBLEM DEFINITION	1
1.2. SCOPE OF THE THESIS	3
1.3. OUTLINE.....	4
2. DESCRIPTION OF THE METHODS	5
2.1. MULTIMODAL FACIAL FEATURE EXTRACTION FOR AUTOMATIC 3D FACE RECOGNITION	5
2.1.1. <i>Nose Tip Extractions</i>	5
2.1.2. <i>Inner Eye Point Detection</i> :.....	10
2.2. AUTOMATIC FEATURE EXTRACTION FOR MULTIVIEW 3D FACE RECOGNITION	13
2.2.1. <i>Nose Tip Estimation</i>	13
2.2.2. <i>Inner Eye Points Detection</i>	15
2.3. MULTIPLE NOSE REGION MATCHING FOR 3D FACE RECOGNITION UNDER VARYING FACIAL EXPRESSION	16
2.3.1. <i>Inner Eye Points Detection</i> :.....	16
2.3.2. <i>Nose Tip Detection</i> :	18
2.4. 3D FACE DETECTION USING CURVATURE ANALYSIS.....	18
2.4.1. <i>Determination of candidate eyes and noses</i>	18
2.4.2. <i>Generation of candidate face triangles by combining eyes and noses</i>	24
2.4.3. <i>Face registration by using actual face triangles</i> :	24
2.4.4. <i>Determining the actual face triangle by using PCA analysis</i> :	26
3. DATA SET	28
3.1. DATA SET	28
4. EXPERIMENTS & RESULTS	30
4.1. RESULTS	30
4.2. SUCCESS RATE ANALYSIS	32
4.3. COMPLEXITY ANALYSIS	41
5. CONCLUSIONS	47
5.1. CONCLUSIONS.....	47
5.2. PROPOSED FUTURE WORK	48
REFERENCES	49

LIST OF TABLES

Table 1 Detection percentage for the nose and the inner eye pits	31
Table 2 Detection errors, examined for the nose and the inner eye pits	32

LIST OF FIGURES

Figure 1 Hair is the global maximum point	5
Figure 2 3D representation of the image that is the global maximum point is the hair.....	6
Figure 3 Highest Z value points in each row are connected	7
Figure 4 Histogram for selecting midline.	7
Figure 5 Vertical profile of the image on midline	8
Figure 6 Histogram of the non-decreasing paths on midline	9
Figure 7 Horizontal profiles for the three nose tip candidates	10
Figure 8 Possible locations of the inner eye points	11
Figure 9 Founded nose tip and inner eye points are illustrated.....	12
Figure 10 Pose angle quantization	14
Figure 11 Candidate nose tips are signed by small points and the most frequently observed candidates are signed with big points.....	14
Figure 12 Gaussian curvature of an image.....	16
Figure 13 Mean Curvature of an image	17
Figure 14 Possible Inner eye points after Mean and Gaussian Classification	19
Figure 15 Possible Nose Tips after Mean and Gaussian Classification.....	20
Figure 16 Nose candidate regions after eliminating small regions.....	21
Figure 17 Eye candidate regions after eliminating small regions	21
Figure 18 For nose candidates, each region is colored in a different gray level	22
Figure 19 For eye candidates, each region is colored in a different gray level.....	22
Figure 20 Possible candidate points for nose	23
Figure 21 Possible candidate points for eyes	23
Figure 22 a) Range representation of a face and b) The region extracted around nose..	25
Figure 23 The image obtained after the reproduction by using the eigenvectors of the face	26
Figure 24 The difference between the original image and the reconstructed image	27
Figure 25 Mouth stretch (a), disgust(b), Inner brow raiser (c), happiness(d), yaw rotations of +30 ° (e), +45 ° (f), +60 ° (g), +75 ° (h), pitch rotations of strong upwards (i), slight upwards(j), slight downwards(k), strong downwards(l), bottom right(m), upper right(n), eye occlusion(o), mouth occlusion(p), eye glasses(r).	29
Figure 26 Success of the four methods for nose and eyes detection	33
Figure 27 Success rate percentages of four methods	34
Figure 28 Z values on the selected line are increased.....	36
Figure 29 Destruction around eyes	36
Figure 30 Z values are increased around the eyebrows	37
Figure 31 Puffiness in cheeks.....	37
Figure 32 Hand on Eye	39
Figure 33 Hand on Mouth.....	40
Figure 34 Eyeglasses.....	40

CHAPTER 1

INTRODUCTION

1.1. Problem Definition

Human identification is very important for security issues. There are a lot of identification techniques for recognizing people. Some of these recognition techniques can be listed as signature, face recognition, voice, and fingerprint. Akarun et al.'s, investigated the performance of the nine identification techniques (Signature, 2D Face, Voice, Iris, Fingerprint, Retina, Gait, Hand, 3D Face) and decided that 3D Face recognition is the best among them because of its cost and reliability.

Up to the 3D sensors, 2D face recognition systems are used [12], [23], [16], [9], [27]. Against to 3D systems, 2D face recognition systems increased the speed of the recognition process. Boehnen and Russ [12] presented a method utilizing the registered 2D color image of a face. Turk and Pentland [23] developed a face recognition system using Eigenfaces technique.

Principal Component Analysis (PCA) is widely used in 2D face recognition systems. The purpose of PCA is to reduce the large dimensionality of the data space to the smaller dimensionality of feature space to describe the data economically [18]. Zhang and Zhou [16] used an alternative PCA technique, which is called as 2DPCA. Main idea behind 2DPCA is to utilize 2D matrices as opposed to the standard PCA, which is based on 1D vectors [16]. Considering a performance criteria for face recognition by PCA, Feng et al.'s [9] applied PCA on wavelet subband to decrease the computational load of PCA.

2D face recognition systems have promising performance under frontal pose but they encounter difficulties for large variations of pose or illuminations. 3D face recognition methods overcome the challenges caused by different head pose, illumination and expression. Iterative Closest Point (ICP) is an algorithm employed to

match two clouds of points. When data is a 3D point cloud, Iterative Closest Point algorithm is commonly used for face recognition [6], [24], [25], [26], [8]. ICP is used by Niese, Al-Hamadi and Michaelis [13] to recognize the pose of the face and normalize it.

Thanks to the improvements in 3D sensors and faster processors, three dimensional face recognition techniques are frequently used in face recognition systems. Although recognition by 3D techniques takes longer time compared to 2D systems, it is still acceptable.

Superiority of the 3D recognition systems is mainly due to utilizing depth information of the images. By using depth information, surface properties of the image can be determined. Since surface curvatures do not change with pose, 3D facial detection systems use surface curvatures to classify surfaces. Dorai and Jain [5] proposed the term “Shape Index” to represent the surface curvature at each point of the image. Shape index is commonly used in 3D recognition algorithms. Besides the shape index, Mean and Gaussian Curvature properties of the image is used for determining the surface properties [27] and [28].

To increase the success rate of face recognition, some methods combine 2D and 3D techniques [10], [11], [21]. Sun at al.’s [10] used PCA for both 2D and 3D data and combined them to increase the success rate. Besides that, Mian at al.’s proposed a 2D and 3D multimodal hybrid face recognition algorithm [11].

Since there are a lot of methods proposed related with face recognition, many surveys and comparisons are proposed related with this issue. Kittler at al’s [29] proposed a survey on 3D face recognition techniques. Bowyer at al’s [30] focuses on face recognition using three-dimensional data, either alone or in combination with two-dimensional intensity images. Chang at al.’s [20] compared 2D, 3D, 2D+3D methods. They claim that although 2D and 3D have similar recognition performance when considered individually, performance of 2D+3D face recognition is significantly better than that of 3D only or 2D only systems.

Face recognition systems can be divided into two approaches. One approach uses the whole face data and uses the data directly for discriminating faces [13],[15] . PCA is commonly used for this approach. Yambor, Draper, and Beveridge examined the role of Eigenvector selection and Eigenspace distance measures on PCA-based face recognition

systems [19]. Heshner [17] explored PCA approach using different numbers of eigenvectors and image sizes.

Second approach for face recognition investigates the relations between the anchor points. Recognition is based on the relationship between human facial features such as eyes, mouth, nose, profile silhouettes and face boundary. Generally, nose tip and inner eye points are selected as anchor points because these points are motionless under varied facial expressions. Success of this approach highly depends on the success of the anchor point detection algorithm. An error in extracting a single anchor point makes it impossible to recognize face depending on the anchor points.

Colbry, Stockman, and Jain [14] used “shape index” property of the image for determining the anchor points. Lu and Jain [2] rotated the image around Y axis and detected the nose tip as the frequently obtained closest point to the camera in all rotations. After finding the nose tip, possible locations of the two inner eye pit points are calculated statistically. In the search region, inner eye pits are selected by shape index and cornerness measurements. In another method of Lu and Jain [1], nose tip is found out by rotating the image many times around y-axis. Lee et al.’s performed 3D face recognition by locating the nose tip and forming a feature vector which is based on contours along the face at a sequence of depth values [7]. Chang and Bowyer [3] used shape index property for detection of inner eye points and extracted the nose tip by using the previously detected inner eye point locations. Colombo et al.’s [4] calculated anchor point candidates by using shape index properties and used PCA to determine which ones are the exact anchor points. Similar to [4], Nagamine [22] detected five feature points and used those feature points to standardize face pose, then finally matched the various curves to the face data.

1.2. Scope of the Thesis

Anchor point detection accuracy is very important for the face recognition methods. [1] and [2] determine the nose tip first and then the inner eye points. [3] calculates the inner eye points first and finds the nose tip by using the detected eyes. Unlike [1],[2] and [3], [4] combines the nose and eye candidates and detect the correct combination by using predefined relations between these points.

Although there are a lot of methods proposed for the detection of anchor points, there is no information about the comparison of the success rates and computational complexities of these methods. In this study, our purpose is to implement some representative methods [1], [2], [3], [4] from the literature and compare them with an objective view, by quantitative analysis. We not only compare them in terms of detection rates but also in terms of localization accuracy and computational load. Besides, weak and strong points of these methods are also described and some improvements for each of them are proposed.

1.3. Outline

Introduction is given within this chapter (Chapter 1).

Chapter 2 describes the details of the methodologies. In Section 2.1 “Multimodal Facial Feature Extraction for Automatic 3D Face Recognition” [1] is described. Section 2.2 describes the details of the methodology described in “Automatic Feature Extraction for Multiview 3D Face Recognition.” [2]. “Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression” [3] is described in Section 2.3. Last method “3D face detection using curvature analysis”[4] is described in Section 2.4.

Chapter 3 describes basic properties of the data used by these methods. Number of the images used to compare success rates and the methodology of comparing the results are also given in chapter 3.

In chapter 4, results of the implemented methods are given and a brief discussion is presented on the results and success rates of the methodologies. One of the methodologies is decided as the best of the tested methods. Moreover, computational complexities of these four methods are analyzed in this chapter.

CHAPTER 2

DESCRIPTION OF THE METHODS

2.1. Multimodal Facial Feature Extraction for Automatic 3D Face Recognition

First approach to determine anchor point is described in [1]. In this method, a robust nose tip locator is presented and a statistical 3D feature location model is applied after aligning the model with nose tip. Positions of the inner eye points are determined by combining the shape index response and cornerness response.

2.1.1. Nose Tip Extractions

Generally, in frontal scans, nose tip has the maximum value in Z direction. Although the nose tip always satisfies to be a local maximum point, it cannot be guaranteed that nose tip is the global maximum point due to the factors such as bear, hair or noise. In Figure 1, an image is illustrated with the hair has the maximum z -value. Figure 2 illustrates the same image along x -axis.

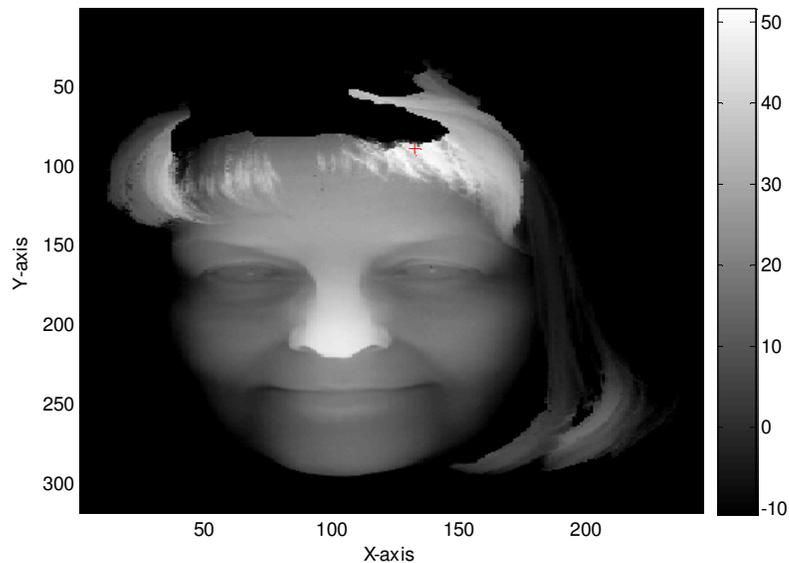


Figure 1 Hair is the global maximum point

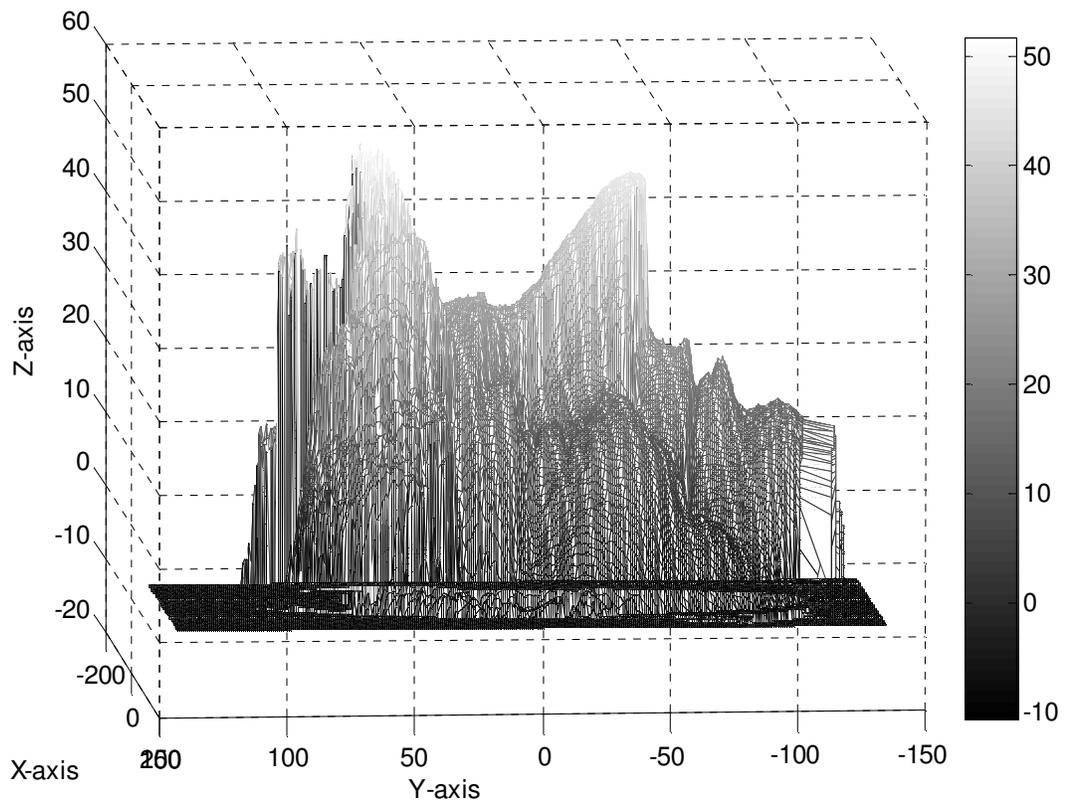


Figure 2 3D representation of the image that is the global maximum point is the hair

In the frontal pose, nose tip is located on the mid-line that is created by connecting the points which has maximum z values. For each row, the position with maximum z value is determined, as shown in Figure 3. Then for each column, numbers of these maximum points are counted and a histogram is drawn. See Figure 4 for the histogram. Mid-line is chosen as the column that has maximum number of maximum z value. (It can be defined as the column at which the histogram reaches the peak). The value that the histogram reaches the peak is selected as midline.

