August 2013
Thesis for the Degree of Masters of Science

RGB-D Fusion Based on Facial Normal Vectors Using 3D Point Cloud

Graduate School, Chosun University
Department of Computer Engineering
Muhammad Naveed Iqbal Qureshi

A u g u s t 2 0 1 3

Thesis for the Degree of Masters of Science

RGB-D Fusion Based on Facial Normal Vectors Using 3D Point Cloud

Graduate School, Chosun University

Department of Computer Engineering

Muhammad Naveed Iqbal Qureshi

3D 포인트 클라우드의 얼굴 정규 벡터들에 기반한 RGB-D 융합

RGB-D Fusion Based on Facial Normal Vectors Using 3D Point Cloud

2013년 8 월 23 일

조선대학교 대학원

컴 퓨 터 공 학 과

쿠레쉬 무하마드 나비드 이끼발

3D 포인트 클라우드의 얼굴 정규 벡터들에 기반한 RGB-D 융합

지도교수 이 상 웅

이 논문을 석사학위신청 논문으로 제출함

2013년 4 월

조선대학교 대학원

컴 퓨 터 공 학 과

쿠레쉬 무하마드 나비드 이끼발

쿠레쉬 무하마드 나비드 이끼발

석사학위논문을 인준함

위원장 조 선 대 학 교 교수 문 인 규 인

위 원 조 선 대 학 교 교수 양 희 덕 인

위 원 조 선 대 학 교 교수 이 상 웅 인

2013년 5 월

조선대학교 대학원

Acknowledgement

I would like to express my deep and sincere gratitude to my advisor, Professor, Dr. Sang Woong Lee for hosting my research in Korea. His wide knowledge and his logical and critical way of thinking have been of great value to me. His understanding, encouragement and personal guidance have provided a perfect basis for the present thesis.

I owe my loving thanks to my mother and my brothers. They have lost a lot due to my research abroad. My special and deepest gratitude is due to my wife Sairah Naveed for her loving support, who become a source for me to exile in life. Without her encouragement and understanding it would have been impossible for me to finish this work.

Contents

List of Figures

List of Tables

초 록

3D 포인트 클라우드의 얼굴 정규 벡터들에 기반한 RGB-D 융합

무하마드 나비드 이끼발 쿠레쉬
지도교수 : 이 상 웅 교수
조선대학교 컴퓨터공학과

다양한 얼굴 인식 알고리즘이 개발되고 독립형 시스템으로 사용할 수 있도록 개발되었다. 그러나 전 세계에서 강인한 얼굴 인식 시스템은 모든 보안 및 정보 기관에서 여전히 꿈으로 여겨지고 있다. 2D 얼굴 인식 시스템은 조명, 다양한 얼굴의 자세, 노화와 얼굴 노출의 차단 등의 문제로 성능에 문제가 있다. 위의 문제뿐만 아니라, 2D 시스템을 사용한 실제 얼굴 존재의 입증은 불가능하다. 이러한 문제점들로 인해 얼굴인식 연구자들은 2.5D와 3D 얼굴 인식 시스템의 개발에 대한 연구를 진행하여 왔다. 이러한 시스템은 광원에 의한 실제 얼굴에서의 명암, 얼굴 윤곽, 깊이 정보 등을 사용하기 때문에 실제 얼굴 대신에 사진으로 얼굴 인식을 시도하는 것은 불가능하다. 현재까지 이미 많은 2.5D와 3D 얼굴 인식 시스템이 시장에 존재하지만, 실시간 동작의 어려움과 센서의 매우 높은 가격은 여전히 주된 문제점이다. 3D 얼굴 모델의 렌더링 과정은 처리과정이 매우 복잡하고 요구되는 하드웨어가 많다. 이러한 문제점으로 인해 2.5D와 3D 얼굴 인식

시스템을 실생활에서 사용하는 것은 쉽지않다.

본 논문에서는 키넥트 센서의 RGB-D 데이터의 알파 블렌딩의 융합을 위한 새로운 방법을 제안한다. 제안된 방법은 2D나 3D 얼굴에서 성능이 잘 나오지 않는 상황에서 유용하며, 특별히 손상된 데이터의 경우에 좋은 성능을 보여 준다. 제안된 방법은 2D 와 3D 얼굴의 모든 정보를 사용고 있으며, 데이터 수집을 위해 마이크로소프트사의 키넥트 센서를 사용한다. 키넥트 센서는 다른 센서들에 비해 매우 저렴한 가격으로 이용 가능하기 때문에 응용 제품을 개발하는 측면에서 매우 유용하다.

본 논문에서는 2D-3D 데이터 융합과 Fisherfaces 알고리즘을 기반으로 하여 알파 블렌딩과 연계하여 특정한 상황에 강인한 얼굴 인식 방법을 제안한다. 이를 위해 알파 블렌딩 융합 방법과 EURECOM의 Kinect의 얼굴 데이터베이스를 사용하였으며 깊이 데이터의 3 차원 클라우드의 법선 벡터는 융합을 위한 알파 마스크 변수를 만드는 데 사용된다.

이러한 알파 블렌딩 융합 방법과 EURECOM 데이터베이스를 바탕으로 회색조, 제안 방법, 제안 방법 후 가우스함수 적용 등의 3가지 실험데이터를 합성하였다. 또한 조명 환경에 대한 변화를 측정하기 위해서 4가지 조명 조건에서 인식 결과를 평가하였다. 또한 얼굴 정규화 과정에서의 두가지 관심 영역(ROI)을 설정하여 인식 결과를 비교하였다.

이러한 다양한 설험 데이터를 통하여 실험한 경과 제안된 방법은 정상적인 경우에는 기존 방법에 비해 유사한 결과를 보여주었으며, 특히 눈썹이 가려지거나 조명에 의해 일부 영역이 가져지는 등의 가려진 얼굴 데이터의 경우에는 상대적으로 좋은 성능을 보였다.

ABSTRACT

RGB-D Fusion Based on Facial Normal Vectors Using 3D Point Cloud

Muhammad Naveed Iqbal Qureshi
Advisor: Prof. Sang Woong Lee, Ph. D.
Department of Computer Engineering
Graduate School of Chosun University

Various face recognition algorithms have been developed and become available as standalone systems. However, a robust face recognition system is still a dream of all the security and intelligence agencies around the globe. 2D face recognition systems fail to perform well in certain cases due to the effect of illumination, various face poses, aging and different type of occlusions. In addition to above mentioned problems, verification of actual face presence is also impossible by using 2D systems, therefore, researchers move towards the development of 2.5 and 3D face recognition systems. It is harder to forge these systems by using photograph instead of the actual face because they utilize multimodal features such as optical flow, curvature profile and depth information etc. which cannot be acquired by conventional 2D sensors. Many 2.5 and 3D face recognition systems are already available in the market, but lack of real time operation and very high sensor cost are the main issue with most of them. Further, rendering of a 3D face model is computationally very expensive and require immense hardware resources. Due to these issues its acceptance for being deployed in common use become undesirable.

We have developed a new method for the alpha blending based fusion of RGB-D data of Kinect sensor. Our fusion approach gives us promising results in situations where the performance of only 2D and 3D is not well; specially in the presence of corrupted data. Our system relies both on the 2D and 3D information of the face. We have used the state-of-the-art Kinect sensor from Microsoft for the data acquisition. Since this sensor is available in very low price as compared to the other 2.5 and 3D sensors, its utilization is highly desirable for this application.

A robust face recognition system is implemented in this research that works with alpha blending based 2D-3D data fusion and Fisherfaces algorithm. We have utilized the EURECOM Kinect face dataset with a new alpha blending fusion method. Normal vectors of 3D point clouds of depth data are used to create alpha mask for fusion.

With the help of this mask we have created three different dataset, namely; grayscale, fusion and fusion with Gaussian smoothing dataset. We have evaluated the recognition results with four different illumination conditions. Two regions-of-interest (ROI) for size normalization are also used to evaluate the recognition results. This system can work either offline with our custom made variants of fusion dataset or online by using RGB-D data acquisition from Kinect sensor.

Our method mainly insist on the robustness of the fusion data. In case of corrupted RGB information, fusion dataset set gives us up-to 64% recognition rate. The other advantage of using this fusion method is its requirement of low computational power. It makes this method desirable for use in realtime environment on ordinary machines without any graphical processing unit (GPU).

1. Introduction

Face as a biometric has the distinct advantage over other modalities such as fingerprint, palm print, DNA, footprint, hand geometry, voice, pattern of vain and iris recognition, in that the acquisition stage is non-intrusive and can be achieved with readily available equipment. 3D representations of the human face have the potential to overcome many of the obstacles such as pose and illumination sensitivity, which have prevented the widespread adoption of face recognition technology [1]. There are various alternative biometric techniques that perform very well, as mentioned above, but these methods require the participants to cooperate and follow a relatively strict data acquisition protocol. In a face recognition scenario participants are not necessarily required to cooperate or even be aware of being scanned and identified, which makes face recognition a less intrusive and potentially more effective identification technique. Finally, the public's perception of the face as a biometric modality is more positive as compared to the other modalities [1]. RGB-D fusion method proposed in this thesis makes the recognition process more efficient in unfavourable conditions such as, in the presence of noise and/or the corrupted data.

The correct identification of individuals based on their faces would make many aspects of our life a lot simpler. First of all, in terms of the need to identify oneself, it would minimize the number of information one must carry or remember for one's day-to-day survival in the modern world. For example, it might be possible in the future to move in and out of a country without requiring a passport and voting might not require registration. Furthermore, in terms of protecting information, logging on to a network or a computer would be automated while access to a dataset or one's personal files would not require a password. When it comes to law enforcement and security applications, trained policemen would not have to go through thousands of hours of CCTV footage in order to identify suspects; as a real-time identification system processing footage from various locations would be able to spot wanted persons and notify police of their whereabouts. Finally, the use of face recognition technology could be beneficial in content tagging. Some online services already offer to tag member's pictures based on the identity of the subject portrayed using face recognition technology [2].

In addition to a need for more sophisticated 3D recognition algorithms, there is also a need for more sophisticated multi-modal combination. Those studies that suggest that 3D allows greater accuracy than 2D also suggest that multi-modal recognition allows greater accuracy than either modality alone. A 2D camera is typically already present as a part of a 3D sensor, so it seems that 2D can generally be acquired along with 3D. Thus, the more productive research issue may not be 3D versus 2D, but instead the best method to use to combine 3D and 2D. Multi-modal combination has so far generally taken a fairly simple approach. The 3D recognition results and the 2D recognition results each can be produced without reference to the other modality, and then the results are combined in some way. It is at least potentially more powerful to exploit possible synergies between the two modalities in the interpretation of each modality [3].