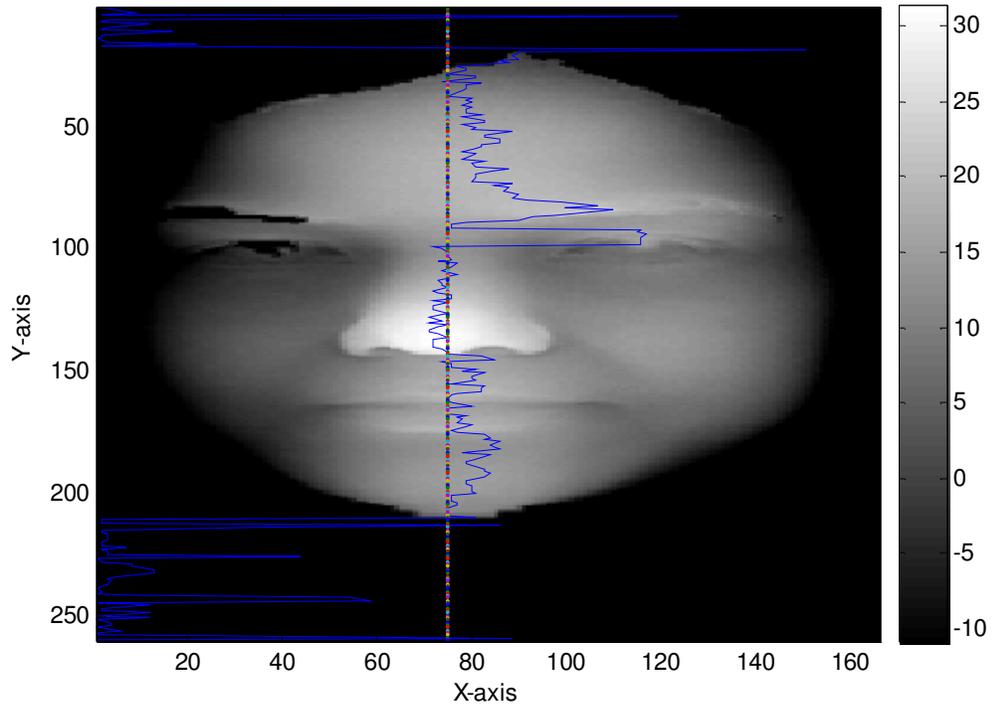


Figure 3 Highest Z value points in each row are connected

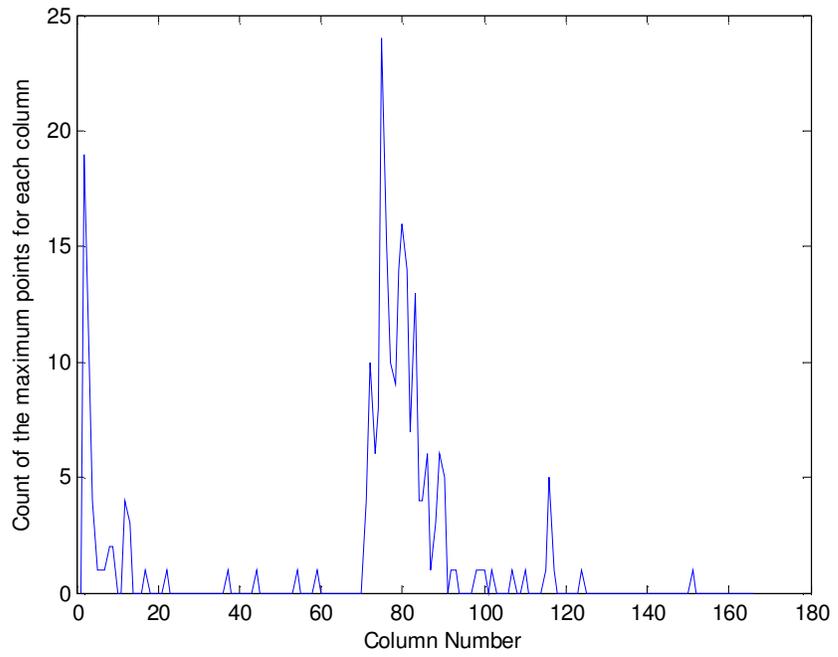


Figure 4 Histogram for selecting midline.

Now, the search region is limited with the mid-line. The vertical z profile analysis along the midline is illustrated in Figure 5. Nose Bridge presents a consecutive increase in z values. The gradients can be calculated by the equation (1) where r is the row index. Figure 6 shows the consecutive increase in z values on mid-line.

$$g(r) = z(r+1) - z(r) \tag{1}$$

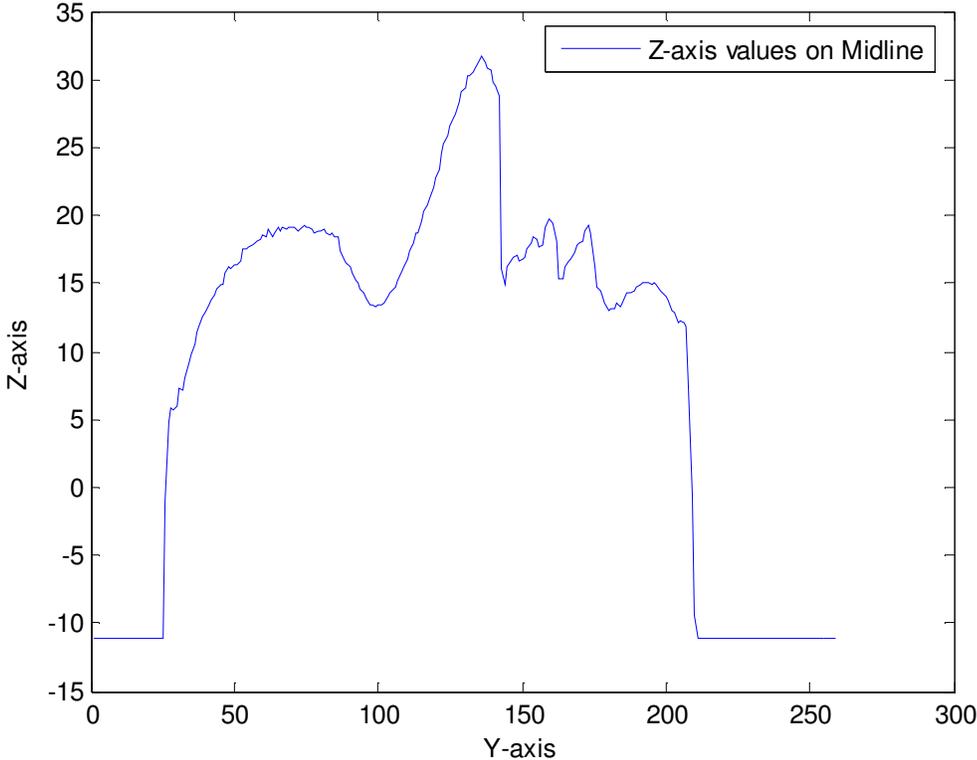


Figure 5 Vertical profile of the image on midline

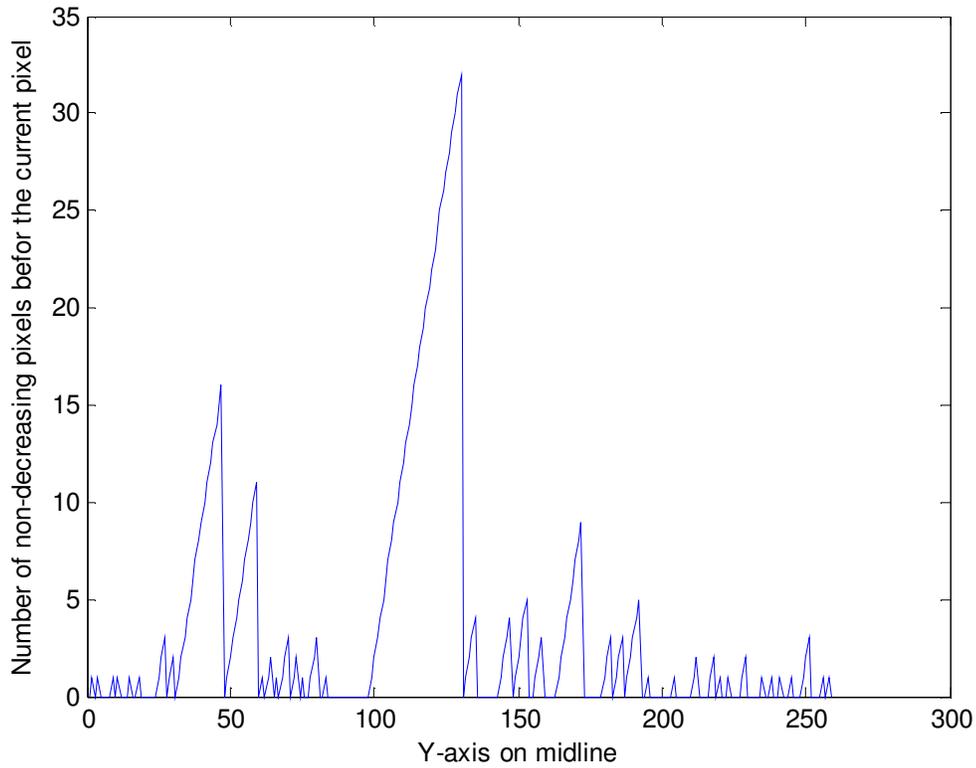


Figure 6 Histogram of the non-decreasing paths on midline

Three highest peaks illustrated in Figure 6 are the candidate nose tip points. But which one is the real nose tip? To determine the nose tip, profile analysis is performed for all three candidate points along each row, i.e., the horizontal z profile analysis.

Corresponding row profiles for candidate nose tip are illustrated in Figure 7. Among three candidate nose tip, the one which present most sharpness characteristic is determined as nose tip. Sharpness is calculated by variations of the row profile with respect to each nose candidate. For the image shown in Figure 7, selected nose tip is signed.

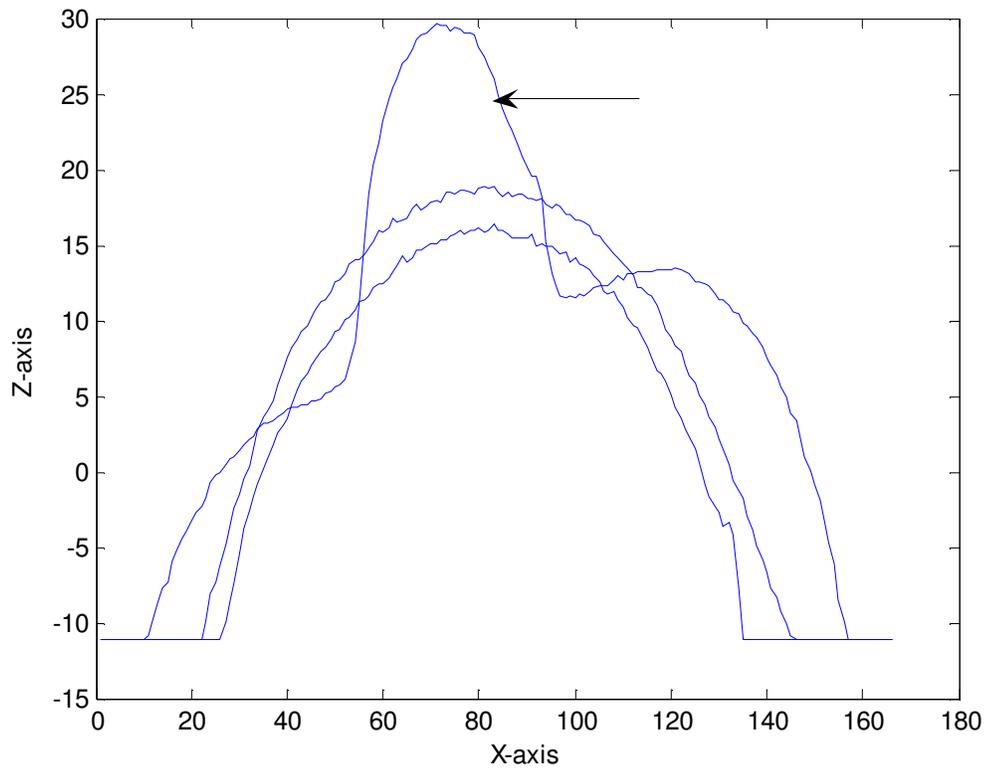


Figure 7 Horizontal profiles for the three nose tip candidates

2.1.2. Inner Eye Point Detection:

To reduce the search area for inner eye points, a statistical model is used. Gaussian mixture model is used to determine the ellipsoid that the anchor points are possibly in. Nose tip and inner eye points of 40 frontal images are used as the input of Gaussian mixture model. Figure 8 shows the ellipsoid for a test image.

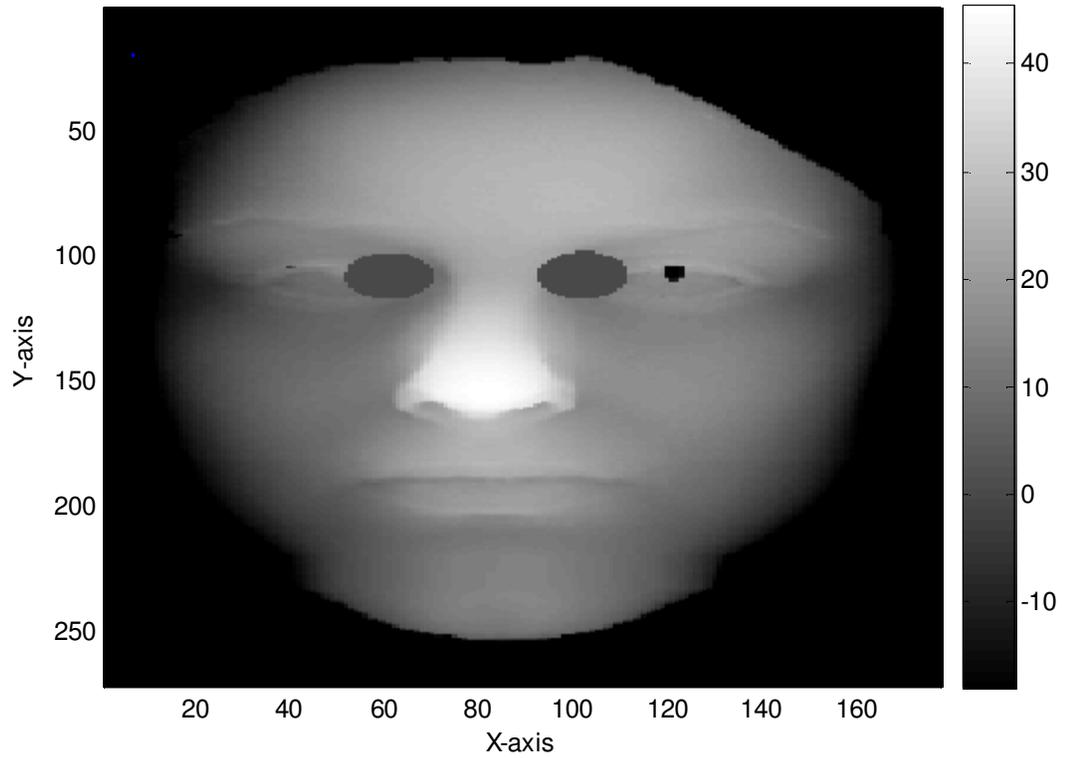


Figure 8 Possible locations of the inner eye points

Inner eye points are determined by combining the shape index response and cornerness response.

Initially, normalized shape index response and cornerness response are calculated for an image. Final score of each point in the image is determined by integrating each score by using the sum rule as shown in (2)

$$F(p) = (1 - S(p)) + C(p) \quad (2)$$

In each search region, the point with the highest $F(p)$ is identified as the corresponding feature point. Figure 9 shows an example of the founded nose tip and inner eye points.

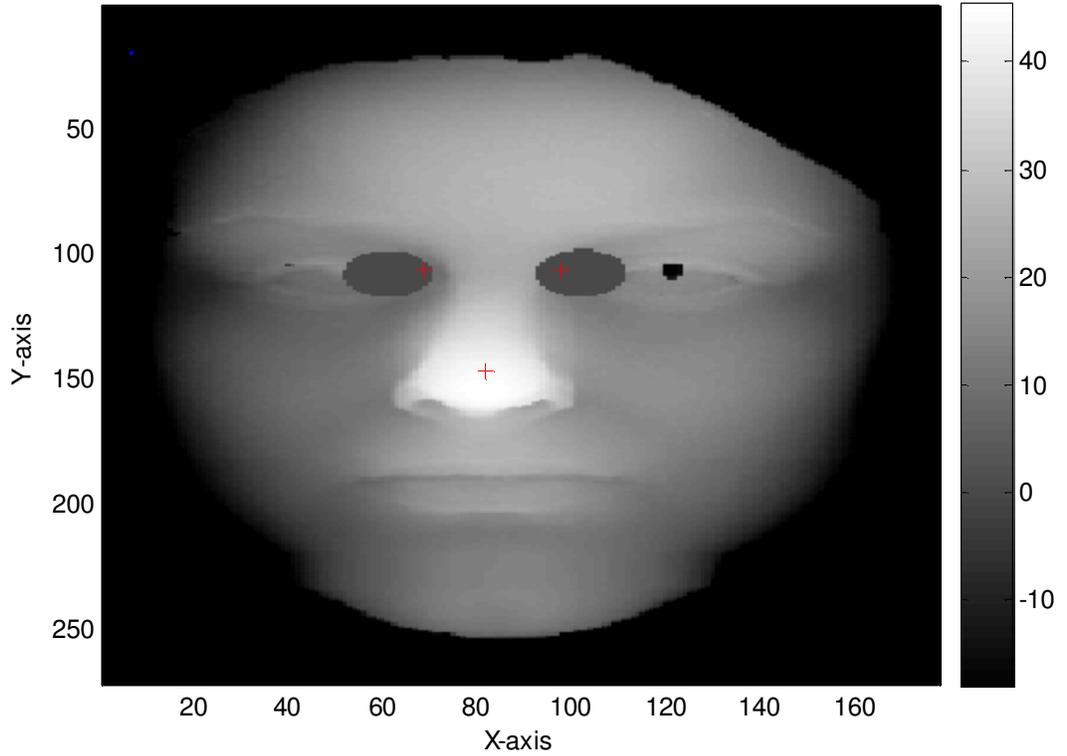


Figure 9 Founded nose tip and inner eye points are illustrated

Techniques used for determining inner eye points are described below. For more detail, see [1]

2.1.2.1. Shape index:

Shape index of an image is derived at each point based on the range map. It is calculated using the maximum (k_1) and minimum (k_2) local curvature values. See (3)

$$S(p) = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \quad (3)$$

2.1.2.2. Cornerness:

The inner eye points form a strong corner-like pattern. Harris corner detector is used to obtain inner eye points. In this thesis, (4) is used to calculate cornerness response.

$$C(p) = \frac{\frac{\partial^2 I}{\partial x^2} \frac{\partial^2 I}{\partial y^2} - \left(\frac{\partial^2 I}{\partial x \partial y} \right)^2}{\frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}} \quad (4)$$

2.1.2.3. Fusion:

Information obtained from Shape index and Cornerness response is combined to have a better result on obtaining inner eye points. In order to do this, both $S(p)$ and $C(p)$ are normalized. After the normalization, final score is calculated as (5).

$$F(p) = (1 - (S'(p))) * C'(p) \quad (5)$$

2.2. Automatic Feature Extraction for Multiview 3D Face Recognition

2.2.1. Nose Tip Estimation

Although the previous method is powerful for frontal scans, its success rate tends to decrease on rotated images. In order to obtain a better result on rotated images, this method is proposed. It also assumes that the nose tip is the closest point to the screen in frontal scans. However, it does not assume that the sample image is frontal pose image. Following steps describe the methodology for finding anchor points.

2.2.1.1. Pose quantization:

Image is rotated with respect to the Y axis from -90 degrees to 90 degrees. This 180 degree range is quantized into $N_{pose} = 91$ angle with interval of $\partial\theta = 2$ degrees. Figure 10 illustrates the quantization for angles = 90, 60, 30, 0, -30, -60 and -90.

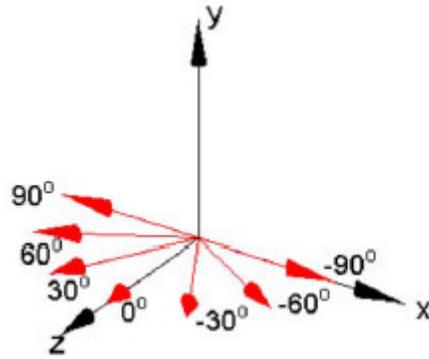


Figure 10 Pose angle quantization

2.2.1.2. Directional maximum:

At each pose angle θ_j ($j = 1, \dots, N_{pose}$), maximum projection value along the corresponding pose direction is determined as nose tip candidate. For N_{pose} angles, M (M is less than N_{pose}) candidate points are detected because of the multiple candidate points. Figure 11 illustrates the distribution of the M candidate nose tips. It is obvious that directional maximum may happen with the same face point P at multiple θ s. In such case, the angle with the largest projection value is selected as the pose angle to be associated with the point P .

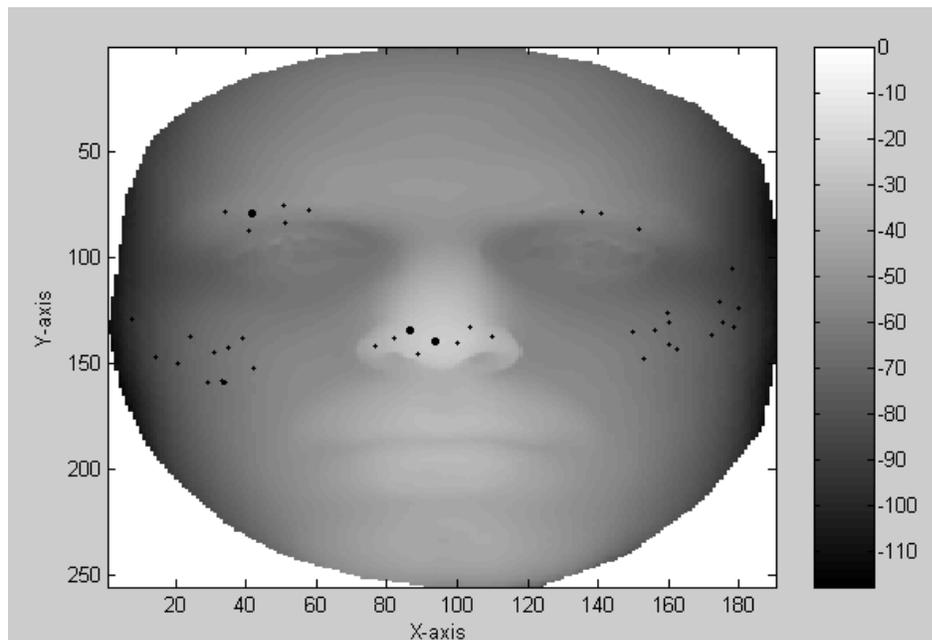


Figure 11 Candidate nose tips are signed by small points and the most frequently observed candidates are signed with big points.

2.2.1.3. Pose Correction:

For the most probable, that is most frequently obtained, three nose tip candidate pairs, the coordinates (x, y, z) of all the original face points are transformed to (x', y', z') so that point p is at the origin, and the face points are rotated according to the pose angle θ . After the rotation, face is changed to a frontal scan face.

2.2.1.4. Nose Profile Extraction:

From the pose corrected images for the three candidate nose tip pairs, nose profile at p (remember that p is the origin of the new coordinate system) is the intersection between the facial surface and the Y-Z plane. Let $X'(r, c)$, $Y'(r, c)$, and $Z'(r, c)$ denote the point coordinate matrices after pose correction. For each row r_i , find the closest point to Y-Z plane and construct a line by combining these points. PCA analysis, described in section 2.2.1.5., is applied to this line.

2.2.1.5. Nose Profile Identification:

To identify the nose tip from the candidate points, PCA analysis is applied to candidate nose tips in [2]. However, in this thesis, horizontal z profile analysis is applied as in [1] instead of PCA, since after applying horizontal z profile analysis, it is observed that if the actual nose tip is one of the candidates it is detected successfully.

2.2.2. Inner Eye Points Detection

Inner eye points are detected by using the Shape index and Cornerness response of the images. Gaussian Mixture Model is used to decrease the search region and increase the success rates. In addition, since the rotation angle of the sample image is known, GMM is constructed using the sample images rotated with the same angle with the sample image. As a result of this, although the image is rotated with high angles, inner eye points are detected successfully if the nose tip is successfully determined.

2.3. Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression

2.3.1. Inner Eye Points Detection:

Apart from other methods, this method first finds inner eye points. In order to achieve this, it only uses the mean curvature, H , and Gaussian curvature, K . A nose tip is expected to be a peak ($K > TK$ and $H > TH$) and a pair of eye cavities is to be a pair of pit regions ($K > TK$ and $H < TH$) where $TK = 0.0025$ and $TH = 0.04$. Mean and Gaussian curvature analysis for an image is illustrated in Figure 12 and Figure 13. Remember that for Figure 12 and Figure 13, if the values are bigger than the threshold value, it is set to 1 and set to 0 otherwise.