Early work in 3D facial recognition emerged in the late 1980's but it was not until recently that substantial research dataset have become available. The face recognition grand challenge [4] aims to address this issue and provides both a common dataset and experiment methodologies to enable accurate comparisons of different algorithms. However, there is no fusion dataset publically available till date for any comparative analysis with proposed method of this research. In [5] the authors present a good summary of the current research in 3D and composite 2D-3D recognition, in particular they note that while it is accepted that a combination of 2D and 3D gives greater performance, it is still unclear which modality performs better in isolation. In the following work more focus was applied on the fusion domain, however the methodology is equally applicable to combination with a more sophisticated 2D recognition engine. In general, approaches to 3D recognition fall into 3 main categories [5]: those that use 3D correspondence matching explicitly to provide discrimination [6, 7]; those that extract 3D features such as curvature directly from the face; and those that treat the range image as a 2D image in order to extract features [8]. The later has the advantage that a considerable number of well tested image processing algorithms can be directly applied, and we have also followed this approach in our method.

A. State of the Art

Qin Jin *et al.* have developed an algorithm to extract facial features with depth active appearance model (AAM) [9]. Their main idea is to accurately locate the human head and estimate the head pose using Kinect. Pixels in a depth image indicate the calibrated depth in the scene, rather than the measure of intensity or color. They combined the RGB and depth data in the same image, depth information was compressed as the alpha channel. Chuanjun Wang *et al.* have proposed Fourier series based expression deformation model for 3D face recognition using depth maps and prinicipal component analysis (PCA) [10]. They utilize Fourier representation of depth map to acquire Fourier series residues. This information was further processed with PCA. Dominik. Jelsovka *et al.* proposed that using 3D range images, 3D mesh diagrams can be generated [11]. Ajmal Mian has examined that hole, spikes and noise cause problem while dealing with 3D data for feature detection and face recognition. He had trained the Haar cascade classifier for detecting the face features [12]. Peter W. Hallinan *et al.* have investigated in creating 3D representations of data, it is often necessary to define data over evenly spaced 2D grids [13]. We, hereby envision developing a robust face recognition system to ensure the security access authorization with minimum risk of wrong authentication. There are many other approaches for score level fusion rather than fusion of input data. Gede Putra Kusuma *et al* [14] proposed the multimodal face recognition method in which decision is based on the fusion at score level. They have run two separate classifiers and combined the result together later on.

Some of the common 2D+3D multimodal fusion approaches are presented as follows in Table 1.

Sr. No.	Method
1	Feature level fusion
2	Score-level fusion
3	Data and score level fusion
4	Data-level fusion (our method)
5	Pixel and score level fusion
6	Score and feature level fusion

Table 1. Multimodal 2D+3D fusion approaches used in previous researches [15]

Stereo based systems can have trouble getting a true sampling of face surface. Systems that depend on structured light typically have trouble in regions such as eyebrow and often generate spike artifacts when light undergoes multiple reflections [3]. We have utilized this information in our work as the main goal to achieve higher recognition rates by using the fusion data in presence of noise or image corruption on the eyebrow area. In the following of this thesis, we will elaborate this concept in detail.

Existing 3D sensors are certainly capable of supporting advanced research in this area, but are far from ideal for practical applications. An ideal 3D sensor for face recognition applications would combine at-least the following properties: (1) Image acquisition time is similar to that of typical 2D camera, (2) a large depth of field; e.g, a meter or more in which there is essentially no loss in accuracy of depth resolution, (3) robust operation under a range of ‘‘normal’’ lighting conditions, (4) no eye safety issues arising from projected light, (5) dense sampling of depth values; perhaps 1000 x 1000, and (6) depth resolution of better than 1 mm. Evaluated by these criteria, we do not know of any currently available 3D sensor that could be considered as ideal for use in face recognition

[3].

Similarly, three-dimensional face recognition needs better algorithms. Here, ''better'' means more tolerant of real world variety in the pose, facial expression, eye-glasses, jewelry and other factors. At the same time, ''better'' also means less computationally demanding. Three-dimensional face recognition in general seems to require much more computational effort ''per match'' than does 2D face recognition [3].

We have utilized Kinect sensor in our method because the depth information acquisition from it is as fast as RGB, and it is extremely low cost as compared to other 3D sensors in the market.

B. Motivation

The main motivation of using fusion method is its robustness against noise. RGB-D fusion of the data is a long awaited requirement of many computer vision researcher. Some of the recent fusion approaches like the work of Richard A. Newcombe [16] are computationally very expensive and are not favourable for use in face recognition researches. We have proposed a low computational cost RGB-D data fusion algorithm. The rates of recognition using the proposed fusion dataset which is generated by our proposed fusion algorithm is similar to that of grayscale dataset in normal conditions. However, in the presence of noise or data corruption such as the removal of eyebrows; our method outperform other methods. The importance of eyebrows in the face recognition was discussed in detail in the literature by Pawan Sinha [17]. Figure. 1 show the details of this phenomenon. Sadr *et al.* [18] have shown that just one feature such as eyes or eyebrows can be enough for recognition of many famous faces.

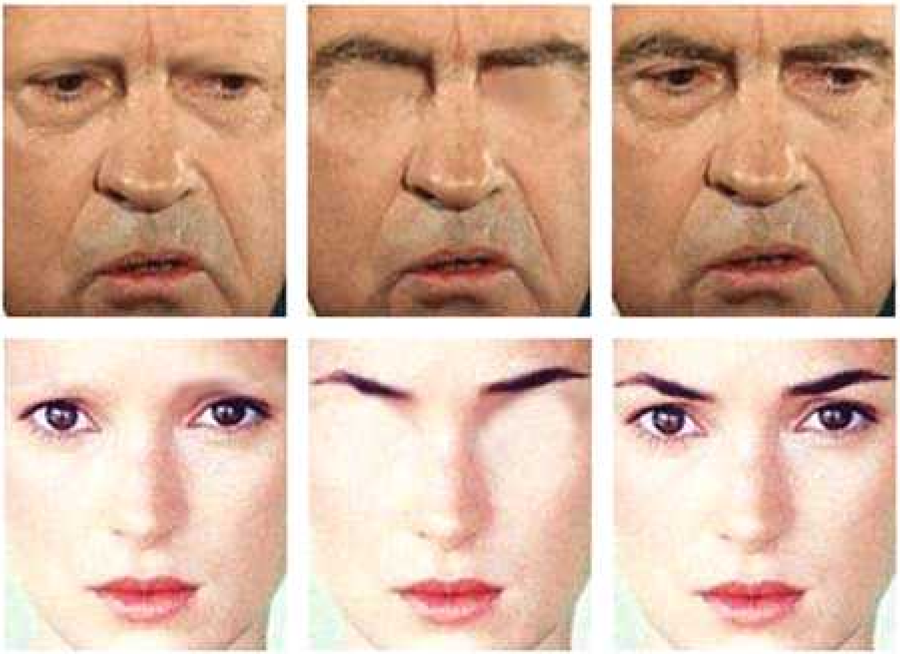

Figure 1. Sample stimuli from Sadr et al.'s [18] experiment assessing the contribution of eyebrows to face recognition: original images of President Richard M. Nixon and actor Winona Ryder, along with modified versions lacking either eyebrows or eyes [17]

Face recognition depends heavily on the particular choice of features used by the classifier. One usually starts with a given set of features and then attempts to derive an optimal subset (under some criteria) of features leading to high classification performance with the expectation that similar performance can also be displayed on future trials using novel (unseen) test data [19].

Progress made in face recognition would also provide important techniques for other disciplines and applications. Some statistical face models can already identify certain genetic syndromes which subtly affect the face [20]. In addition a model that can simulate the aging of a person [21] can be useful in the search for people who have been missing for many years. Finally, research in this area would also lead to advances in facial expression recognition which could have profound effects on human-computer-interaction. It is for these reasons that face recognition is such an inter-disciplinary research area with scientists from various fields contributing to the literature, from psychophysics and psychology, to mathematics and computer science. Another reason for the growing interest in face recognition has been the emergence of affordable hardware, such as digital photography and video, which have made the goals of such research more feasible. Affordable systems that capture the 3D geometry of a surface, like the one used in our study and presented later in the thesis, make 3D face recognition a “tolerable” research interest, which combined with cheap computing power makes the intense calculations involved in surface processing realistic. Finally, the field of face recognition has inherited a wealth of algorithmic tools that had been traditionally developed for other disciplines, such as medical imaging, something which has allowed research to move forward relatively fast [22].

Face recognition can benefit the areas of: law enforcement, travel and transportation, military, homeland defense, and national security, government (national, state, municipal), benefit distribution, physical security, facility management, risk management, IT security, logical access, transactions, authentication, forensics, Id management (IdM), HR, personnel management, wage fraud, time/attendance, financial, healthcare, construction, consumer/retail sectors, gaming and hospitality, high tech and telecom, airport security, access control, driver's licenses & passports, customs & immigration and scene analysis. The following paragraphs give details of some of these topics respectively.

Law Enforcement: Today's law enforcement agencies are looking for innovative technologies to help them stay one step ahead of the world's ever-advancing terrorists.

Airport Security: Face Recognition can enhance security efforts already underway at most airports and other major transportation hubs (seaports, train stations, etc.). This includes the identification of known suspects before they get onto an airplane or into a secure location.

Access Control: Face recognition can enhance security efforts considerably. Biometric identification ensures that a person is who they claim to be, eliminating any worry of someone using illicitly obtained keys or access cards.

Driver's Licenses & Passports: Face recognition can leverage the existing identification infrastructure.

This includes, using existing photo dataset and the existing enrollment technology (e.g. cameras and capture stations); and integrate with terrorist watch lists, including regional, national, and international "most-wanted" dataset.

Homeland Defense: Face recognition can help in the war on terrorism, enhancing security efforts. This includes scanning passengers at ports of entry, integrating with CCTV cameras for "out-of-the-ordinary" surveillance of buildings and facilities, and more.

Customs & Immigration: New laws require advanced submission of manifests from planes and ships arriving from abroad; this should enable the system to assist in identification of individuals who should, and should not be there [23].

2. Face Detection on Depth Maps

Face detection is one of the most thoroughly-explored questions in computer vision, with applications ranging from consumer cameras to human-computer interaction [24]. The faster and more reliable we can make face detection, the better our interaction with the people owning these faces will be. The computer vision community addresses face detection from a monocular image or video-centric perspective. Most algorithms are designed to detect faces using one or more camera images, without additional sensor information or context. A Kinect sensor, however, has multiple sensors in addition to its cameras, including sensors that can provide depth information. Laser range finders, stereo camera pairs, the new Kinect sensor [25], and the SwissRanger camera [26] can all provide depth information of varying types and qualities. All of these additional cues can be used to make face detection more efficient and, potentially, more accurate. In this thesis we have utilized an extension to the classic Viola-Jones face-detection algorithm [27] that considers depth and RGB information for searching the faces in an image. This work is similar to that of Dixon *et al.* [28].

A. The Computer Vision Approach for Face Detection in Depth Images

The Viola-Jones algorithm [28] involves exhaustively searching an entire image for faces, with multiple scales explored at each pixel. Given no other knowledge about the scene, an exhaustive search is reasonable, and the Viola-Jones algorithm is efficient. With available video data from Kinect sensor, the basic algorithm can be made more efficient by tracking the detected faces, and searching only a local neighborhood around faces found in previous frames. Areas of the image corresponding to points further than a given distance are likely to be blurry, or to contain too few face pixels for reliable detection. Since we cannot detect faces in these areas, even if they are present, there is no point in trying, and we can avoid running the face detector over these image areas [29].

Figure. 2 show a typical depth map obtained from Kinect Sensor with background suppression.

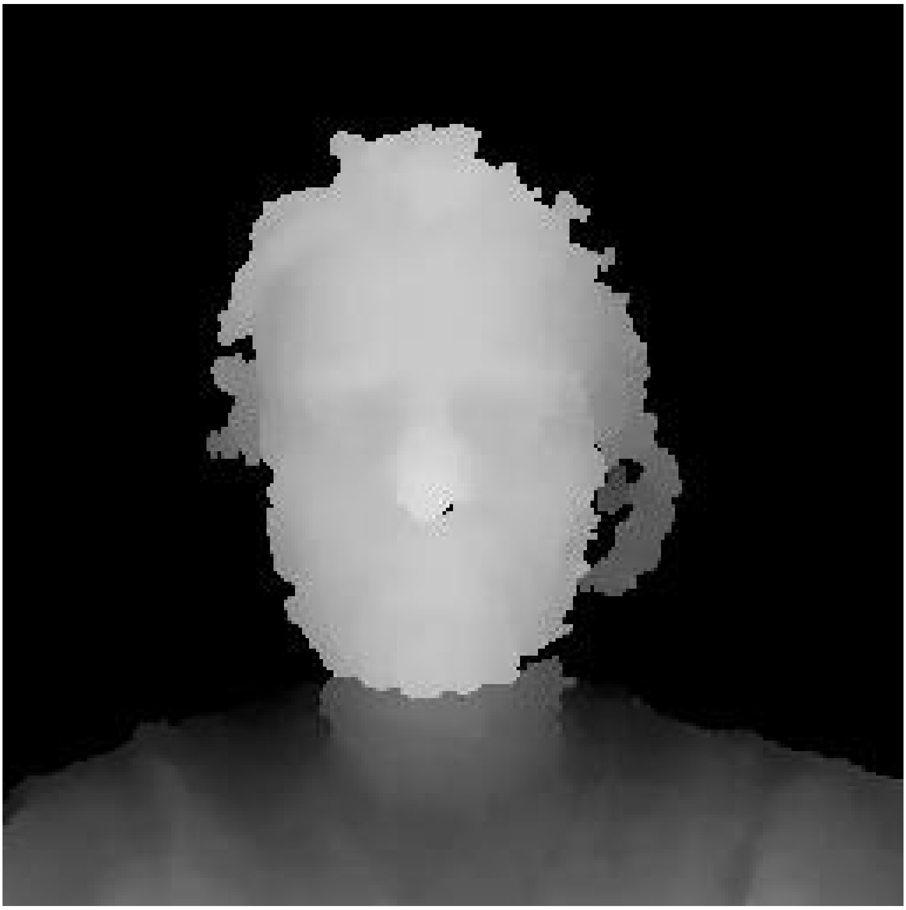

Figure. 2 A typical depth map from Kinect sensor [30]

B. Face Detection on Grayscale and Depth Dataset

Face detector for grayscale data was implemented with the help of MATLAB's cascade object detector and frontal face CART algorithms [31]. Before registration and recognition task, face images were normalized with six different sets of parameter values for illumination and ROI selection. We did the normalization of size with the position of eyes information from Kinect EURECOM face dataset [30]. Background and lighting conditions in grayscale dataset were also normalized before registration for recognition task.

A 2.5D depth image is 2D representation of 3D data points set xyz where each pixel in xy plane store a depth value z. It can be considered as a gray scale image in which black pixels represents the farthermost distance and white pixels shows the nearest distance of the object surface from the sensor [32]. Nose is considered as the center of face in depth dataset because of the highest pixel value and least distance from the Kinect sensor. Although there exists many algorithms for face detection on texture data, but for detecting the face in depth images we had trained a new cascade classifier. We had used the depth images from EURECOM Kinect Face Dataset for training the classifier. The detection accuracy was up-to 99 %. Face images cropped by the detector from the regions where depth values get changed significantly. Depth face size was normalized by using the resizing technique to keep the standard resolution of 100x100 pixels. It was not required to normalize the depth maps for lighting conditions because they did not get affected by illumination changes.

C. Real Face Verification with Depth Data

In the 2D face recognition; it is easy to forge the system by using photographs of the users instead of the original user to forge the system. We can use depth data from the Kinect sensor to verify the presence of real user in-front of it. Kinect sensor projects extremely low power infrared laser light patterns on the object. Depth sensor receives the reflected patterns and estimates the depth by measuring their intensity. Another benefit of using the depth data is that these patterns are in-sensitive to the external lighting conditions, therefore the depth data from Kinect sensor gives promising recognition results. Figure. 3 gives the pictorial explanation of this phenomenon. These results are acquired by implementation of our concept using MATLAB®.

D. Visual Results for Real Face Verification

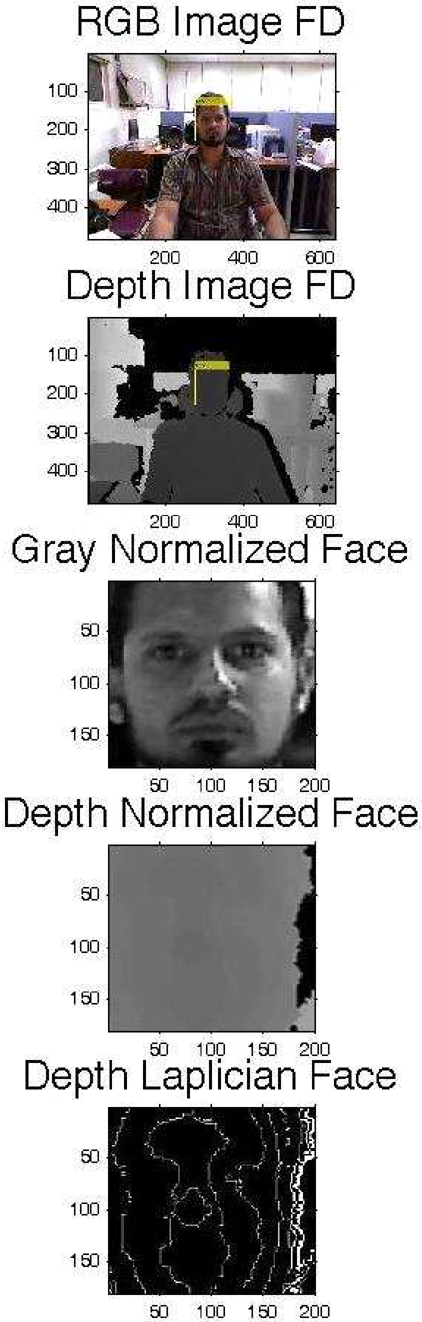

Figure. 3(a) Presence of real face can be verified by both texture and depth data

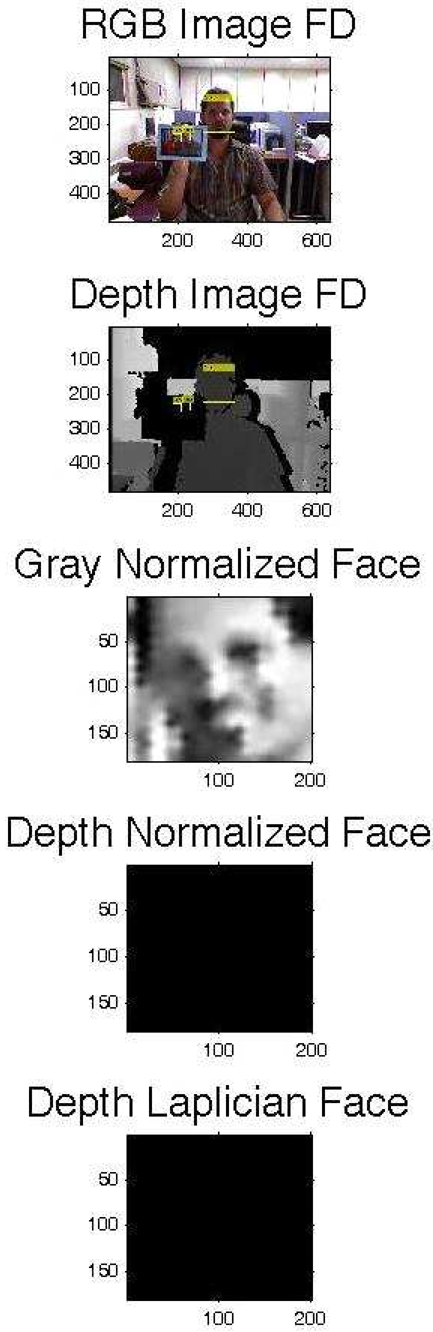

Figure. 3(b) Face detected on a photograph has no depth information. In this case the system will not proceed for further recognition process.

3. Face Recognition Using RGB-D Fusion

We have utilized the Fisherfaces algorithm for recognition task. Our method mainly insist on the robustness of the fusion with Gaussian smoothing data. In case of corrupted RGB information, fusion with Gaussian smoothing dataset gives us promising face recognition results. Our experiment consists of four parts, namely; generation of fusion mask for alpha blending, dataset normalization and registration based on the position of eyes, online recognition using Kinect sensor, and offline recognition using normalized EURECOM Kinect face dataset. This dataset is further modified in our work for noise input and shaved eyebrows. For all three mentioned dataset we have three further classes named as grayscale, fusion, and fusion with gaussian smoothing.

A. Fusion of Texture and Depth Data

Many texture and depth data fusion techniques are available, such as fusing RGB and depth data by row wise concatenation in spatial domain or with PCA or FLD [14]. The Image of human face contains rich information in dimensions with maximum variances, therefore we have proposed a new alpha blending algorithm for this purpose. Fisherfaces becomes highly desirable algorithm for the recognition task, because it take benefit from both principal component analysis to reduce the data dimension and Fisher’s linear discriminator to manage the between and with-in class scatter. Our fusion technique will be explained in detail in the following sub-chapter.

i. Alpha Blending

We have used the 3D point cloud data of Kinect sensor to create the normal vectors. 3D point cloud data gives us details of the depth information of corresponding face image. The alpha blending approach for fusion of the grayscale (texture) and depth information is explained by following Equation 1.

$$fusion = \alpha \times I_G + (1-\alpha) \times I_D \quad (1)$$

Where, $\alpha = \dfrac{Normal\,Vx}{\sqrt{Normal\,Vx^2 + Normal\,Vy^2 + Normal\,Vz^2}}$

I_D is the depth image and I_G is the grayscale texture image. All the images had 256 x 256 dimensions. $Normal\,V_x, Normal\,V_y, \text{and}\,Normal\,V_z$ are the normal vectors obtained from the 3D point cloud of depth data by using the Delaunay triangulation method. All of these normal vectors were reshaped to 256 x 256 dimensions and treated as an image in the following of this work. Following sub-chapter will explain the process of normal vector generation in detail.

ii. Delaunay Triangulation for Normal Vector Generation

We have used this method to create the triangulation of 3D point cloud data which will become input of normal vector generation function of the depth map surface.