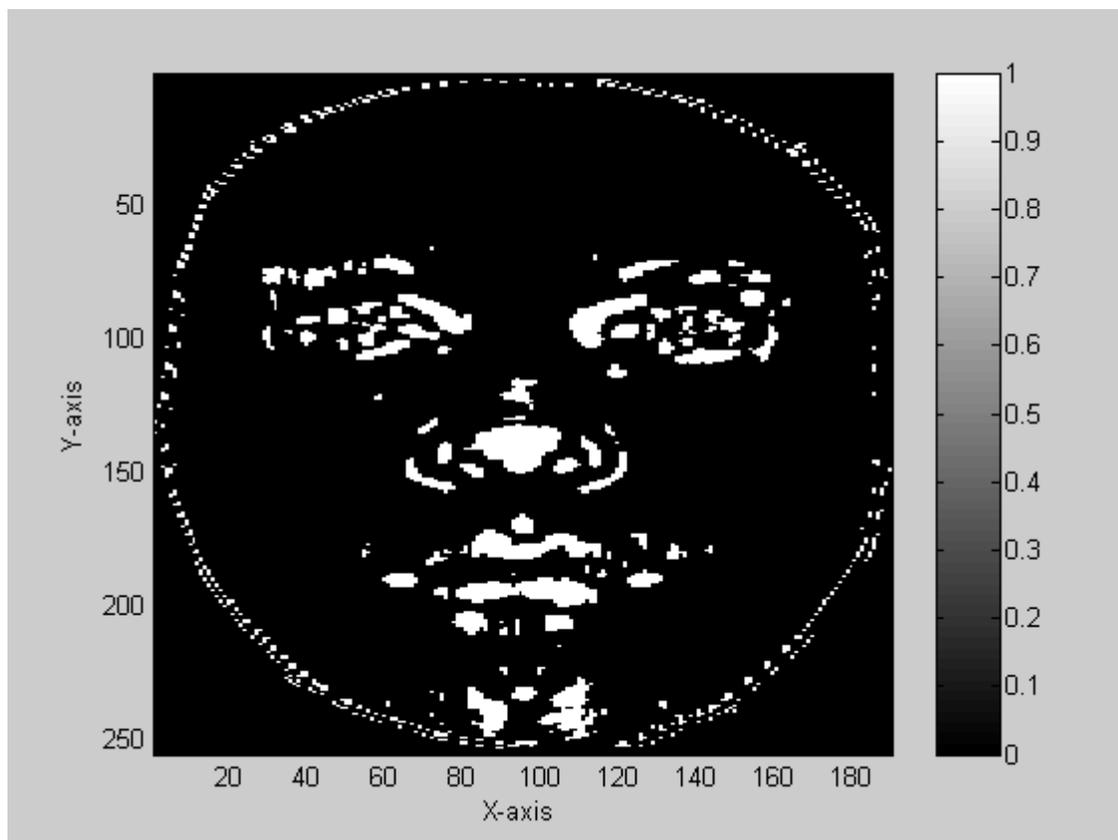


Figure 12 Gaussian curvature of an image

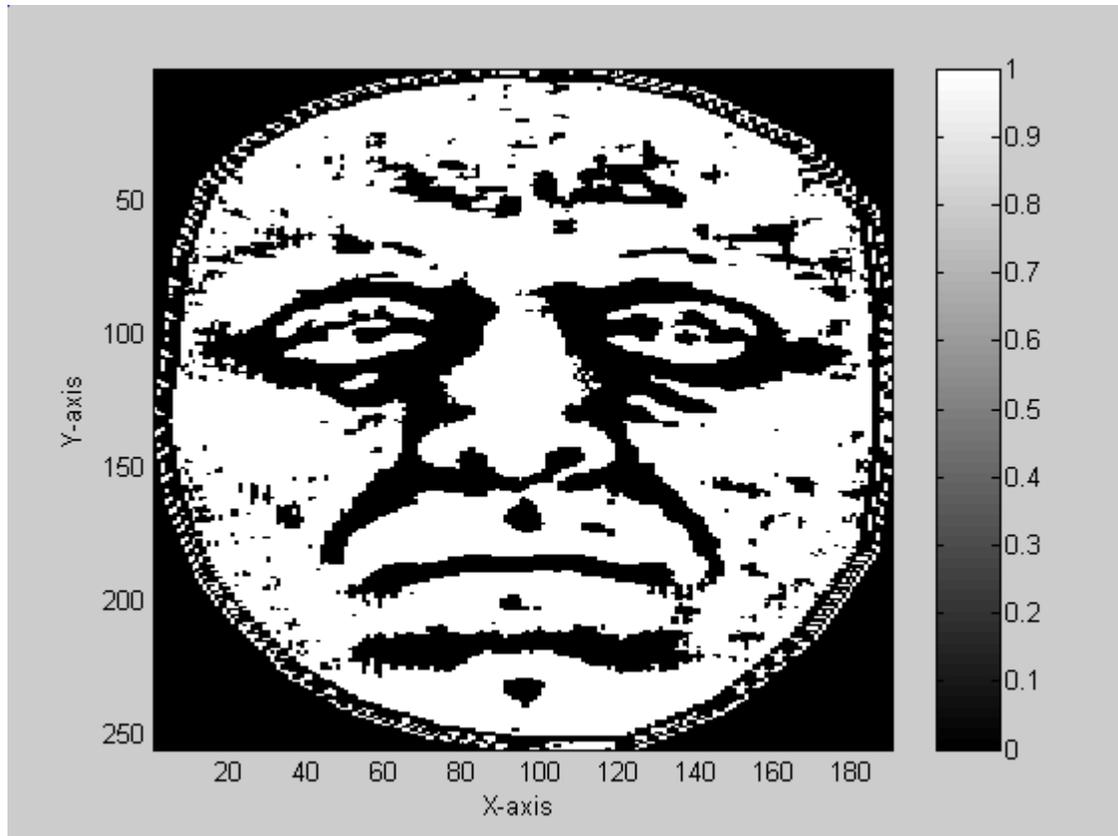


Figure 13 Mean Curvature of an image

By using the Mean and Gaussian curvature of an image, a combination image ZI is created as specified in (6).

$$ZI(i, j) = (K(i, j) > 0.0025 \& \& H(i, j) < 0.00005) \quad (6)$$

There may have many pit region candidates. To eliminate false pit regions, small pit regions are removed. This process is realized by filtering the Mean and Gaussian Curvature images with the filter specified in (7).

$$H = \begin{pmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix} \quad (7)$$

After observing the filtered image, a pair of points that has similar Y and Z values is chosen as the inner eye points. If there are more than one candidate pit regions, the ones with higher Y values are chosen.

Improvement: In this thesis, if the distance between inner eye point candidates is smaller than 15 mm, these candidates are eliminated.

2.3.2. Nose Tip Detection:

The nose tip is detected after finding the pit regions. Starting from the center of two inner eye points, the column is searched to find the point with the largest difference in Z value from the center of the pit regions.

2.4. 3D face detection using curvature analysis

Another approach to determine anchor point is proposed in [4]. Unlike the other three methods, this method determines the Nose-Eye-Eye triangle at the same time. Basic steps of this method can be summarized as;

1. Determination of candidate eyes and noses
2. Generation of candidate face triangles by combining eyes and noses
3. Face registration by using actual face triangles
4. Determination of actual face triangle by using PCA analysis

2.4.1. Determination of candidate eyes and noses

Mean (H) and Gaussian (K) curvature analysis is used for determining candidate noses and eyes. Candidate point detection method of M4 is similar with M3. To eliminate possible wrong nose and eye candidates, image is smoothed before Mean and Gaussian curvature analysis. Following filter is used for smoothing.

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 36 & 36 & 36 & 36 & 36 & 36 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 36 & 36 & 36 & 36 & 36 & 36 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 36 & 36 & 36 & 36 & 36 & 36 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 36 & 36 & 36 & 36 & 36 & 36 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 36 & 36 & 36 & 36 & 36 & 36 \end{pmatrix} \quad (8)$$

A nose tip is expected to be a peak ($K > TK$ and $H < TH$) and a pair of eye cavities to be a pair of pit regions ($K > TK$ and $H < TH$). Different from M3, TK and TH values are 0:005 and 0:04 respectively for M4. Possible Inner eye points are illustrated in Figure 14 and Nose tip candidates are shown in Figure 15.

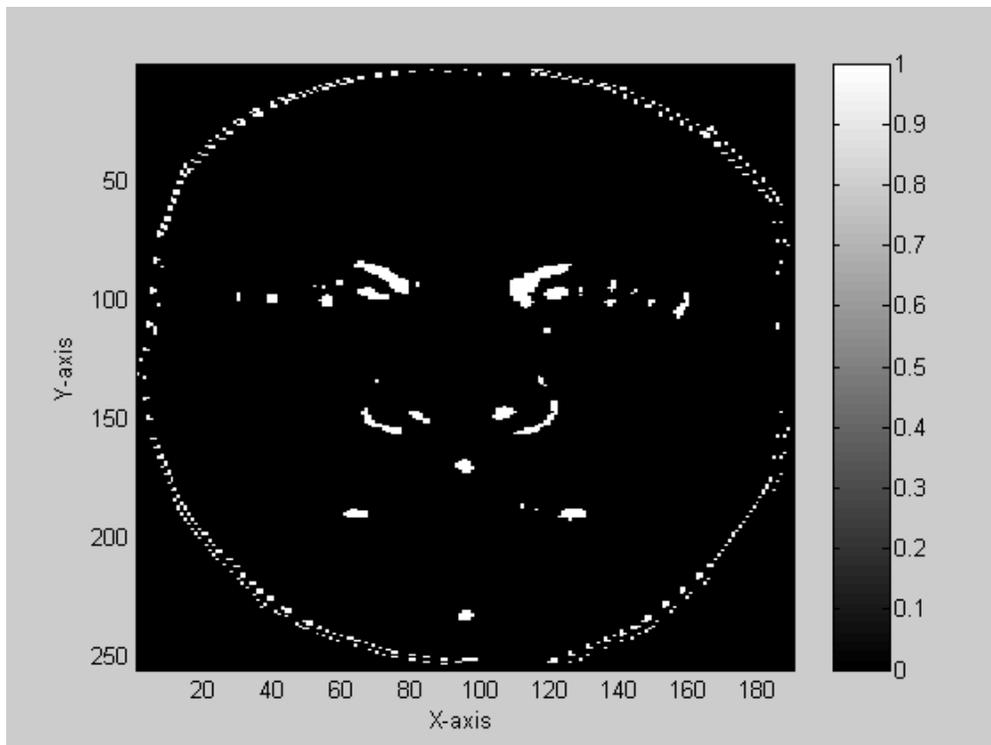


Figure 14 Possible Inner eye points after Mean and Gaussian Classification

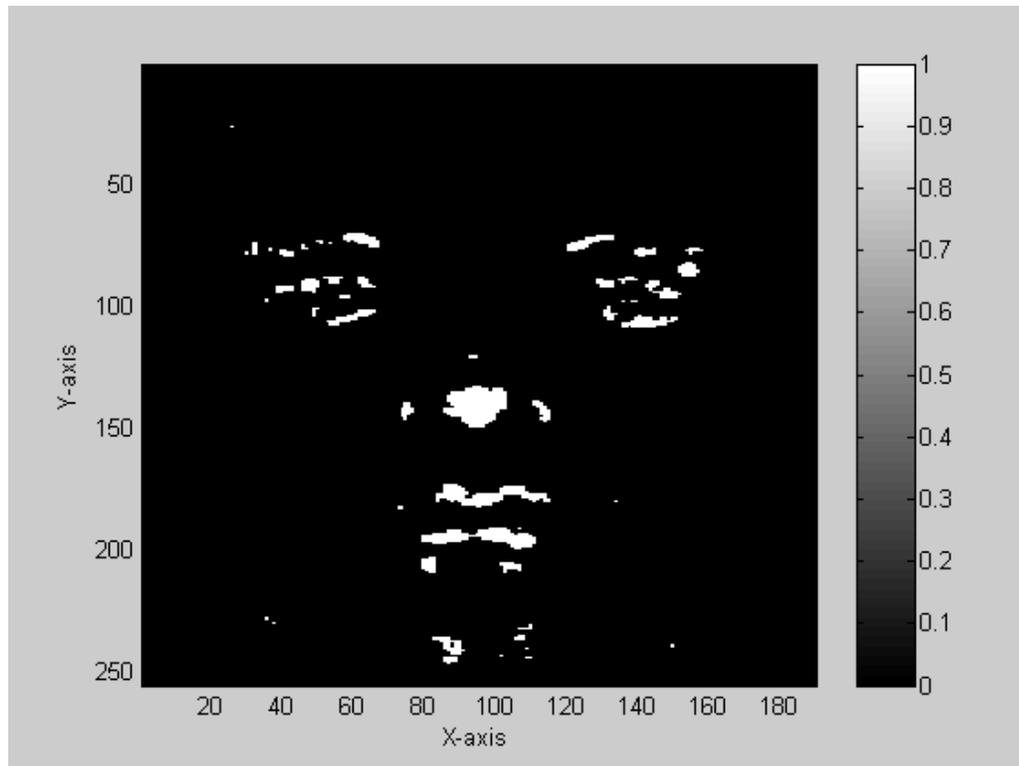


Figure 15 Possible Nose Tips after Mean and Gaussian Classification

As seen in Figure 14 and Figure 15, there are a lot of possible candidates. In order to decrease the number of candidate anchor points, small regions are removed from the candidate list. For nose, the regions with the number of connected component smaller than 30 pixels are removed. This threshold is 20 pixels for the eye candidates. After removing the small regions, Figure 16 and Figure 17 are obtained.

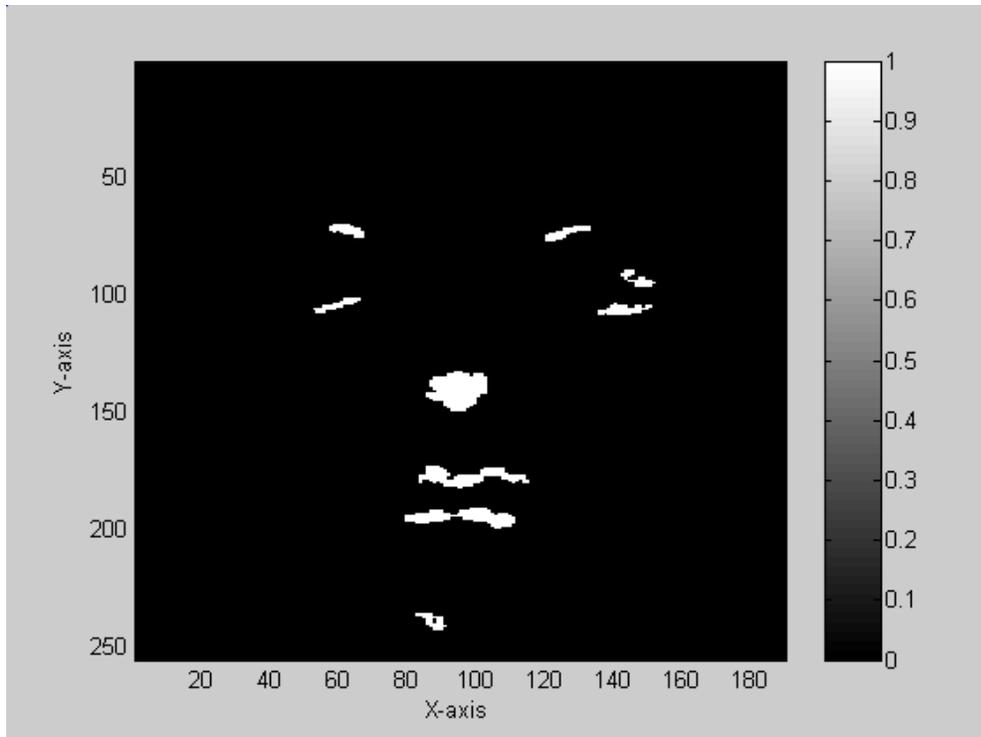


Figure 16 Nose candidate regions after eliminating small regions

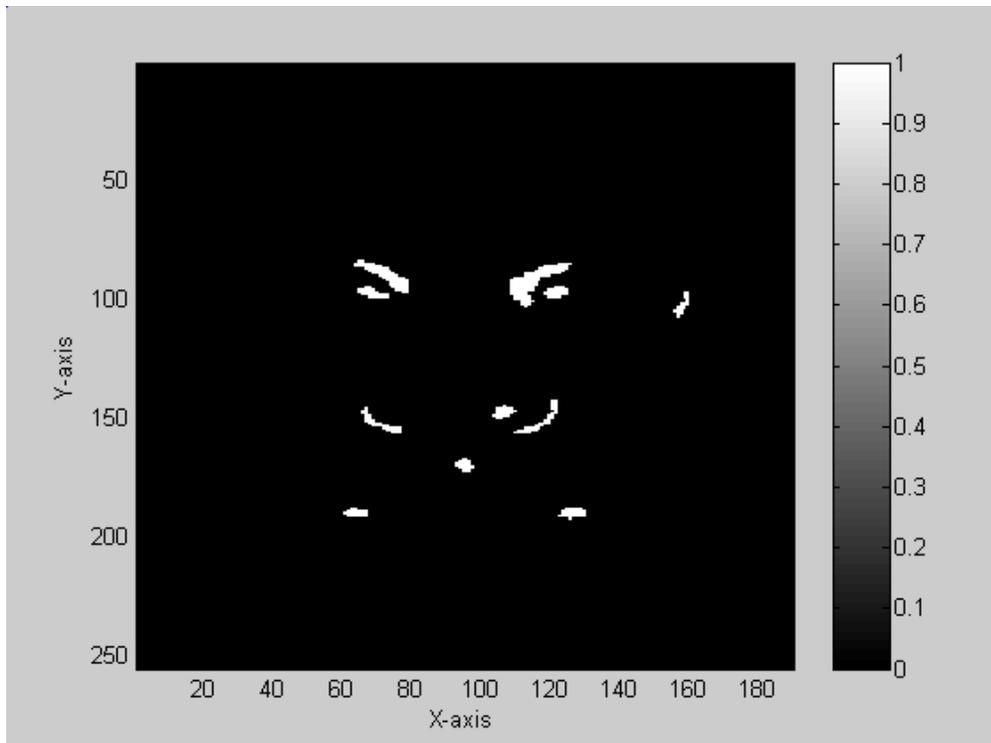


Figure 17 Eye candidate regions after eliminating small regions

Next step is to determine the center of masses for each region. Figure 18 and Figure 19 illustrate the steps for center of mass detection.