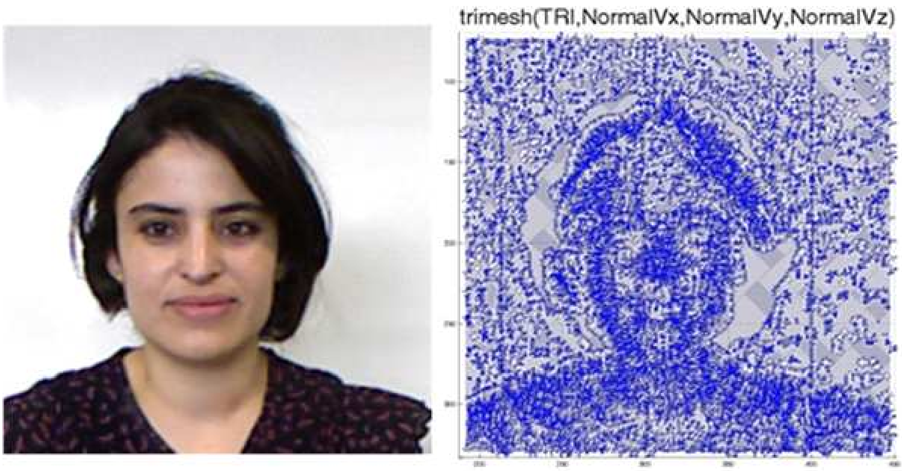

Figure. 4 RGB image and it corresponding normal vectors of 3D point cloud data generated by using Delaunay triangulation

In mathematics and computational geometry, a Delaunay triangulation for a set P of points in a plane is a triangulation DT(P) such that no point in P is inside the circumcircle of any triangle in DT(P). Delaunay triangulations maximize the minimum angle of all the angles of the triangles in the triangulation. The triangulation is named after Boris Delaunay for his work on this topic from 1934 [33, 34]. Figure. 4 gives an overview of the generated normal vectors.

iii. Gaussian Smoothing of Fusion Dataset

We have performed the recognition experiment on fusion image dataset both in its original form and after applying Gaussian smoothing filter to it as shown below in Figure. 5.

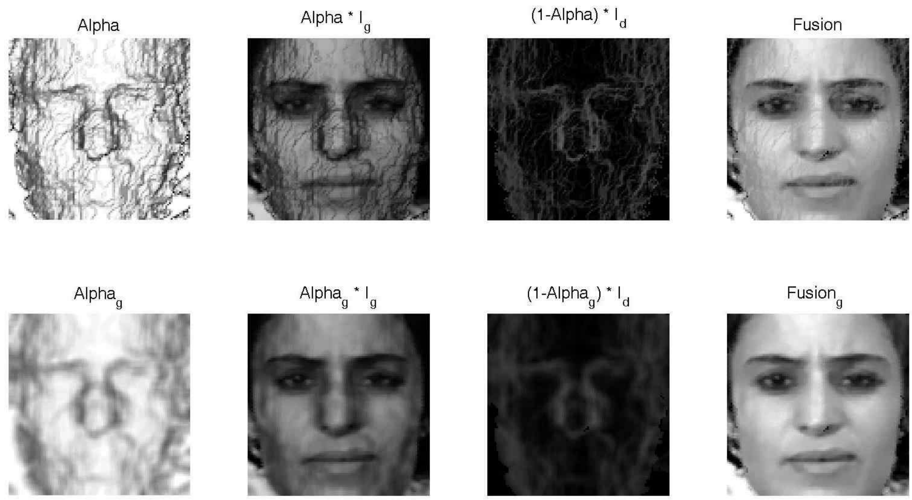

Figure. 5 Fusion dataset before and after Gaussian smoothing. First image from left is the fusion mask α obtained by 3D point cloud. Second image is the α-blending of grayscale image. Third image is the α-blending of depth image. Fourth image is the final result of α-blending fusion as given in equation 1.

We had tested various input parameter combinations for Gaussian smoothing filter and finally selected ROI as 5 x 5 pixels and value of smoothing parameter δ (sigma) as 5.

Gaussian smoothing dramatically increased the recognition rate of fusion dataset up to a significant figure With the noisy or corrupted grayscale input. We have verified this proposition in this thesis by inserting black dots random noise on the cheeks and roughly shaved eye brows in the grayscale input. Details along with the figures of these dataset will be discussed chapter 4.

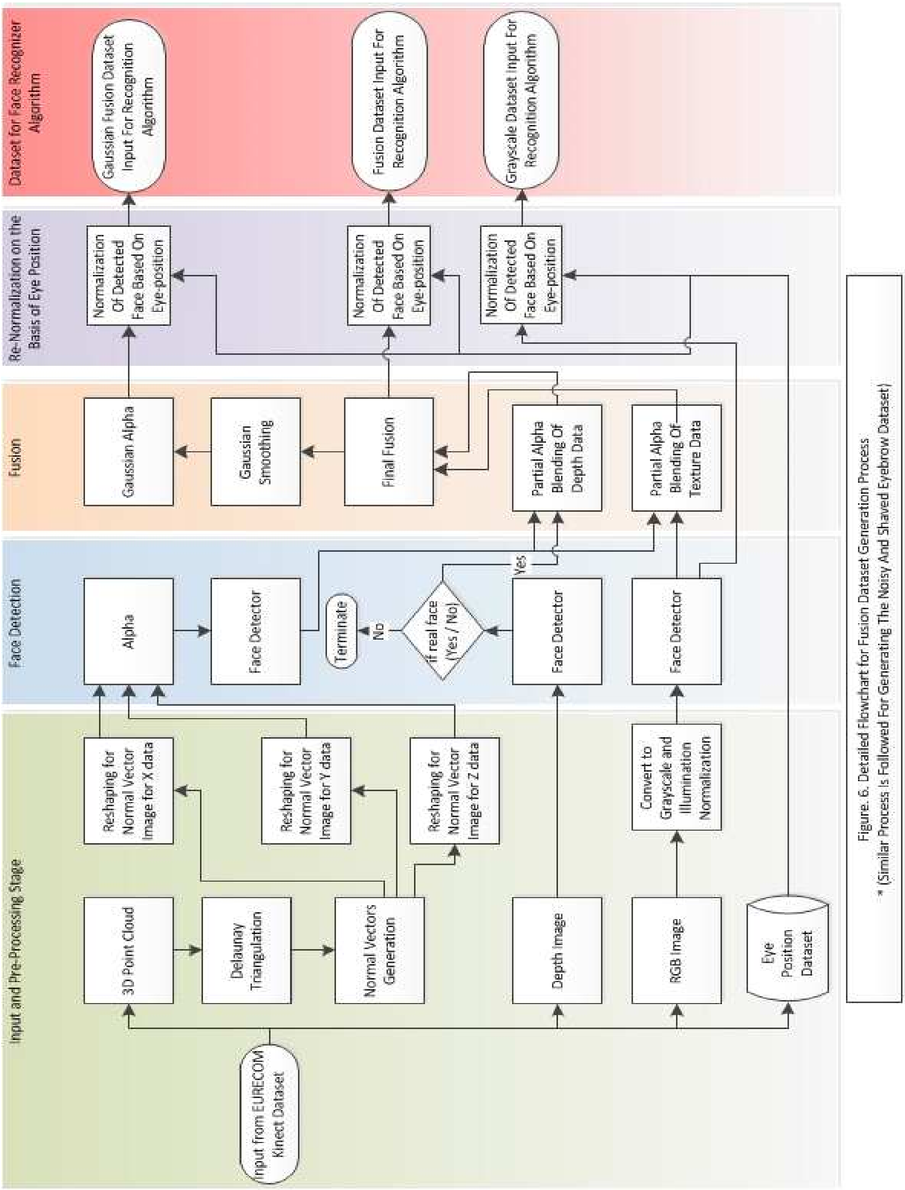

Figure. 6. Detailed Flowchart for Fusion Dataset Generation Process
* (Similar Process Is Followed For Generating The Noisy And Shaved Eyebrow Dataset)

Figure. 6 Whole process of dataset generation for this experiment. We have created three different dataset by using the EURECOM face dataset. Details will be discussed in chapter 4.

B. Preliminary Knowledge

We have choose Fisherfaces as our recognition algorithm for this research. It is actually the combination of two commonly known pattern recognition algorithms; namely, Principal Component Analysis (PCA) and Fisher's Linear Discriminator (FLD) or Linear Discriminator Analysis (LDA). In the following of this section we will describe Fisherfaces recognition algorithm in detail.

i. Fisherfaces Algorithm

Fisherfaces is a statistical face recognition method proposed by Belhumeur et al. [15]. It employs Fisher's Linear Discriminant (FLD) to derive the optimal projection matrix that maximizes the ratio of the determinants of the between-class and within-class scatter matrixes of the projected samples. Fisherfaces views each class as a random variable and sample faces of the class provides observations of the random variable.

Fisherfaces algorithm is ideal for face recognition problem because in face space it is necessary to maximize the distance between different faces as much as possible and minimize the distance between similar faces.

This algorithm relies on the combination of two most popular pattern recognition algorithms i.e. Fisher's linear discriminator (FLD) and principal component analysis (PCA). PCA is performed prior to FLD. Since the within-class scatter matrix is most of the time singular, due to a small sample size problem, PCA aligns the major axis of data on the primary axes of coordinate space and moves it so that data center lies on origin. It projects the sample set to a lower dimensional subspace so that the resulting within-class scatter matrix is non singular. The PCA projection matrix is composed by the eigenvectors of the total scatter matrix. FLD minimize inter or within class scatter S_w of the data (faces) and maximize the between class scatter S_b as well as the ratio of between class scatter and with-in class scatter is maximized.

a. Details of Principal Component Analysis (PCA)

PCA is a dimensionality reduction linear projection that maximize the scatter of all projected samples. Consider a set of N sample images $\{X_1, X_2, X_3, ..., X_N\}$ taking values in an n-dimensional image space, and assume that each image belongs to one of c classes $\{\varsigma_1, \varsigma_2, \varsigma_3, ..., \varsigma_c\}$. Let us have a linear transformation mapping of the original n-dimensional space into a m-dimensional feature space, where $m < n$. The new feature vectors z_k is defined as;

$$z_k = (W)^t X_k \qquad k = 1, 2, 3, ..., N$$

Where, $W \in R^{n \times m}$ is the Eigen vectors matrix with orthogonal columns. z_k is the feature vector, X_k represents the sample image. If S_t is the total scatter covariance matrix defined as:

$$S_t = \sum_{k=1}^{N} (X_k - \mu)(X_k - \mu)^t$$

n is the number of sample images and $\mu = \frac{1}{N} \sum_{k=1}^{N} X_k$ is the mean image of all the samples. After applying the linear transformation W^t, the scatter of the transformed feature vectors $\{z_1, z_2, ..., z_N\}$ is $W^t S_t W$. The Projection W_{opt} is chosen to maximize the determinant of the total scatter matrix of the projected samples as

$$W_{opt} = W_{pca} = \arg\max_W |W^t S_t W|$$

$$= [W_1, W_2, W_3, ..., W_m]$$

where $\{W_i | i = 1, 2, ..., m\}$ is the set of *n*-dimensional Eigenvectors of S_t corresponding to m largest Eigenvalues. These Eigenvectors have the same dimensions as the original face images, therefore they are also known as Eigenfaces.

b. Details of Fisher's Linear Discriminator

After applying the projection y_k, $S_b = \sum_{i=1}^{c}(\mu_i - \mu)(\mu_i - \mu)^t$ is the between class scatter matrix, and $S_w = \sum_{i=1}^{c}\sum_{X_k \in X_i}(X_k - \mu_i)(X_k - \mu_i)^t$ is the within class scatter matrix.

$\mu_i = \frac{1}{N_i}\sum_{X_k \in \varsigma_i} X_k$ is the mean image of class X_i, and N_i is the number of samples in the class X_i. If S_w is nonsingular, the optimal projection W_{opt} is chosen as the matrix with orthogonal columns which maximizes the ratio of the determinant of the between class scatter matrix of the projected samples to the determinant of the with-in class scatter matrix of the projected samples as:

$$W_{opt} = \arg\max_w \frac{|W^t S_b W|}{|W^t S_w W|}$$

where $\{W_i | i = 1, 2, ..., m\}$ is the set of generalized Eigenvectors of S_b and S_w corresponding to the m largest generalized eigenvalues $\{\lambda_i | i = 1, 2, ..., m\}$ as $S_b W_i = \lambda_i S_w W_i \qquad i = 1, 2, 3, ..., m$

with at most $c-1$ non-zero generalized eigenvalues, and so the upper bound on m is $c-1$ where c is the number of classes.

c. Details of Fisherfaces (PCA + FLD) Method

For the formulation of Fisherfaces; faces are projected with PCA to reduce the dimension of the feature space with rank of S_w to be maximum

$Rank(S_w) \leq N-c$ and then applying the standard FLD to reduce the dimension to $c-1$ as discussed above. Now the $W_{opt} = W_{fld}^t W_{pca}^t$

Where, $W_{pca} = \arg\max_w |W^t S_t W|$

and $W_{fld} = \arg max_w \dfrac{|W^t W_{pca}^t S_b W_{pca} W|}{|W^t W_{pca}^t S_w W_{pca} W|}$

The optimization of W_{pca} was performed over $n \times (N-c)$ matrices with orthogonal columns, and optimization of W_{fld} was performed over $(N-c) \times m$ matrices with orthogonal columns. While computing the W_{pca} we have thrown away only the smallest $c-1$principal components.

Further details about this method can be found in the literature of Belhumeur et al. [15]. In our experiment, we have used 2 features per class for both test and training data. We set the values equals to 50 for both FLD and PCA as Eigen vectors and discriminators respectively.

4. Experimental Results and Analysis

In this section we will discuss about the dataset that we have used for our experiment. Results for normalized dataset by six different methods applied on three different custom made dataset namely, normal, noisy and shaved eyebrow dataset. To sum up we will have eighteen recognition rate results and will be discussed in last sub-chapter of this chapter.

A. Facial Dataset

The EURECOM Kinect Face Dataset [30] was used for experimentation, having color and depth images obtained using Kinect device. The EURECOM Kinect Face dataset contains face images of 52 people (14 females, 38 males) taken in two sessions. In each session, the people are captured with nine states (neutral, smile, open mouth, left profile, right profile, occlusion eyes, occlusion mouth, occlusion paper, light on), besides the depth image, the raw depth level sensed by Kinect is also provided in a .txt file for better precision. The dataset also includes 6 manually located landmark points on the face (left eye, right eye, tip of the nose, left and right side of the mouth, the chin). Structure of the dataset is explained in the Figure. 7.

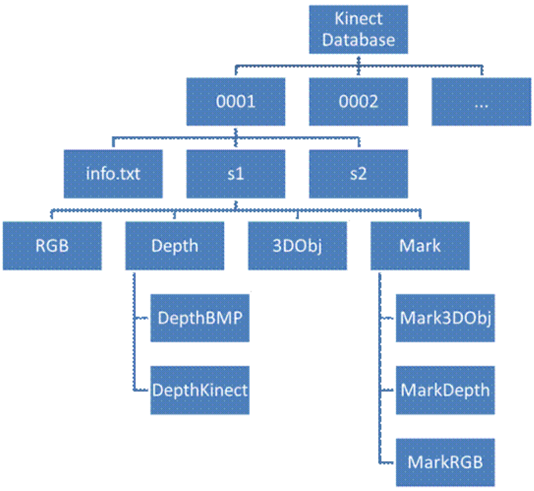

Figure. 7 Structure of the EURECOM Kinect Face Dataset [30]

We have used only the information about position of eyes to normalized the face image for tilt and angle variations. For every subject we treat s1 as test data and s2 as training data in our recognition experiment.

B. Dataset Normalization

We have used position of eyes as the index to normalize the face position of every test and training image. Neutral face poses had variations of 10 degrees up down, and 15 degrees left right movement from the center of face and they were not corrected for the tilt and size in the EURECOM Kinect Face Dataset.

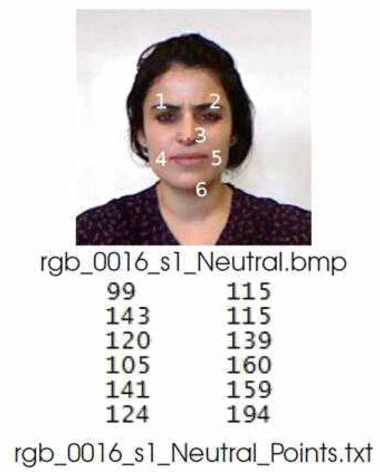

Figure. 8 The marker point positions in the neutral image and the corresponding .txt file [30]

For correct face registration we have utilized the “register_face_based_on_eyes” algorithm from the Ph.D. face recognition toolbox of Matlab® [35, 36]. Position of eyes information was already available in the EURECOM Kinect Face Dataset as shown in the Figure. 8. Parameters of normalization methods are explained in Table 2 on next page.

Table 2. Normalization Parameters

S. No.	Illumination		Parameters for Ph.D. Face Detector Tool [35, 36]				Frequency
	During Fusion	After Face Detection	Forehead	Chin	Left Cheek	Right Cheek	All Dataset
1	no	no	0.5	1.6	0.7	0.7	3
2	no	yes	0.5	1.6	0.7	0.7	3
3	no	yes	0.5	1.6	0.8	0.8	3
4	yes	no	0.5	1.6	0.8	0.8	3
5	yes	yes	0.5	1.6	0.7	0.7	3
6	yes	yes	0.5	1.6	0.8	0.8	3

Table 2. Parameters of normalization of face ROI for face detection and illumination reduction

C. Recognition Results

In this section we will discuss some of the significant results one by one according to the type of dataset used. Dataset are generated by the normalization methods explained above in table 2. Normalization parameters in the following sub-sections can be identified by the serial numbers of table 2.

i. With Normal Dataset

Recognition was done through the Fisherfaces algorithm by running it on normal, fusion, and fusion with Gaussian smoothing dataset and final recognition

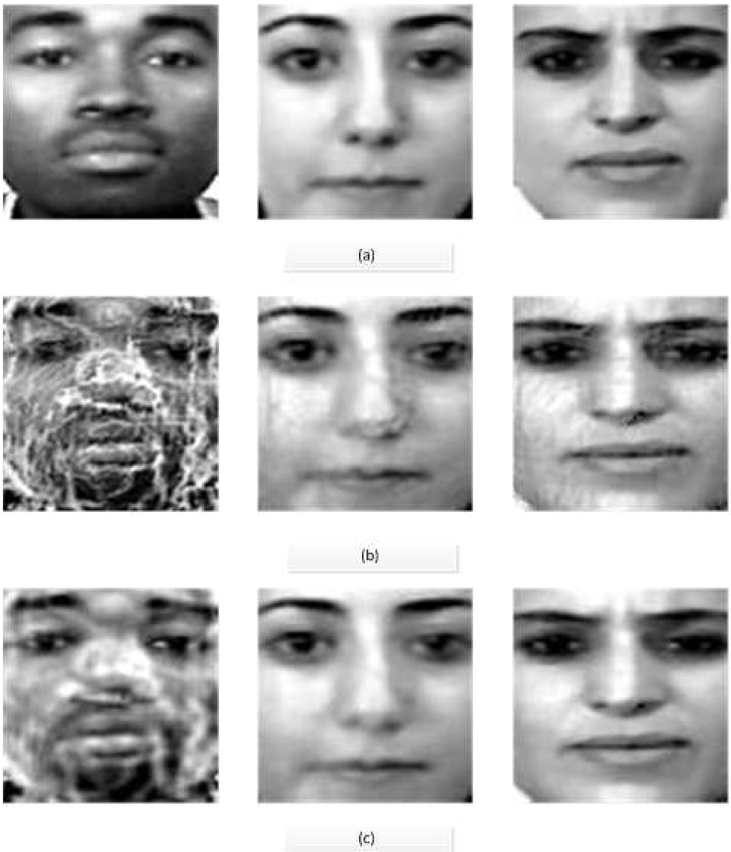

Figure. 9 Result for normal dataset (a) grayscale (b) fusion, and (c) fusion with Gaussian smoothing

rates were analysed. We have analyzed that the recognition rate of grayscale and fusion with Gaussian smoothing dataset are very high as compared to fusion dataset. We have achieved up to 95% correct recognition results for normal dataset with the normalization parameters set to serial number 2 of table 2. Figure. 9 shows the results of normalization for each dataset in normal condition. Figure. 10 shows three best recognition rate plots out of six for each dataset in normal condition.

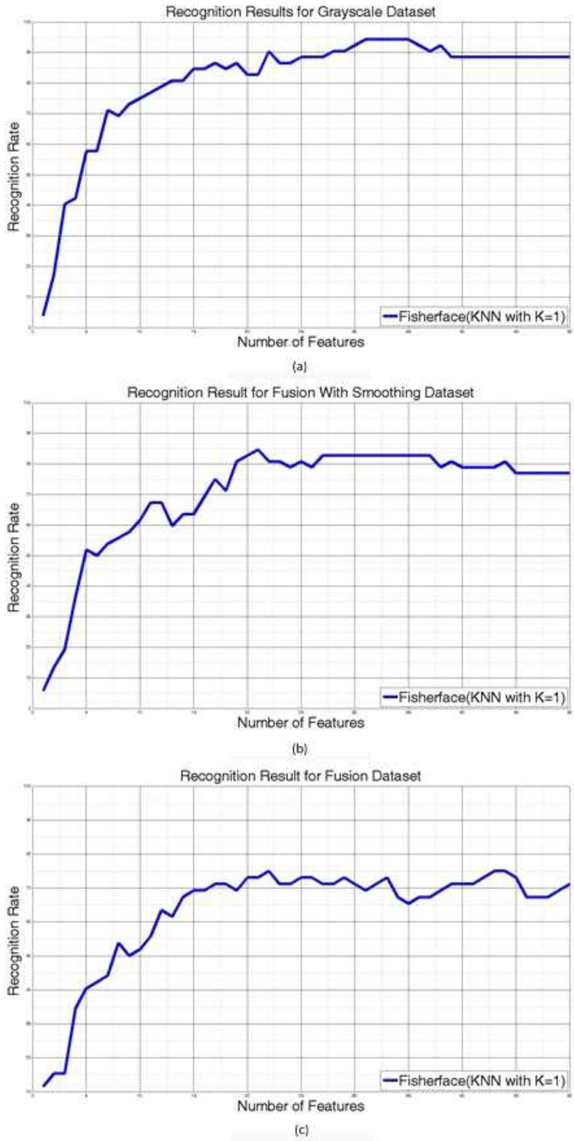

Figure. 10 The recognition rate plots for grayscale (10a), fusion with gaussian smoothing (10b), and fusion (10c) dataset

ii. With Manually Inserted Noise on Cheeks Dataset

After manually inserting the noise on cheeks in grayscale dataset we have achieved up to 87% correct recognition results with the normalization parameters set to serial number 2 of table 2. Figure. 11 Show the normalization result for each noisy dataset. Figure. 12 shows the recognition plots of each noisy dataset.