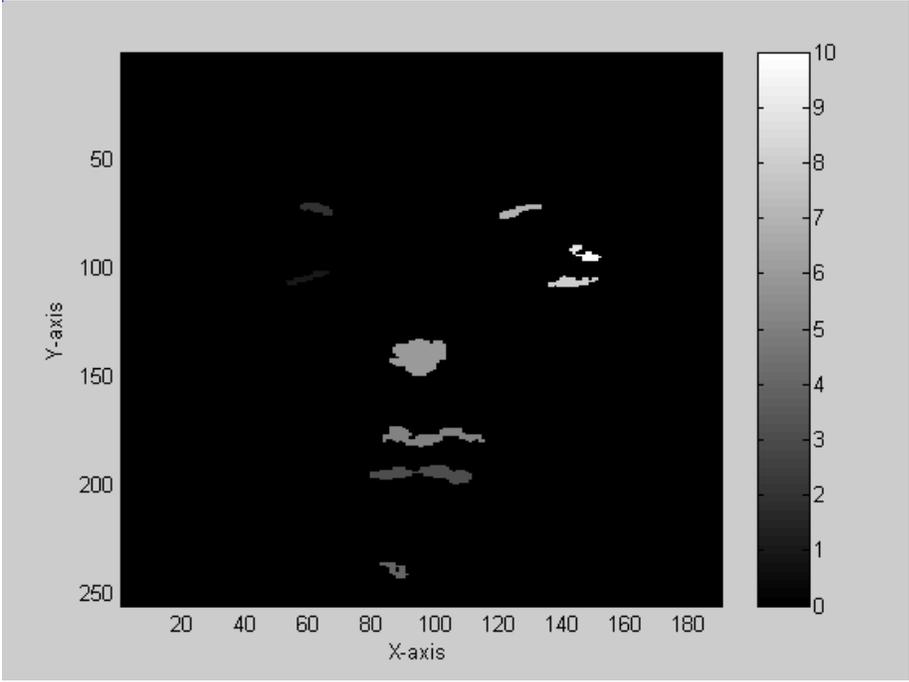


Figure 18 For nose candidates, each region is colored in a different gray level

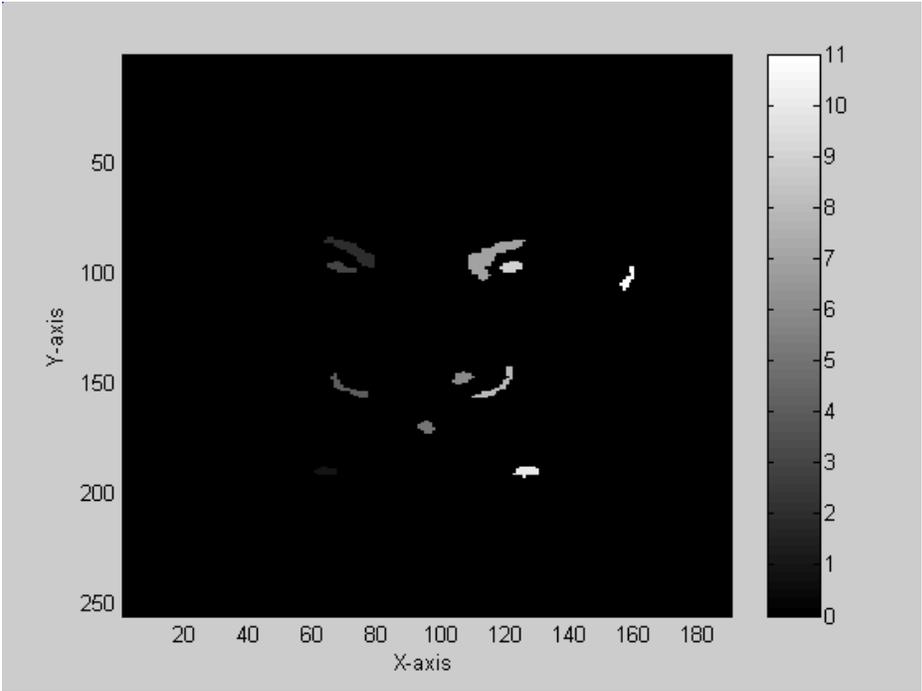


Figure 19 For eye candidates, each region is colored in a different gray level

In last step, Nose tip and Inner eye point candidates are obtained by calculating the center of masses for each region. Figure 20 and Figure 21 illustrate the center of masses of the candidate regions. These points are used as input for creating face triangles.

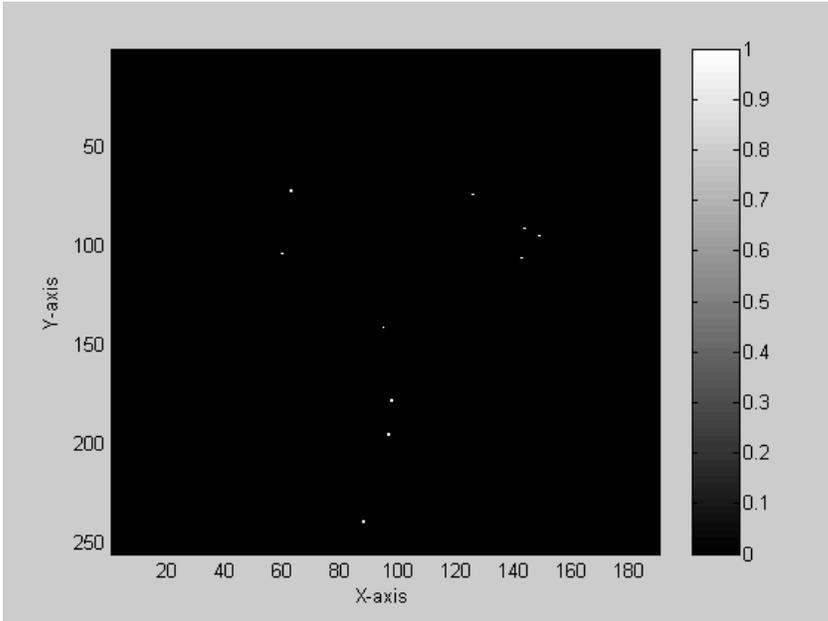


Figure 20 Possible candidate points for nose

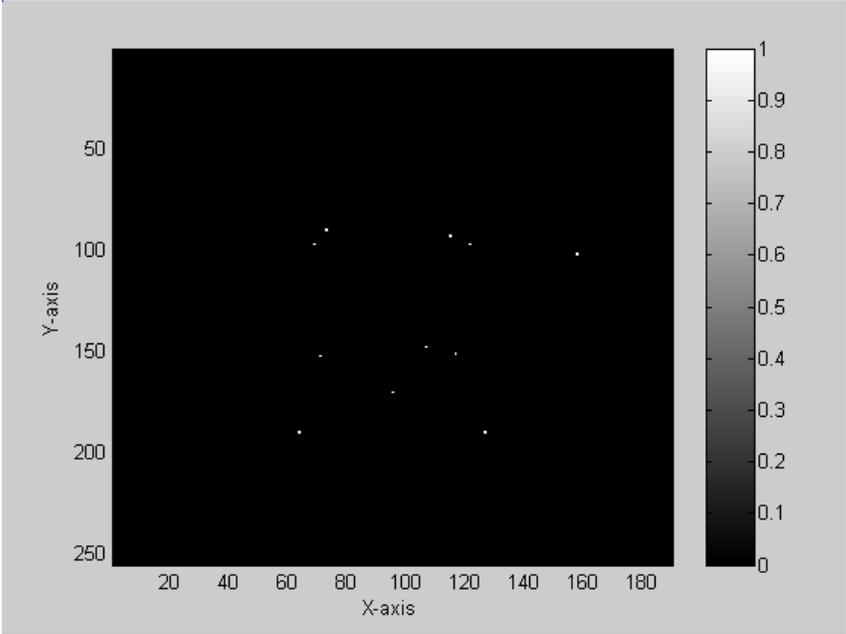


Figure 21 Possible candidate points for eyes

2.4.2. Generation of candidate face triangles by combining eyes and noses

Let us assume that we have p nose candidates and t eye candidates. By using these candidates pt^2 face triangle can be created. However, how many of these triangles are meaningful? Is it possible to have a distance between two eyes smaller than 10 mm? To reduce the computational overhead of the following steps, face triangles with abnormal distances are rejected. The distance between two eye points shall be bigger than 10 mm and smaller than 50 mm. Similarly, the distance between an inner eye point and the nose tip shall be between 15 mm and 80 mm.

After eliminating the impossible triangles, n face triangles are obtained. (n is probably smaller than pt^2)

2.4.3. Face registration by using actual face triangles:

Next step is to choose the correct face triangle among all the candidates. For this purpose, using 50 faces and corresponding face triangles, registration is done. Since the faces are freely oriented, a reference system is created as follows:

- The x-axis is oriented from the right eye to the left one.
- The z-axis is oriented in the same direction as the normal vector of the face triangle.
- The y-axis is computed using the cross product between the other two axes.
- The origin of the axes is translated to the tip of the nose.
- Finally the system is rotated about the x-axis by 45 degree

After standardizing the images, PCA analysis is performed. First, a region around the nose with 80*90 sizes is extracted from the original image as shown in Fig10. Then the extracted image is used for both registration and recognition.

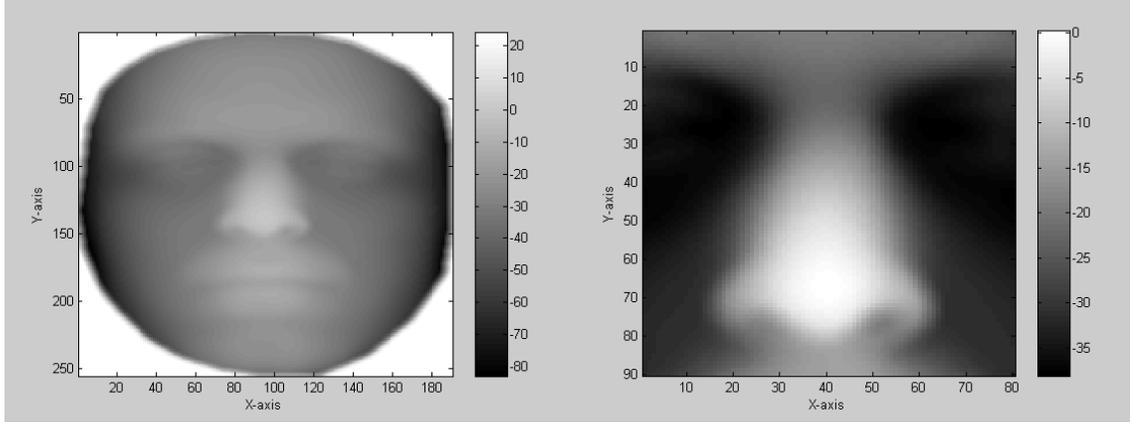


Figure 22 a) Range representation of a face and b) The region extracted around nose

For each one of M images, (M is 50 for our case) training vector is constructed by linking each row of the image together. (There are 50 training vector with size 7200×1) These training vectors, named as $(\Gamma_1, \dots, \Gamma_M)$, are used to construct the “face space”. The mean vector is computed as:

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad (9)$$

Then, the covariance matrix \mathbf{C} is built as:

$$\mathbf{C} = \frac{1}{M} \sum_{n=1}^M (\Gamma_n - \Psi)(\Gamma_n - \Psi)^T \quad (10)$$

The eigenvectors \mathbf{u}_i of \mathbf{C} form an orthogonal base for the face space. Each eigenvalue λ_i represents the variance of the direction indicated by \mathbf{u}_i . To reduce the dimension, k eigenvectors with greater eigenvalues are selected. k shall be selected to satisfy a predefined threshold $r(k)$. See (11) for the definition of $r(k)$.

$$r(k) = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^M \lambda_i} \quad (11)$$

2.4.4. Determining the actual face triangle by using PCA analysis:

The base of the reduced space is denoted as $U = (u_1, \dots, u_k)$. For each one of the candidate triangles, face is standardized. After the standardization, test face vector can be projected in the reduced space as follows:

$$\omega = U^T (\Gamma - \Psi) \quad (12)$$

With some loss of information, the projected vector can be re-projected in the original space as follows

$$\Gamma = \Psi + U\omega \quad (13)$$

After the re-projection, Figure 23 is obtained for a correct face triangle. It can be easily seen from the re-projected image that the current triangle is the correct one.