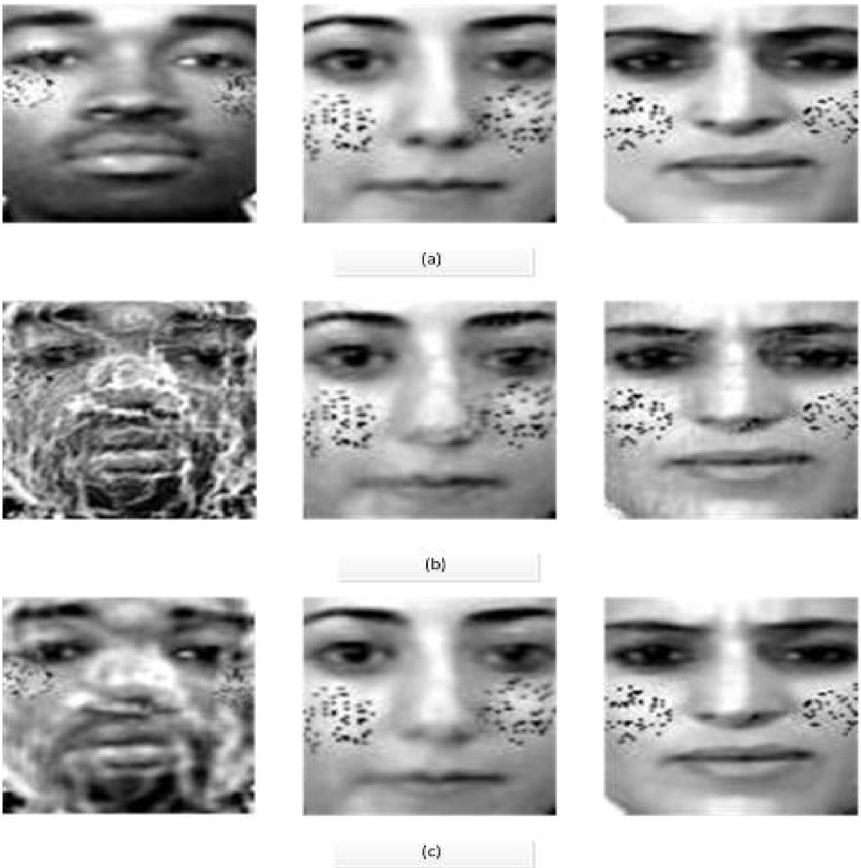

Figure. 11 Dataset with noise inserted on cheeks (a) grayscale (b) fusion, and (c) fusion with Gaussian smoothing

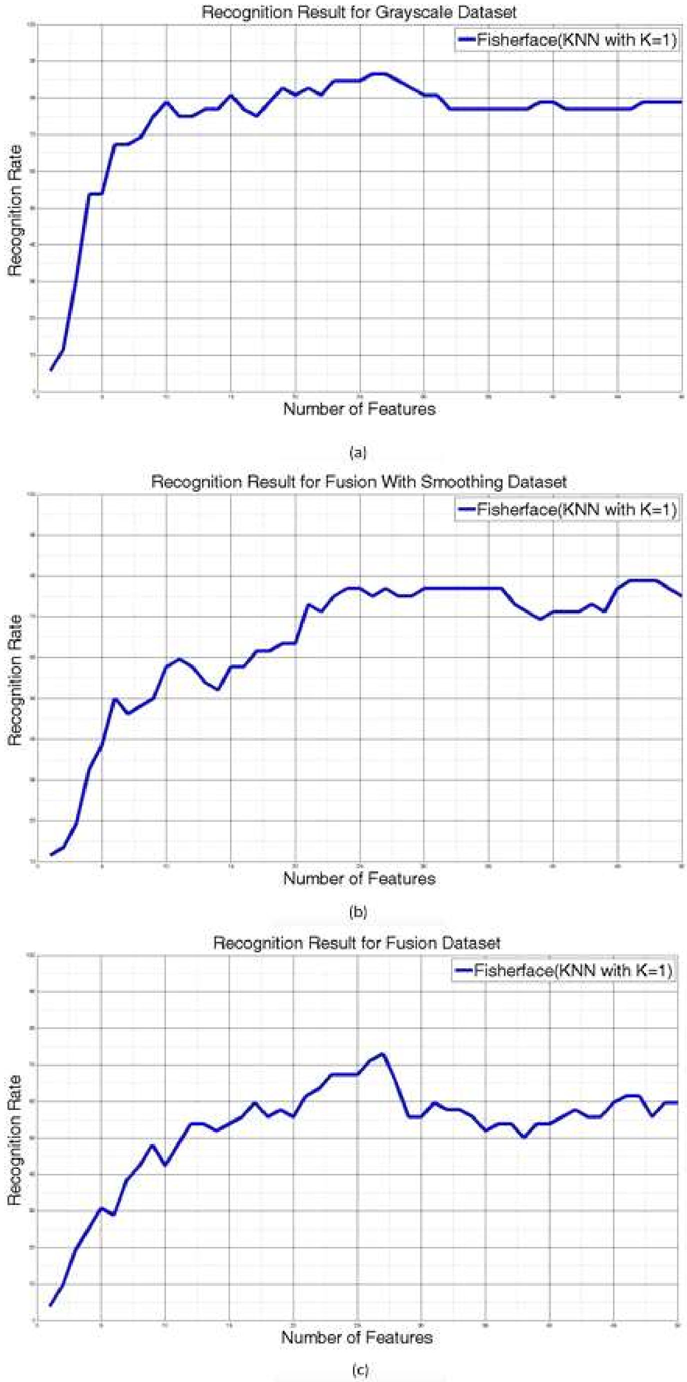

Figure. 12 The recognition rate plots for grayscale (12a), fusion with gaussian smoothing (12b), and fusion (12c) dataset.

iii. With Manually Shaved Eyebrows Dataset

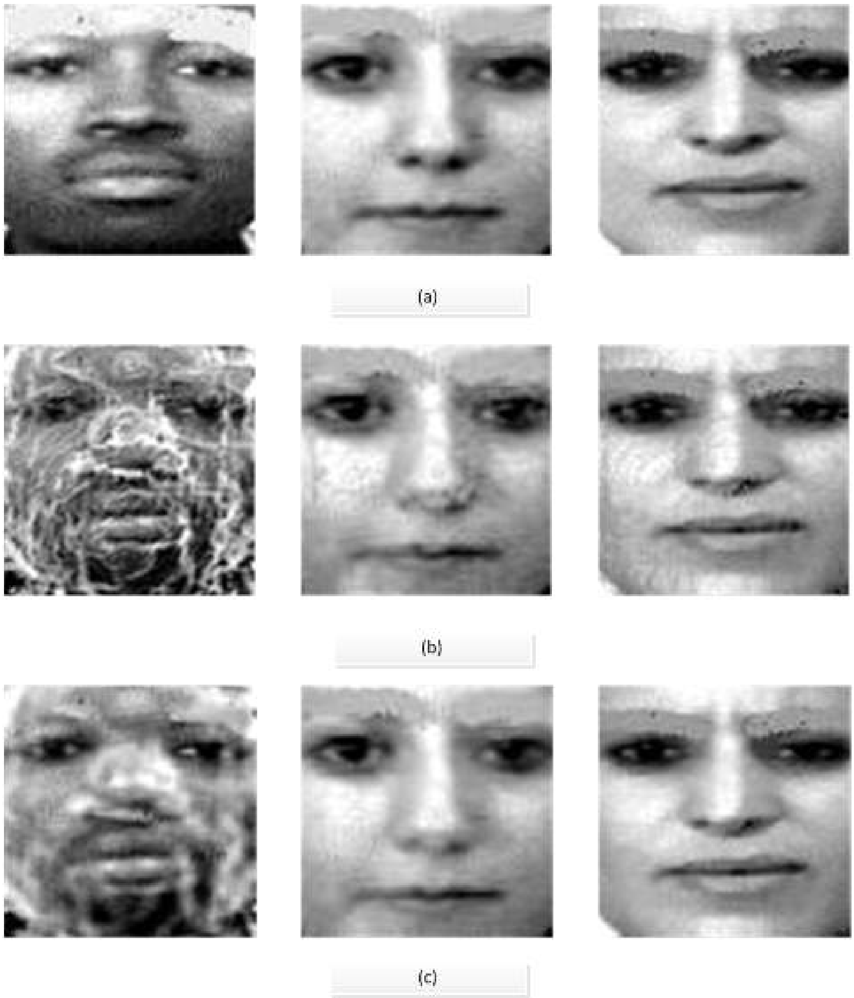

Figure. 13. Manually normalized data with shaved eyebrows

In this part of our experimental analysis we have further verified that if we change the effective area of eye brows in normalization and/or shaved the eyebrows manually, recognition rate changes dramatically; but our algorithm for alpha blending based fusion gives promising results which are better than of grayscale. Hence proved the effectiveness of our proposed fusion method with Gaussian smoothing scheme. We have achieved up to 64% correct recognition results with the normalization parameters set to serial number 3 of table 2.

Shaved eyebrow input of face recognition system shown above in the Figure. 13. Figure. 14 shows the recognition plots of each dataset after manually shaving the eyebrows before fusion.