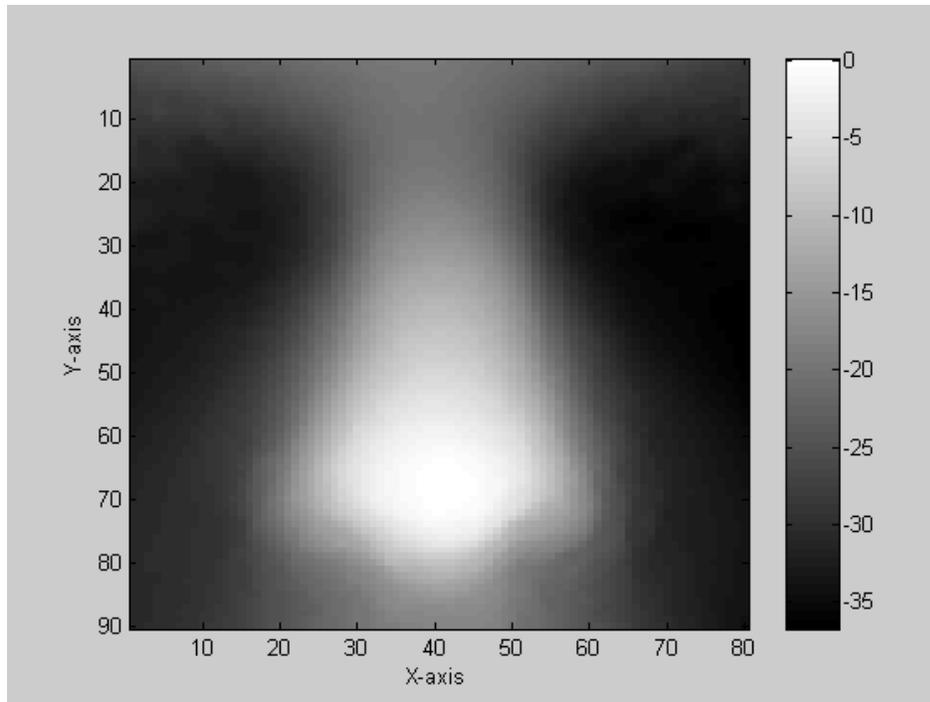


Figure 23 The image obtained after the reproduction by using the eigenvectors of the face

Finally, the amount of information loss is calculated by taking the difference between the original image and the re-projected image. The difference for a correct face triangle is shown in Figure 24. Smaller the difference, more correct the face triangle. The base for this approach is that a non-face cannot be represented by using the eigenvectors of the face space.

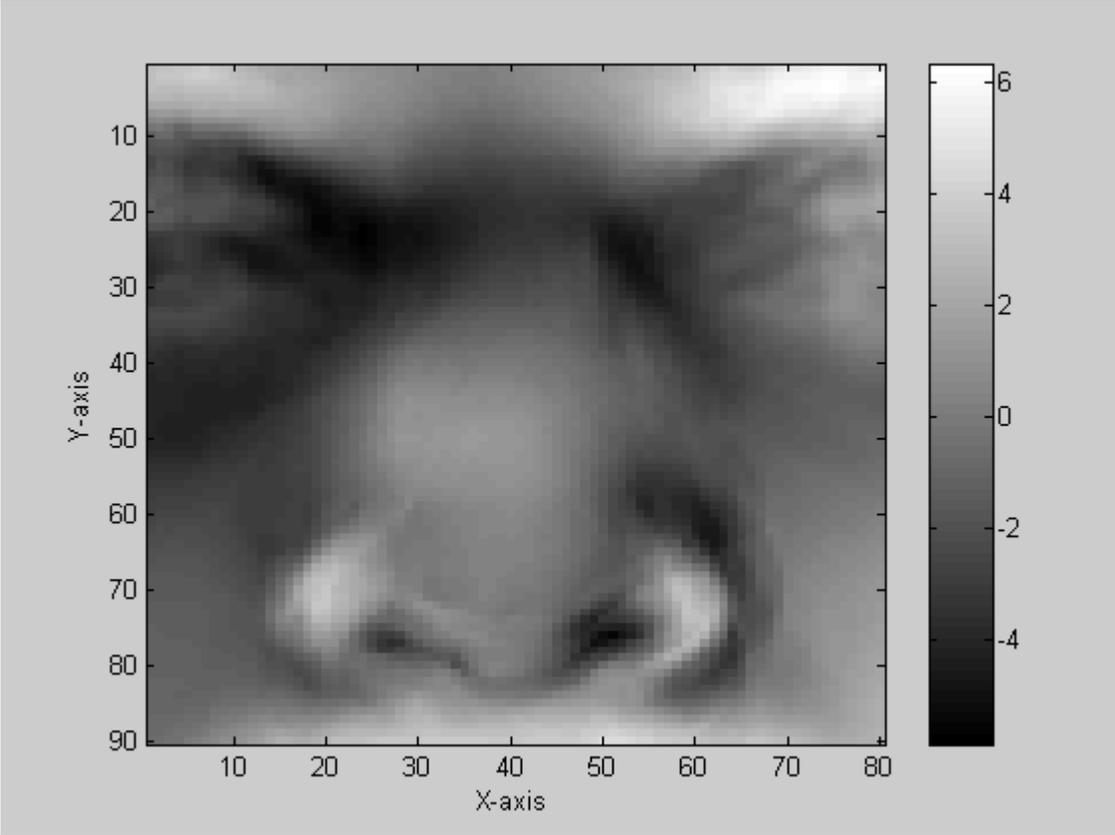


Figure 24 The difference between the original image and the reconstructed image

CHAPTER 3

DATA SET

3.1. Data Set

All face recognition techniques described in this document are the 3D face recognition techniques. Because of this, data used as the input of these systems are the 3D point cloud (i.e. data consists of points that have X, Y, and Z coordinates.) A piece of source data is shown below:

<u>X</u>	<u>Y</u>	<u>Z</u>
-15.569	-89.860	-52.452
-14.782	-89.877	-52.784
-13.992	-89.884	-52.916
-13.202	-89.891	-53.066
-12.407	-89.881	-52.878
-11.619	-89.900	-53.161
-10.825	-89.901	-53.094
-10.038	-89.923	-53.403
-9.249	-89.939	-53.617
-8.462	-89.958	-53.894
-7.670	-89.963	-53.880
-6.874	-89.959	-53.711
-6.087	-89.968	-53.870
-5.296	-89.966	-53.821

First, second, and third columns are X, Y, and Z coordinates of the points respectively.

Data set consists of 17 different poses and expressions of 78 people. Images are grouped according to their poses and expressions. These poses are described in Figure 25.

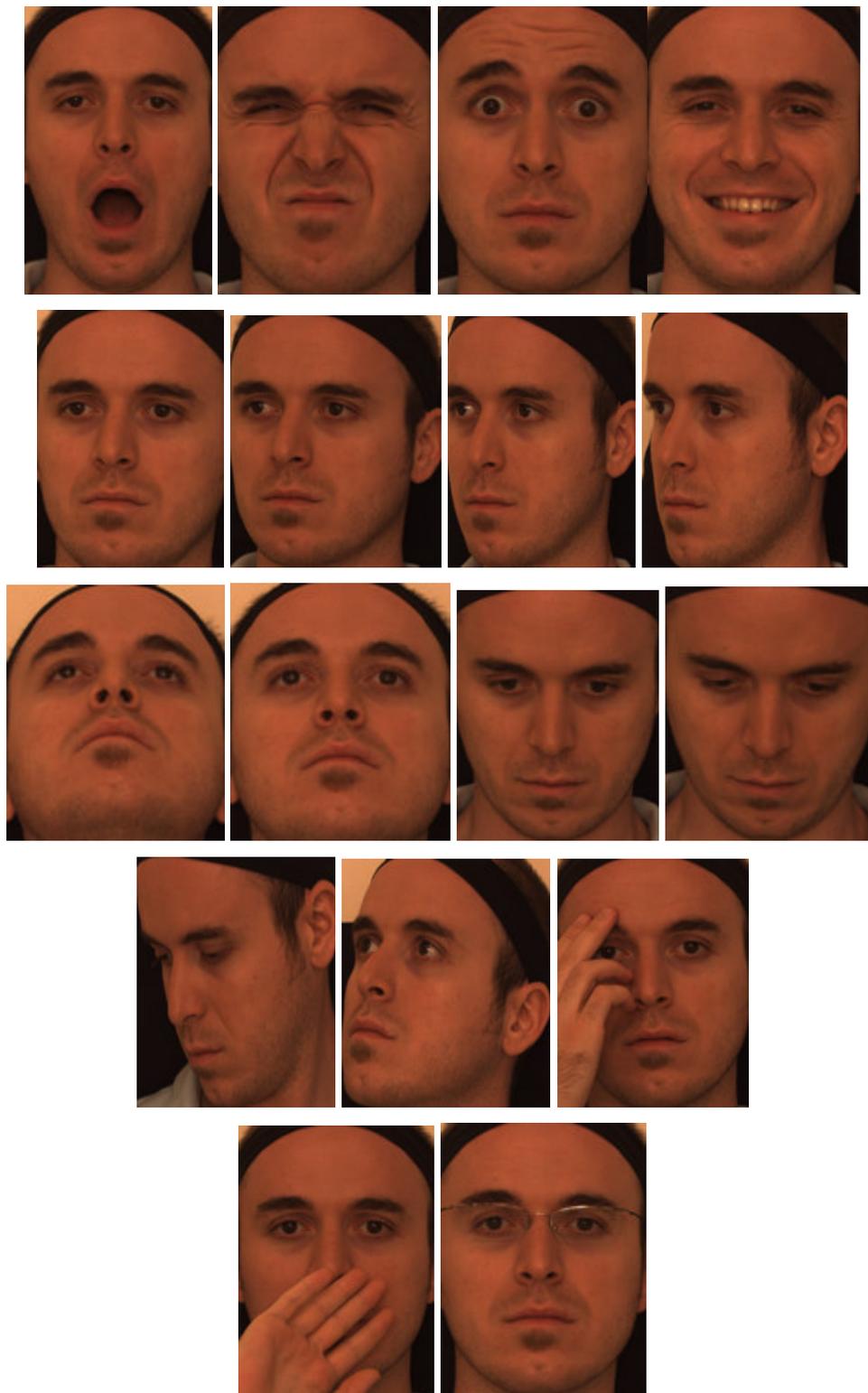


Figure 25 Mouth stretch (a), disgust(b), Inner brow raiser (c), happiness(d), yaw rotations of +30 ° (e), +45 ° (f), +60 ° (g), +75 ° (h), pitch rotations of strong upwards (i), slight upwards(j), slight downwards(k), strong downwards(l), bottom right(m), upper right(n), eye occlusion(o), mouth occlusion(p), eye glasses(r).

CHAPTER 4

EXPERIMENTS & RESULTS

4.1. Results

There are two performance criteria for comparing the methodologies. One of them is the percentage of the success rates and the other one is the mean of the distances between ground truth points and the detected point. Both criteria are examined for each pose and expression separately.

Percentage of the success rate is determined by dividing the number of successful face detection to the number of total images. If the distance between the actual and detected points is smaller than 20 mm, it counts as successful detection. For the second criteria, the mean of the distances between ground truth points and the detected points is calculated for only successfully detected faces.

Detection percentage for the nose and the inner eye pits are compared among all methods. The results are given in Table 1. M1 is Xiaoguang and Jain's method [1]. M2 is Xiaoguang and Jain's other method [2]. M3 is Chang et al.'s method [3]. Finally M4 is Colombos et al.'s method [4]. Best results for each pose are underlined and bolded. Beside the detection rates, Detection errors, examined for the nose and the inner eye pits, are compared for all methods. The mean and the standard deviation of the error (the absolute distance between the marked points and the calculated points) are given in Table 2. The values are in millimeters. The mean is written over the standard deviation. Best mean values for each pose are bolded and underlined.

For each pose, the method with the biggest success rate is signed as successful. For Look Right & Down 45° pose, M4 is signed as successful for Nose detection but its success rate is only 35,7 percent. 35,7 percent is not an acceptable success rate for a face recognition system. However, since the other methods have less success rates for this pose, M4 is signed as successful.

Table 1 Detection percentage for the nose and the inner eye pits

%	Nose				Left Eye				Right Eye			
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
Neutral Pose	90	90	67,1	<u>100</u>	<u>88,6</u>	<u>88,6</u>	81,4	78,6	88,6	<u>90</u>	82,6	80
Mouth Open	97	75,8	80,3	<u>100</u>	<u>97</u>	77,3	87,9	75,8	<u>97</u>	74,2	84,8	75,8
Disgust	89,7	92,6	58,8	<u>100</u>	91,2	<u>92,6</u>	69,1	63,2	91,2	<u>92,6</u>	77,9	64,7
Eyebrows Up	79,7	92,8	68,1	<u>100</u>	79,7	<u>92,8</u>	63,8	72,5	81,2	<u>92,8</u>	72,5	71
Smile	82,6	78,3	65,2	<u>98,6</u>	<u>82,6</u>	78,3	72,5	62,3	<u>85,5</u>	79,7	78,3	68,1
Look Right 30°	70	81,4	4,3	<u>100</u>	<u>81,4</u>	<u>81,4</u>	54,3	64,3	75,7	<u>80</u>	8,6	68,6
Look Right 45°	48,6	72,9	7,1	<u>100</u>	65,7	<u>72,9</u>	41,4	<u>72,9</u>	55,2	67,2	7,5	<u>71,6</u>
Look Right 60°	7,4	54,4	0	<u>60,9</u>	10,1	55,1	24,6	<u>63,8</u>	9,3	39,5	2,3	<u>44,2</u>
Look Right 75°	0	<u>32,2</u>	0	2,9	0	<u>42,6</u>	14,7	39,7	0	0	0	0
Look Up 30°	78,3	43,5	39,1	<u>97,1</u>	<u>76,8</u>	44,9	71	68,1	<u>76,8</u>	44,9	75,4	<u>76,8</u>
Look Up 15°	95,7	82,9	61,4	<u>100</u>	<u>88,6</u>	82,9	71,4	68,6	<u>90</u>	82,9	70	74,3
Look Down 15°	81,4	52,9	75,7	<u>100</u>	87,1	52,9	<u>88,6</u>	74,3	<u>87,1</u>	51,4	81,4	80
Look Down 30°	59,7	17,9	40,3	<u>98,5</u>	77,6	14,9	<u>92,5</u>	76,1	76,1	16,4	<u>88,1</u>	80,6
Look Right & Down 45°	14,3	22,2	4,7	<u>35,7</u>	24,3	28,6	11,4	<u>41,4</u>	20,1	26,7	10,3	<u>53,3</u>
Look Right & Up 45°	16,2	<u>51,5</u>	0	22,9	20	<u>58,6</u>	32,9	42,9	20,1	<u>44,1</u>	4,4	32,1
Hand on one Eye	53,6	23,2	57,9	<u>95,7</u>	59,4	26,1	59,4	<u>65,2</u>	60,7	25	64,3	<u>75</u>
Hand on mouth	59,7	14,9	16,4	<u>100</u>	62,7	14,9	74,6	<u>82,1</u>	62,7	14,9	59,7	<u>79,1</u>
With Glasses	89,7	30,9	75	<u>100</u>	<u>92,5</u>	31,3	67,2	83,6	<u>91</u>	29,9	59,7	<u>79,1</u>