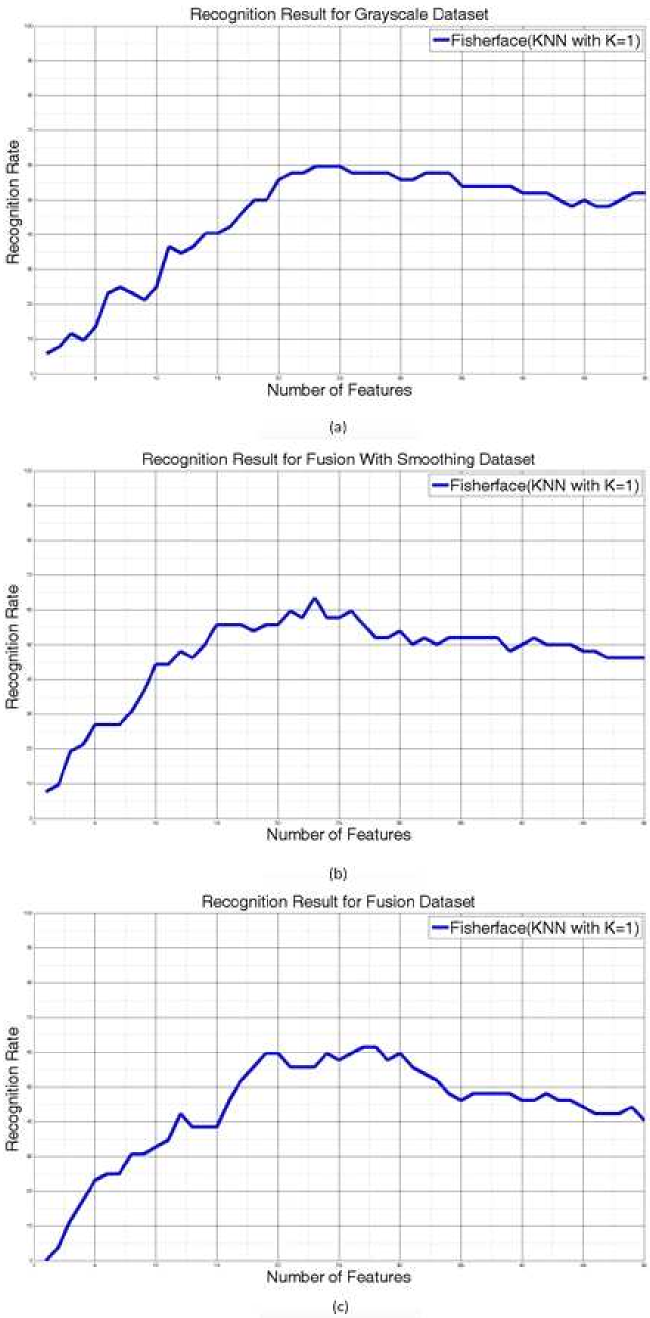

Figure.14. The recognition rate plots for grayscale (14a), fusion with gaussian smoothing (14b), and fusion (14c) dataset.

D. Comparative Analysis of Results

In this section we will present all the eighteen results for every normalization condition according to the values mentioned in table 2 for each dataset. We have compare them graphically as well for better visual understanding of our approach.

Sr. No.	dataset	Grayscale	Fusion With Gaussian	Fusion
1	Normal	80	68	66
2	Normal	95	85	76
3	Normal	88	82	70
4	Normal	80	76	68
5	Normal	87	78	70
6	Normal	86	78	70

Table. 3 Recognition rates of normal dataset in details

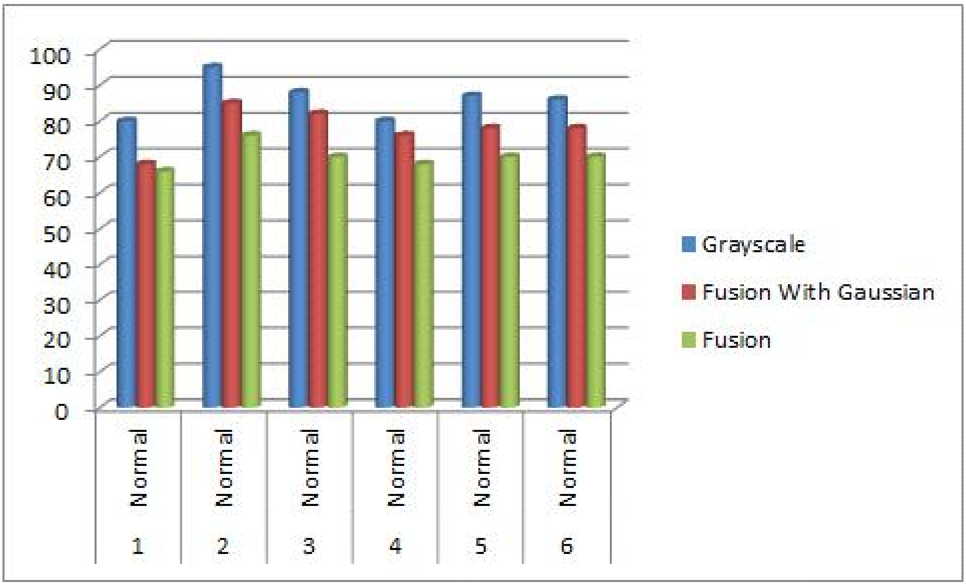

Figure 15. Summary of recognition results by using six different normalization methods on our normal dataset

Sr. No.	dataset	Grayscale	Fusion With Gaussian	Fusion
1	Noise	7.7	5.8	7.8
2	Noise	87	79	74
3	Noise	78	70	64
4	Noise	7.8	7.8	5.7
5	Noise	78	79	66
6	Noise	77	76	64

Table. 4 Recognition rates of noisy dataset in details

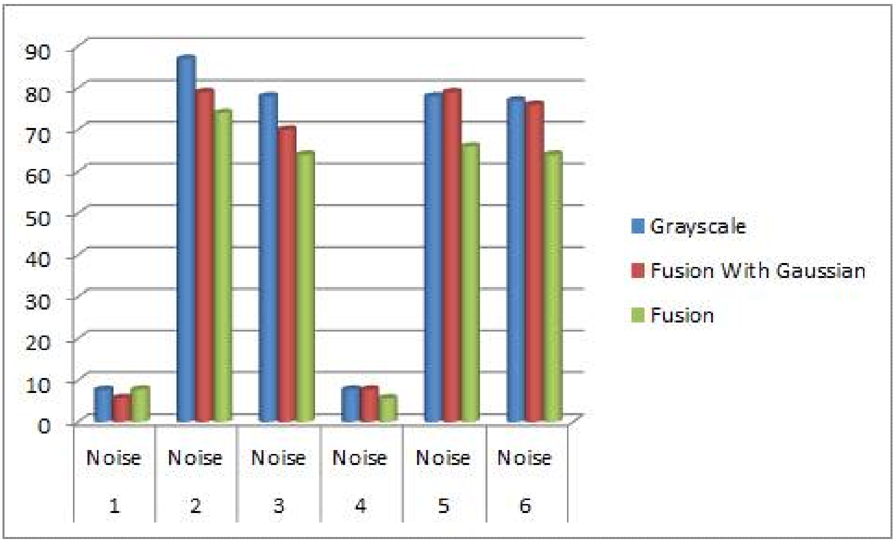

Figure 16. Summary of recognition results by using six different normalization methods on our noisy dataset

Sr. No.	dataset	Grayscale	Fusion With Gaussian	Fusion
1	Shaved Eyebrow	13	5.8	5.8
2	Shaved Eyebrow	67	66	66
3	Shaved Eyebrow	60	64	63
4	Shaved Eyebrow	13	11.8	7.8
5	Shaved Eyebrow	60	64	64
6	Shaved Eyebrow	60	64	62

Table. 5 Recognition rates of shaved eyebrows dataset in details

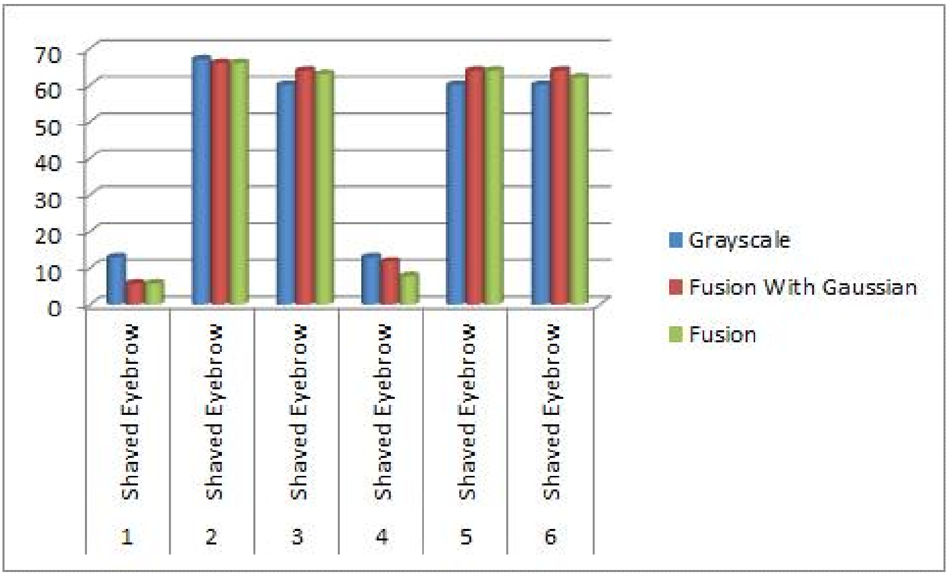

Figure 17. Summary of recognition results by using six different normalization methods on our shaved eyebrows dataset

E. Optimal Condition for All Dataset

After carefully analysing the comparison of different normalization conditions on each dataset as shown in previous section, we can conclude that the following set of normalization parameters generate the optimal results for all the dataset.

Sr. No.	dataset	Grayscale	Fusion With Gaussian	Fusion
5	Shaved Eyebrow	60	64	64
5	Noise	78	79	66
5	Normal	87	78	70

Table. 6 Optimal Condition for all Dataset

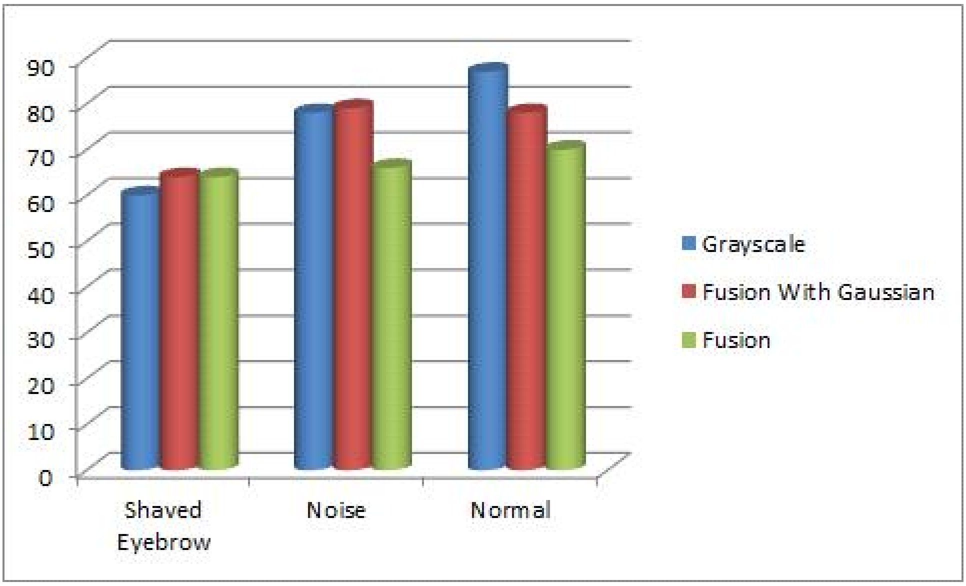

Figure 18. Summary of recognition results by using optimal normalization parameters for all dataset

5. Conclusion and Future Work

This system was our first step towards the final goal of developing a robust multimodal 2D-3D face recognition system using Kinect sensors which can perform well in the noisy environments. Performance can be enhanced by adding more rules for data normalization and adjusting the threshold of recognition parameters. We have to maintain a balanced tradeoff between speed and accuracy of the system. In future work we propose to use more noisy, occluded, and random face poses of RGB data for fusion and check their effects on recognition rates with the proposed fusion algorithm. We hope with a better dataset our algorithm can produce more promising results and higher recognition rates can be achieved.