Table 2 Detection errors, examined for the nose and the inner eye pits

μ (mm) σ (mm)	Nose				Left Eye				Right Eye			
	M1	M2	M3	M4	M1	M2	M3	M4	M1	M2	M3	M4
Mouth Open	3,96 2,54	5,35 2,40	4,27 2,94	<u>3,05</u> 1,71	<u>6,08</u> 2,95	6,91 3,99	8,51 4,09	7,57 4,86	6,13 3,23	<u>5,55</u> 3,17	11,30 3,84	5,98 4,55
Disgust	6,09 3,51	6,20 3,01	7,38 4,88	<u>4,34</u> 2,67	5,94 4,49	<u>5,69</u> 4,08	9,41 5,00	7,60 5,15	5,56 3,72	<u>5,12</u> 2,97	11,16 4,29	6,71 4,24
Eyebrows Up	3,80 2,80	5,29 2,92	4,51 2,87	<u>3,21</u> 2,16	<u>6,36</u> 3,10	6,38 3,17	9,35 4,54	7,01 4,42	6,17 3,02	<u>5,91</u> 2,96	11,36 4,07	7,40 5,16
Smile	5,65 3,85	5,09 2,47	6,58 4,25	3,90 2,12	6,21 3,92	6,37 3,86	8,74 4,41	6,60 3,99	5,84 3,00	<u>5,52</u> 2,98	10,36 3,93	6,88 4,84
Look Right 30°	7,39 3,97	<u>4,61</u> 2,99	9,58 1,05	4,65 2,88	<u>7,51</u> 3,95	7,66 3,89	14,51 3,48	7,91 4,22	5,84 3,94	<u>5,25</u> 2,93	13,58 4,70	7,69 5,52
Look Right 45°	10,93 4,87	<u>5,66</u> 3,71	15,10 3,85	7,78 3,86	7,85 4,04	<u>6,86</u> 3,27	12,61 4,38	8,69 4,62	7,68 4,13	7,91 4,92	13,34 4,22	<u>7,64</u> 5,34
Look Right 60°	10,44 3,81	<u>8,45</u> 3,85	-	9,81 3,98	<u>6,94</u> 5,15	6,71 3,29	12,93 5,84	9,85 5,95	11,19 5,07	<u>7,78</u> 3,62	19,20 -	<u>7,04</u> 4,95
Look Right 75°	-	<u>13,73</u> 4,46	-	13,76 7,69	-	<u>8,23</u> 5,19	11,42 6,63	9,28 5,31	-	-	-	-
Look Up 30°	3,48 2,71	6,06 2,93	4,53 2,77	<u>3,44</u> 2,41	8,75 4,97	<u>5,57</u> 2,72	9,29 3,99	6,55 4,07	8,49 4,14	<u>5,14</u> 3,27	11,24 4,16	7,12 4,69
Look Up 15°	3,48 3,00	4,74 2,24	4,84 3,21	<u>3,07</u> 2,12	6,30 2,92	<u>5,28</u> 2,87	9,33 4,18	7,80 4,60	6,27 3,52	<u>5,35</u> 3,22	11,21 3,43	6,71 4,44
Look Down 15°	6,17 3,90	5,72 3,00	5,46 3,24	<u>4,10</u> 2,42	<u>6,19</u> 3,80	7,03 4,17	7,29 4,02	6,29 4,42	6,03 4,10	5,95 3,33	9,20 3,96	<u>5,34</u> 3,84
Look Down 30°	7,46 5,32	5,83 2,59	<u>6,04</u> 3,75	4,61 2,48	4,67 3,71	<u>4,64</u> 3,92	6,87 3,40	4,74 3,24	5,42 3,47	5,47 3,70	7,51 3,50	<u>4,47</u> 3,23
Look Right & Down 45°	11,46 4,77	8,87 4,00	14,14 4,80	<u>8,78</u> 4,11	8,24 4,93	7,24 4,27	10,10 4,84	<u>6,62</u> 4,36	10,09 4,22	7,38 4,05	<u>7,35</u> 4,51	7,51 5,15
Look Right & Up 45°	13,53 4,23	<u>9,22</u> 4,41	-	11,35 4,28	9,99 4,36	<u>7,37</u> 4,44	12,47 4,79	9,32 5,26	11,7 6,57	8,84 6,69	11,37 -	<u>5,47</u> 3,18
Hand on one Eye	5,29 3,85	6,01 2,20	5,40 3,75	<u>3,82</u> 2,22	7,39 4,02	7,42 5,35	9,90 4,96	<u>6,62</u> 4,33	5,98 3,42	<u>4,48</u> 2,02	10,06 5,11	8,06 4,72
Hand on mouth	4,76 3,35	5,52 2,68	7,69 5,59	<u>3,98</u> 2,36	<u>6,05</u> 3,46	7,70 3,73	10,48 4,84	6,68 4,05	<u>5,24</u> 3,30	5,76 4,26	12,65 4,50	6,17 4,11
With Glasses	4,33 2,75	5,14 2,58	4,49 2,66	<u>3,03</u> 1,70	7,16 4,01	<u>6,41</u> 3,04	10,63 4,82	7,49 3,90	<u>7,38</u> 4,31	7,82 3,67	11,94 4,38	7,61 4,13

4.2. Success Rate Analysis

There are 54 (3x18) results tabulated for each method at Table 1, which are for 18 poses and 3 landmark points for each pose. The best detection rates are underlined and bolded. It is observed that M4 has the best rating for 28 out of 54 results, which are

mostly for the detection of the nose tip. Method M2 has best results in 14 cases out of 54. Similarly method M1 has the best for 13 cases. Method M3 has the best result for only 3 cases. Comparison and the analysis of the success rates are investigated in details in the following section. Figure 26 gives a clear illustration of the success rates. Values in Figure 26 are the number of successful detections. To compare the total success rates of these methods in a clear graphic, Figure 27 is drawn.

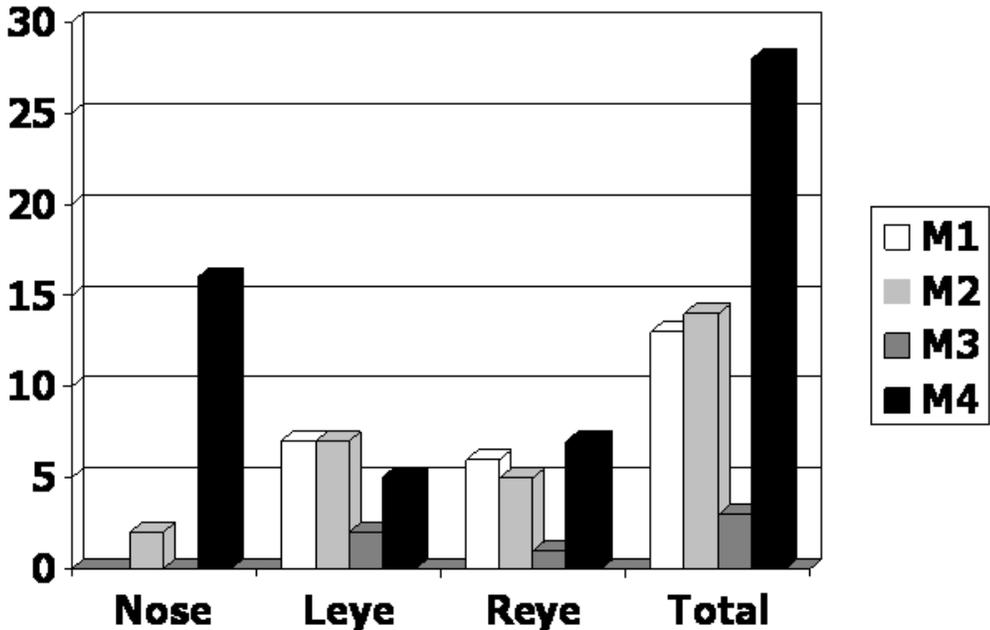


Figure 26 Success of the four methods for nose and eyes detection

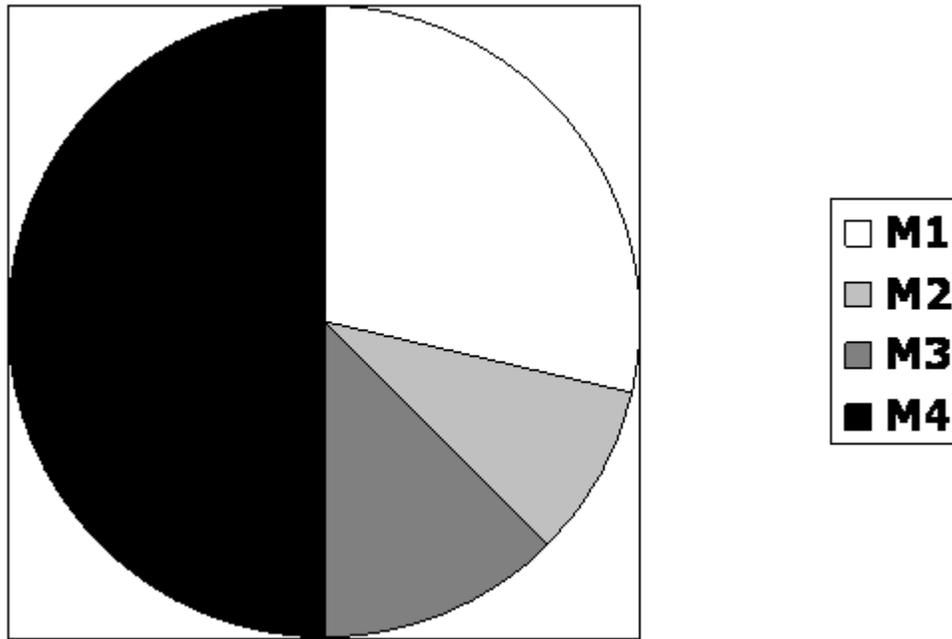


Figure 27 Success rate percentages of four methods

Neutral pose:

Nose:

Methods M1 and M2 have acceptable success rates for neutral poses. The reason for this is that these methods have some assumptions and these assumptions are valid only for frontal neutral poses. One of these assumptions is that nose tip is the closest point to the scanner (i.e. nose tip has the maximum Z value among the other points in face). The other assumption is valid for M1 and this method assumes that nose tip is on the column that has the maximum average Z value.

M3 has worse performance in frontal images compared to other methods. This method first determines the inner eye points, then find the nose tip by using the inner eye points. An error in pit region determination misleads the place of the nose tip.

M4 has perfect success rates for nose tip detection. This method is most complicated one and it uses PCA analysis for detection. Before PCA, image is rotated and Nose tip is located at the origin.

Eye:

M3 has considerable success rates for inner eye points. This is expected because it first determines the place of eyes then the nose tip is detected by using the inner eye points. M1 and M2 have similar inner eye point success with nose tip detection. Reason for this is that the eye points are detected after the detection of nose tip.

An interesting result is obtained for M4. Although the eye points and nose tip are detected together, there is a noticeable decrease in success rates of inner eye points when compared with nose tip. The reason for that is, for M4, image is rotated basically with the following properties;

- X-axis is oriented from right eye to left eye.
- Z-axis is oriented in the same direction of the normal vector of the face triangle.

A nose tip and pits around mouth also satisfies the conditions for a face triangle. Because of this, success rate for inner eye points is smaller than success rates of nose tip for M4. This is the weak point of M4.

Mouth open:

For Nose Tip detection, compared to the neutral pose, it is observed that success rates are increased for M1 and M3 and decreased for M2. M4 still has 100% success for this pose.

This pose increases the Z values on the column of nose tip. It is illustrated in Figure 28. Then it eliminates the fake columns that decrease the success rates. Besides that, for M3, this pose arranges the shape curvature to determination of Mean and Gaussian curvature correctly.

For inner eye points, M1 is the most successful method. For M1 and M2, success rates are similar to Nose tip success rates because for these methods, eye points are found after finding nose tips. In comparison to nose tip successes, M3 has better performance and M4 has worse performance in inner eye point detection. The reasons are described in neutral pose comparison section.



Figure 28 Z values on the selected line are increased

Disgust:

This pose decreases success rates catastrophically for M3. The reason for this is that M3 first determines the inner eye points by using eye's standard surface characteristic and this pose destructs the standard characteristic of the surface curvature around eyes. Other methods are also affected from this destruction slightly.



Figure 29 Destruction around eyes

Eyebrows Up:

This pose decreases the success rate of M1. The reason for this is that at the column of nose tip, this pose decreases the Z values relatively on forehead (i.e. this pose increases the Z values of the regions illustrated in Figure 30). There is no remarkable change on other methods' successes.

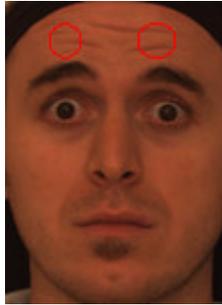


Figure 30 Z values are increased around the eyebrows

Smile:

This pose decreases the success rates for M1 and M2. The reason for decreasing is the puffiness in cheeks. (See Figure 31) Both of these methods assume that the nose tip has the maximum Z value but this puffiness misleads these methods. Success rates for M4 are also decreased slightly because of this puffiness. These regions are considered as possible nose tips. On the other hand, there is not any noteworthy difference in the success rates of M3 because the pose is not related with inner eye points and this method based on the pit regions.



Figure 31 Puffiness in cheeks

Yaw rotations of +30 °, +45 °, +60 °, and +75 °

Success rates of the method on rotated images take more importance because for security issues, it is more possible to have a rotated image than an image with an expression. For that reason, performance of the methods for rotated images is designative for selecting the best of four methods.

M1 has difficulties on rotated images w.r.t. y axis because it determines the closest column to screen as the location for the nose. For rotated images w.r.t. y axis with higher than 30 degrees, closest column changes to another line on cheeks. It attracts attention that M3 has nearly zero success rates for rotated images. Since this method tries to find inner eye points together and one of the inner eye points is not on the screen, this method even fails to find inner eye points. Success rates of M4 are changed slightly up to 45 degree rotations (comparing with neutral pose). Since M4 deals with the surface curvatures and it does not use any assumption, it succeeds to find anchor points if they are seen on the screen. On rotated images more than 45 degrees, one of the inner eye points disappears from the screen and the method fails to obtain correct nose-eye-eye triangle.

After analyzing the results, it is observed that M2 is the best among other methods because it is robust for the pose changes w.r.t. y axis. In comparison to M4, since M2 does not need to find all three points at the same time, although one of the eye points is not seen on the screen, this method has a chance to find the nose tip and the other inner eye point.

Upwards and Downwards (15 ° and 30 °):

Upward and Downward images are the images rotated w.r.t. X-axis. Success rates for these poses are also designative for determining the success rates. M4 is the most powerful method for these images. Since the rotation angles are not bigger than 45 degrees, none of the anchor points are disappeared. Because of this, theoretically there should be no difference between the neutral pose and the upward and downward rotated images. Actually, success rates for these images are nearly equal with the success rates of neutral pose images.

Contrary from the poses rotated w.r.t. y axis, M1 has acceptable success rates. Rotations w.r.t. x axis does not change the assumptions of M1. Column of the nose tip still has the maximum average z values compare to other columns.

M2 has troubles with upward and downward images. Chin and forehead are counted as possible nose tip candidates and the correct nose tip is eliminated before the last phase of M2 (i.e., actual nose tip can not succeed to be a nose tip candidate).

Success rate of M3 does not change on a large scale because the surface characteristic of the face does not change for these poses.

Bottom Right 45 ° , Upper Right 45 ° :

These poses are fully problematic for all methods. Since these poses are combinations of the rotations w.r.t. Y-axis and w.r.t. X-axis, all destructive conditions related with rotations specified previously are also valid for those poses.

Hand on Eye:

Eye occlusion is a big problem for M1 and M2. In fact, the problem is not caused from the disappearance of eye, but the unrelated hand located in pose. It destroys the assumptions of both M1 (Nose tip is not on the line closest the camera) and M2 (Nose tip is not the point closest to camera). M3 is not affected on a large scale because it only deals with the surface characteristic and it does not have any assumptions related with the places of the anchor points. Similar to most of the other poses, M4 is the powerful method for this pose. Its success rate is nearly same with the neutral pose.



Figure 32 Hand on Eye

Hand on Mouth

Mouth occlusion is problem for all methods except M4. Since this method does not have any assumption and all of the anchor points appear on the image, this destruction does not affect the success rate of this method. Although this pose seems as a destructive pose for M1 and M2, it is not destructive for M1. The reason for this is, the hand is around the column that nose tip is located and hand does not mislead to find the midline. It is not valid for M2. For all rotated versions of the images (i.e. rotations from

90 degrees to -90 degrees around y -axis), a point on the hand is found as the closest point to the screen. Therefore, the performances of M2 for all anchor points are smaller than 15 percent. Eye point detections for M3 shall not be affected from mouth occlusion because it first determines the inner eye points. It is observed that as expected M3 has good success rates for inner eye points. However, after obtaining the inner eye points; since the nose tip is selected as the point with highest z value, M3 determines a point on hand as nose tip.



Figure 33 Hand on Mouth

Eye Glasses:

Although M1 and M3 are not affected from eyeglasses negatively, success rate of M2 decreases dramatically. Since only the frames of the eyeglasses are obtained from scanner, M3 does not have any problem related with eyeglasses. M1 overcomes this problem as well, because eyeglasses do not change the observed midline. (Midline is the column with maximum average Z value). However, since M2 rotates the image and observe the maximum z value for each rotated image, different parts of the eyeglasses frame is selected as the candidate nose tips for this pose. Because of this, success rate of M2 is nearly zero. M4 is again the best and it achieves a success similar to neutral pose.

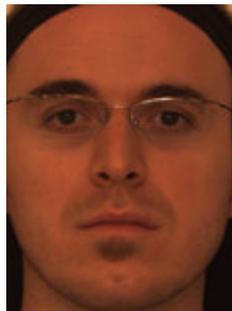


Figure 34 Eyeglasses

4.3. Complexity Analysis

In this section, complexities of four algorithms are investigated in terms of speed and scaling. Only test phases of the algorithms are considered because any training phase is completed before the implementation of the method.

M1

M1 algorithm can be divided into 2 parts. Nose Tip detection and Inner Eye points detection.

Nose Tip Detection:

```
for i=1:M,
    for j=1:N,
        mid-line determination
    end
end
for i=1:M,
    Z profile obtained
end
for i=1:M,
    Consecutive increases
end
for i=1:k1,
    for i=1:k2,
        Nose tip is found
    end
end
```

Thus the time complexity for Nose tip detection step can be stated as;

$$O(.) = (MN + M + M + k_1k_2).$$

Since k_1 and k_2 can be taken as constants, the complexity in terms of scaling is $O(MN)$

Inner Eye Point Detection:

```

for i=1:M,
    for j=1:N,
        Shape index and Cornerness detection
    end
end
for i=1:M,
    for j=1:N,
        GMM for Detecting Inner Eye Points
    end
end

```

Time complexity for Inner Eye Point detection is $O(\cdot) = (MN + MN)$. The complexity in terms of scaling is $O(MN)$

Then, Total Time Complexity of the method M1 is $O(\cdot) = (MN + M + M + k_1k_2 + MN + MN)$ and the complexity in terms of scaling is $O(MN)$

k1: Most possible nose candidates

k2: Elements used in Nose horizontal profile

M2

M2 can also be divided in two parts as Nose Tip detection and Inner Eye point's detection.

Nose Tip Detection:

```

for k=1:k1,
    for i=1:M,
        for j=1:N,
            Rotate the each point w.r.t. y axis
        end
    end
end
end

```

for k=1:k1,

 Choose the Nose Tip

end

Time complexity for Nose tip detection step can be stated as $O(.) = (k_1MN + k_1)$.

k_1 can be taken as constant, the complexity in terms of scaling is $O(MN)$

Inner Eye Point Detection:

for i=1:M,

 for j=1:N,

 Shape index and Cornerness detection

 end

end

for i=1:M,

 for j=1:N,

 GMM for Detecting Inner Eye Points

 end

end

Time complexity for Inner Eye Point detection is $O(.) = (MN + MN)$. The complexity in terms of scaling is $O(MN)$

Then, Total Time Complexity of the method M2 is $O(.) = (k_1MN + k_1 + MN + MN)$ and the complexity in terms of scaling is $O(MN)$

M3

M3 has a simpler algorithm in comparison to M1 and M2. It first determines the Mean and Gaussian Curvatures of the images. By using Mean and Gaussian Curvatures, Inner Eye Points and Nose Tip are detected. The whole algorithm is as follows;

for i=1:M,

 for j=1:N,

 Mean and Gaussian Curvatures determination

 end

end

```

for i=1:M,
    for j=1:N,
        Inner Eye Points Detection
    end
end
for i=1:M,
    Nose Tip Detection
end

```

Then, Total Time Complexity of the method M3 is $O(.) = (MN + MN + M)$ and the complexity in terms of scaling is $O(MN)$

M4

M4 firstly filter the image for smoothing. It then calculates mean and Gaussian curvature values and opening operation is performed to segment connected components clearly.

```

for i=1:M,
    for j=1:N,
        for l=1:k1,
            Filtering for smoothing
        end
    end
end
for i=1:M,
    for j=1:N,
        Mean and Gaussian Curvatures determination
    end
end

for i=1:M,
    for j=1:N,
        for l=1:k2,

```


point candidates and they effect the computation time directly. So, the complexity in terms of scaling is $O(MN p_1 p_2^2)$,

k_1 : Filter size for smoothing

k_2 : Filter size for opening operation

k_3 : Filter size to find connected components

k_4 : Size of the Region for applying PCA

p_1 : Number of Nose Tip Candidates

p_2 : Number of Inner Eye Point Candidates

In summary, complexities of all the methods are;

M1: $O(.)=O(MN + MN) = O(2MN) = O(MN)$

M2: $O(.) = O(k_1MN + k_1 + MN + MN) = O(MN(2 + k_1) + k_1) = O(MN)$

M3: $O(.) = O(MN + MN + M) = O(2MN + M) = O(MN)$

M4: $O(.) = O(MN (k_1 + k_2 + 1+k_3) + p_1p_2p_2MNk_4) = O(MN p_1 p_2^2)$

CHAPTER 5

CONCLUSIONS

5.1. Conclusions

M1 is a robust method for nose tip detection even if the image is rotated around x-axis only. The reason for this is that the midline does not change. However, rotation around y-axis cannot be handled by this method. On the other hand, facial expressions do not decrease the success rate for M1.

M2 has an opposite characteristic with M1 for the nose tip detection on rotated images. Since this method finds the nose tip by rotating the image around y-axis, it handles a rotation around y-axis. However, if the image is rotated around x-axis, this method assumes the forehead and chin as the nose tip. This method has a powerful strategy for inner eye point detection. Since the rotation angle of the image is found besides the nose tip, this method uses this information for extracting the inner eye points more successfully than M1.

M3 has a simple algorithm for detecting inner eye points. It only uses shape index and cornerness characteristic of the image and it does not have an assumption for the positions of the anchor points. Because of these, although it is the fastest method among others, its performance is the worst.

M4 has the best performance in nearly half of the results. Although it has the best success rates, it needs longest time for the extraction of anchor points. The reason for this is described in Chapter 4 in details. Computational complexities of M1, M2 and M3 are $O(MN)$. However, the complexity of the M4 is $O(MN p_1 p_2^2)$. Remember that p_1 and p_2 are the number of possible Nose Tip and Inner Eye Point candidates. If these numbers increase, computational load of M4 increases rapidly. On the other hand, the correctness performances of M1, M2 and M3 are smaller than M4.

Relation between anchor points is a valuable information and it shall be used in face detection techniques. All of the methods except M4 use this information for detecting an anchor point after finding another one. However, since M4 detects the

anchor points all together, it uses this information through out the algorithm. This information increases the success rate and also the computational cost in return.

Although M4 needs longest time, it has the best performance among the others. With today's technology, with advanced processors, execution time for M4 is still acceptable. In conclusion, if there is a selection chance, M4 should be selected for orientation and expression invariant 3D face detection purpose.

5.2. Proposed Future Work

In this work, four methods are implemented and success rates are compared. Different methodologies and approaches are examined for the extraction of anchor points. Beside that, some improvements are offered to some methods. Based upon the experience gained through this research, a hybrid face detection system may be constructed using the advantageous properties of each proposed methods.

REFERENCES

- [1] X. Lu and A. K. Jain, "Multimodal Facial Feature Extraction for Automatic 3D Face Recognition," Technical Report MSU-CSE-05-22, Computer Science and Engineering, Michigan State University, August 2005.
- [2] X. Lu and A. K. Jain, "Automatic Feature Extraction for Multiview 3D Face Recognition," *Proc. 7th IEEE International Conference on Automatic Face and Gesture Recognition (FG2006)*, pp. 585-590, Southsampton, UK, Apr. 2006.
- [3] K. I. Chang, K. W. Bowyer, "Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression", *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, VOL. 28, NO. 10, OCTOBER 2006
- [4] A.Colombo , C. Cusano , R. Schettini, 3D face detection using curvature analysis, *Pattern Recognition*, v.39 n.3, p.444-455, March, 2006
- [5] C. Dorai and A. Jain. Cosmos - a representation scheme for 3d free-form objects. *IEEE Trans. PAMI*, 19(10):1115– 1130, 1997.
- [6] M. O. Irfanoğlu, B. Gökberk, L. Akarun, "3D Shape-Based Face Recognition Using Automatically Registered Facial Surfaces," in *Proc. ICPR*, vol.4, pp.183-186, 2004.
- [7] Y. Lee, K. Park, J. Shim, T. Yi, 3D face recognition using statistical multiple features for the local depth information, *Proceedings of the 16th International Conference on Vision Interface*, 2003

- [8] 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.
- [9] G. C. Feng, P. C. Yuen, and D. Q. Dai, "Human face recognition using PCA on wavelet subband," J. Electron. Imaging, vol. 9, pp. 226–233, 2001.
- [10] Zhong Jin , Zhen Lou , Jingyu Yang , Quansen Sun, Face detection using template matching and skin-color information, Neurocomputing, v.70 n.4-6, p.794-800, January, 2007
- [11] A. Mian, M. Bennamoun, and R. Owens, "2D and 3D multimodal hybrid face recognition," in Proceeding of the 9th European Conference on Computer Vision (ECCV '06), p. 344, Graz, Austria, May 2006.
- [12] C. Boehnen and T. Russ. A fast multi-modal approach to facial feature detection. In Workshop on Applications of Computer Vision, 2004.
- [13] R. Niese, A. Al-Hamadi and B. Michaelis, A Novel Method for 3D Face Detection and Normalization, Institute for Electronics, Signal Processing and Communications (IESK) Otto-von-Guericke-University Magdeburg D-39016 Magdeburg, P.O. Box 4210 Germany
- [14] D. Colbry, G. Stockman, A. Jain, (2005), "Detection of anchor points for 3D face verification", paper presented at IEEE Workshop on Advanced 3D Imaging for Safety and Security, A3DISS, San Diego, CA.
- [15] T. Papatheodorou, D. Rueckert, Evaluation of 3D face recognition using registration and pca, LNCS 3546: International Conference on Audio- and Video-based Biometric Person Authentication (AVBPA 2005) (2005) 997-1018.

- [16] J. Yang, D. Zhang, A.F. Frangi, J.Y. Yang, Two-dimensional PCA: a new approach to appearance based face representation and recognition, *IEEE Trans. Patt. Anal. Mach. Intell.* 26 (1) (2004)
- [17] C. Heshner, 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.
- [18] K. Kim, "Face Recognition using Principle Component Analysis" Department of Computer Science, University of Maryland, College Park
- [19] W.S. Yambor, B.A. Draper, J.R. Beveridge, Analyzing pca-based face recognition algorithms: eigenvector selection and distance measures, in: *Empirical Evaluation in Computer Vision*, 2000.
- [20] K. I. Chang, K. W. Bowyer and P. J. Flynn, "Face Recognition Using 2D and 3D Facial Data," *MMUA*, pp. 25–32, 2003.
- [21] F. Tsalakanidou, S. Malassiotis, and M. Strintzis, "Integration of 2D and 3D Images for Enhanced Face Authentication," *Proc. Sixth Int'l Conf. Automated Face and Gesture Recognition*, pp. 266-271, May 2004.
- [22] 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.
- [23] M. Turk and A. Pentland, Eigenfaces for recognition, *J. Cognitive Neuroscience*, Vol. 3, 71-86., 1991.
- [24] Lu, X., D. Colbry, A.K. Jain, "Three-Dimensional Model Based Face

Recognition,” in Proc. ICPR, 2004.

- [25] Lu, X. A.K. Jain, “Integrating Range and Texture Information for 3D Face Recognition,” to appear in Proc. IEEE WACV, 2005.
- [26] Papatheodorou, T., D. Rueckert, “Evaluation of automatic 4D face recognition using surface and texture registration,” in Proc. AFGR, pp.321-326, 2004.
- [27] G. Gordon. Face recognition based on depth maps and surface curvature. Geometric Methods in Computer Vision, SPIE 1570:1–12, July 1991.
- [28] J. C. Lee and E. Milios. Matching range images of human faces. Int’l Conf. on Comp. Vision, pages 722–726, 1990.
- [29] Kittler, J. Hilton, A. Hamouz, M. Illingworth, J.: 3D Assisted Face Recognition: A Survey of 3D Imaging, Modelling and Recognition Approaches. Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on Volume 3, Issue , 20-26 June 2005 Page(s): 114 - 114
- [30] Kyong I. Chang, Kevin W. Bowyer, Patrick J. Flynn: Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression. IEEE Trans. Pattern Anal. Mach. Intell. 28(10): 1695-1700 (2006)