Bibliography

[1] Cook, J., McCool, C., Chandran, V., Sridharan, S., "Combined 2D/3D Face Recognition Using Log-Gabor Templates," Video and Signal Based Surveillance, 2006. AVSS '06. IEEE International Conference on , vol., no., pp.83,83, Nov. 2006

[2] Riya. Riya photo search. http://www.riya.com/, 2005

[3] K. Bowyer, K. Chang, P. Flynn, “A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition”, Computer Vision and Image Understanding, Vol. 101, No. 1. (January 2006), pp. 1-15

[4] P. Jonathon Phillips, Patrick J. Flynn, Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin Hoffman, Joe Marques, Jaesik Min, andWilliam Worek, “Overview of the face recognition grand challenge,” in CVPR ’05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) - Volume 1, Washington, DC, USA, 2005, pp. 947–954, IEEE Computer Society

[5] Kevin W. Bowyer, Kyong Chang, and Patrick Flynn, “A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition,” Computer Vision and Image Understanding, vol. 101, no. 1, pp. 1–15, 2006, TY – JOUR.

[6] Kyong I. Chang, K.W. Bowyer, and P.J. Flynn, “Adaptive rigid multi-region selection for handling expression variation in 3D face recognition,” in Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on, 2005, vol. 3, p. 157

[7] Thomas Maurer, David Guigonis, Igor Maslov and Bastien Pesenti, Alexei Tsaregorodtsev, David West, and Gerard Medioni, “Performance of Geometrix ActiveIDTM3D Face Recognition Engine on the FRGC Data,” in CVPR ’05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision

and Pattern Recognition (CVPR'05), Washington, DC, USA, 2005, IEEE Computer Society

[8] Michael Husken, Michael Brauckmann, Stefan Gehlen, and Christoph Von der Malsburg, "Strategies and benefits of fusion of 2D and 3D face recognition," in CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops, Washington, DC, USA, 2005, p. 174, IEEE Computer Society

[9] Qiu Jin; Jieyu Zhao; Yuanyuan Zhang, "Facial feature extraction with a depth AAM algorithm," Fuzzy Systems and Knowledge Discovery (FSKD), 2012 9th International Conference on , vol., no., pp.1792,1796, 29-31 May 2012

[10] Chuanjun Wang; Xuefeng Bai; Tiejun Zhang; XiaMu Niu, "A Fourier series based expression deformation model for 3D face recognition," Natural Computation (ICNC), 2012 Eighth International Conference on , vol., no., pp.130,133, 29-31 May 2012

[11] Jelsovka, D.; Hudec, R.; Breznan, M.; Kamencay, P., "2D-3D face recognition using shapes of facial curves based on modified CCA method," Radioelektronika (RADIOELEKTRONIKA), 2012 22nd International Conference , vol., no., pp.1,4, 17-18 April 2012

[12] Mian, A., "Robust realtime feature detection in raw 3D face images," Applications of Computer Vision (WACV), 2011 IEEE Workshop on , vol., no., pp.220,226, 5-7 Jan. 2011

[13] Peter W. Hallinan, Gaile G. Gordon, A. L. Yuille, Peter Giblin, and David Mumford. Two- and Three-Dimensional Patterns of the Face. A. K. Peters, Ltd., Natick, MA, USA. 1999.

[14] Kusuma, G.P.; Chin-Seng Chua; Toh, H., "Recombination of 2D and 3D Images for Multimodal 2D + 3D Face Recognition," Image and Video Technology (PSIVT), 2010 Fourth Pacific-Rim Symposium on , vol., no., pp.76,81, 14-17 Nov. 2010

[15] Belhumeur, P.N.; Hespanha, J.P.; Kriegman, D., "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.19, no.7, pp.711,720, Jul 1997

[16] Newcombe, Richard A.; Davison, Andrew J.; Izadi, S.; Kohli, P.; Hilliges, Otmar; Shotton, J.; Molyneaux, David; Hodges, Steve; Kim, David; Fitzgibbon, A., "KinectFusion: Real-time dense surface mapping and tracking," Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium, pp.127,136, 26-29 Oct. 2011

[17] Pawan Sinha, Benjamin Balas, Yuri Ostrovsky, Richard Russell “Face recognition by humans: 20 results all computer vision researchers should know about” Proceedings of IEEE Vol. 94, No. 11, 2006

[18] J. Sadr, I. Jarudi, and P. Sinha, “The role of eyebrows in face recognition,” Perception, vol. 32, pp. 285–293, 2003

[19] R. Hietmeyer. Biometric identification promises fast and secure processing of airline passengers. The International Civil Aviation Organization Journal, 55(9):10–11, 2000

[20] P. Hammond, T. Hutton, J. Allanson, A. Shaw, and M. Patton. 3D digital stereophotogrammetric analysis of noonan syndromed. In British Human Genetics Conference. Blackwell Synergy, 2002

[21] A. O’Toole, T. Price, T. Vetter, J. Bartlett, and V. Blanz. Three-dimensional shape and two-dimensional surface textures of human faces: The role of ”averages” in attractiveness and age. Image and Vision Computing Journal, 18(1):9–19, 1999

[22] Theodore Papatheodorou, 3D Face Recognition Using Rigid and Non-Rigid Surface Registration, Ph.D. Thesis, University of London, July 2006

[23] Nishanth Vincent, “INTERACTIVE FACE RECOGNITION,” Ph.D. Thesis, Fairfield University

[24] E. Hjelmas and B. K. Low. Face detection: A survey. Computer Vision and Image Understanding, 83:236-274, 2001

[25] MESA Imaging. SwissRanger sensor. http://www.mesa-imaging.ch

[26] Microsoft Corporation. The Kinect sensor. http://www.xbox.com/en-US/kinect

[27] P. Viola and M. J. Jones. Robust real-time face detection. International Journal of Computer Vision, 57(2):137-154, 2004

[28] M. Dixon, F. Heckel, R. Pless, and W. D. Smart. Faster and more accurate face detection on mobile robots using geometric constraints. In IEEE/RSJ International Conference on Robots and Systems (IROS 2007), pages 1041-1046, 2007

[29] Burgin, W., Pantofaru, C, Smart, W.D, "Using depth information to improve face detection," Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on , vol., no., pp.119,120, 8-11 March 2011

[30] Tri Huynh, Rui Min, Jean-Luc Dugelay, "An Efficient LBP-based Descriptor for Facial Depth Images applied to Gender Recognition using RGB-D Face Data," in ACCV Workshop on Computer Vision with Local Binary Pattern Variants 2012, Daejeon, Korea, November 2012

[31] M.Tanaka, mtanaka@ctrl.titech.ac.jp, Face Parts Detection, Matlab file exchange, 2012

[32] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, 2D and 3D Face Recognition: A survey, Pattern Recognition Letters, Vol 28, I 14, pp.1885–1906, 2007

[33] B. Delaunay: Sur la sphère vide, Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk, 7:793–800, 1934

[34] http://en.wikipedia.org/wiki/Delaunay_triangulation#cite_note-Delaunay1934-1

[35] Struc V. Pavesic, N. “Gabor-Based Kernel Partial-Least-Squares Discrimination Features for Face Recognition”, Informatica (Vilnius), vol. 20, no. 1, pp. 115-138, 2009

[36] Struc V., Pavesic, N. “The Complete Gabor-Fisher Classifier for Robust Face Recognition”, EURASIP Advances in Signal Processing, vol. 2010, 26 pages, doi:10.1155/2010/847680, 2010

Appendix 1

Research Work and Published Papers During MS Degree

Journals/Book Chapters

- **M. N. I. Qureshi,** and S.-W. Lee, "RGB-D Fusion of Kinect Sensor Data for Face Recognition," EURASIP International Journal on Image and Video Processing (JIVP) submitted (to be appear)

- **M. N. I. Qureshi**, J.-T. Lee, and S.-W. Lee, "Facial Identity Encryption with Kinoform and Phase-key Watermarking for Homeland Security Agencies," Lecture Notes in Computer Science: Multidisciplinary Research and Practice for Information Systems, Vol. 7465, August 2012, pp. 525-533

- **M. N. I. Qureshi**, S. Naveed, and S.-W. Lee, "CUDA Implementation of Kernel Independent Component Analysis for Reflection Removal," Journal of the Korea Institute of Next Generation Computing, Vol. 8, No. 3, June 2012, pp. 78-87

Domestic and International Conference Proceedings:

- **M. N. I. Qureshi,** and S.-W. Lee, "RGB-D Fusion of Kinect Sensor Data for Face Recognition," Proc. of 2nd International conference of Smart Media Applications (SMA2013) Kota Kinabalu, Malyasia, (to be appear on) October 14-17 2013

- **M. N. I. Qureshi,** and S.-W. Lee, "Auto Photo Composition Guide for Human Photography Using Kinect," Proc. of 2nd International conference of Smart Media Applications (SMA2013) Kota Kinabalu, Malyasia, (to be appear on) October 14-17 2013

- Heon Jeong, **M. N. I. Qureshi,** and S.-W. Lee, "Tire Overlapping Inspection System based on Laser-lined Curvature Analysis," Proc. of 2nd International conference of Smart Media Applications (SMA2013) Kota Kinabalu, Malyasia, (to be appear on) October 14-17 2013

- **M. N. I. Qureshi** and S.-W. Lee, "RGB-D Fusion Based on Facial Normal Vectors," Proc. of 2013 KMMS Spring Conference, Daejeon, Korea, June, 1st ,2013

- **M. N. I. Qureshi** and S.-W. Lee, "Face Recognition Using Both RGB and Depth Information from Kinect Sensor," Proc. of 2012 KOCTA Fall Conference, Seoul, Korea, November 2012, pp. 93-94

- **M. N. I. Qureshi,** J.-T. Kim, and S.-W. Lee, "A Study on Preventing from a forgery of Personal Identification Using Kinoform and Phase-key," Proc. of The Korean Physical Society Fall Conference, Phoenix Park, Kangwon-do, Korea, October 2012

- **M. N. I. Qureshi,** J.-T. Lee, and S.-W. Lee, "Facial Identity Encryption with Kinoform and Phase-key Watermarking for Homeland Security Agencies," International Workshop on Security and Cognitive Informatics for Homeland Defence, Prague, Czech Republic, August 2012, pp. 525-533

- **M. N. I. Qureshi,** J.-E. Lee, and S.-W. Lee, "Robust Classification Techniques for Connection Pattern Analysis with Adaptive Decision Boundaries Using CUDA," International Conference on Cloud Computing and Social Networking, Bandung, Indonesia, Apr 2012, pp. 24-28

- **M. N. I. Qureshi** and S.-W. Lee, "Automatic Composition Guide for Digital Portrait Photo Capture Using Kinect," Proc. of KISM Spring Conference, Jeju, Korea, April 2012, pp. 283-285 *(Awarded as one of the best publications)*

Memo