# Performance Evaluation of Face Recognition Using Frames of Ten Pose Angles

By

Sherif M. El Seuofi

Submitted in Partial Fulfillment of the Requirements

For the degree of

Master of Computing and Information Systems

In the

Department of Computer Science and Information Systems

Youngstown State University

[December, 2007]

# Performance Evaluation of Face Recognition Using Frames of Ten Pose Angles

Sherif M EL Seuofi



Signature:

Sherif M. EL Seuofi	Date

Approvals:

Dr. Yong Zhang	Date

Dr. John Sullins	Date

Dr. Robert Kramer	Date

Peter J.Kasvinsky, Dean of School of Graduate Studies & Research	Date

# Performance Evaluation of Face Recognition Using Frames of Ten Pose Angles

Sherif EL Seuofi

## Abstract

Face recognition has received much attention recently in the biometrics research. Many studies have shown improvement in recognition rate when 2D and 3D faces were combined. However, the use of the 3D face has a few limitations such as that the 3D data requires much more storage space and long processing time. Therefore, there is a strong interest to explore new methods that can provide similar or better results in the face recognition.

This thesis presents an experimental study by using a sequence of rotating head videos under two different lighting conditions, regular indoor lighting and strong shadow lighting. The experiment were carried out using two sets of data, the first set of over 100 subjects and the second set of 47 subjects. Very promising results have been observed in terms of the recognition performance measured by the cumulative characteristics curves

**Keywords**: Face recognition, Biometrics, video-based face recognition.

# Contents

**Abstract……………………………………………………………………….iii**

**Contents……………………………………………………………………….iv**

**List of Figures……………………………………………………………...vi**

**List of Tables…………………………………………………………………vii**

**Acknowledgments…………………………………………………………...viii**

**1 Introduction ........................................................................................ 1**

**1.1 Statement of the Problem ..................................................................... 1**

**1.2 Research Methods .................................................................................... 2**

**1.3 Related Works .......................................................................................... 3**

**1.4 Objective .................................................................................................... 4**

**1.5 Reading Guide .......................................................................................... 4**

**2 Biometrics Overview ............................................................................. 5**

**3 Face Recognition ..................................................................................... 8**

**3.1 Face Recognition Steps ........................................................................... 8**

**3.2 Face Recognition Methods ...................................................................... 8**

**4 Experiment Setup .................................................................................... 10**

**4.1 Video Collection ....................................................................................... 10**

**4.2 Setup ........................................................................................................... 11**

**4.3 Subjects ...................................................................................................... 14**

**4.4 Procedure ................................................................................................... 14**

**4.5 Frame Selection Criteria ......................................................................... 15**

**4.6 Fusion on Image Level ........................................................ 19**

**5 Experiment Results ........................................................ 21**

**5.1 First Condition Test: ........................................................ 21**

**5.2 Second Condition Test: ........................................................ 21**

**5.3 Discussion ........................................................ 22**

**5.4 Summary ........................................................ 22**

**Bibliography ........................................................ 24**

# List of Figures


**Figure 1: Correct Verification Claim ..... 5**

**Figure 2: False Verification Claim..... 5**

**Figure 3 : Watchlist Example 1..... 6**

**Figure 4: Watchlist Example 2..... 7**

**Figure 5: Identification Example ..... 7**

**Figure 6: Sample frames from two video sequences..... 11**

**Figure 7: Geometry of camera setup. ..... 11**

**Figure 8  Regular lighting...... 12**

**Figure 9: Regular Light in Frontal and Profile view..... 12**

**Figure 10: Shadow lighting...... 13**

**Figure 11: Shadow Light in Frontal and Profile view..... 13**

**Figure 12: Selection of two facial markers for image normalization..... 19**

**Figure 13 : Samples of fused images. ..... 20**

# List of Tables

**Table 1: DATA COLLECTIONS AND LIGHTING CONDITIONS....................................................10**

**Table 2: FRAME SELECTION CRITERIA ..........................................................................................16**

**Table 3: GROUND TRUTH FOR FRAME SELECTION......................................................................17**

# Acknowledgments

Several people have contributed one way or another to the result of this thesis, and I would like to use this opportunity to thank them all for their guidance and support through this process.

I would like to thank Dr. Zhang for introducing me to this subject and for giving me the opportunity to work on this research, for generating my interest in this area of study, for his support and encouragement.

Very special thanks go also to Dr. Sullins who gave me his constant support and encouragement during my studies. Without his support, it would not have been possible for me to pursue the master degree at the Youngstown State University. Sincere thanks also go to Dr. Bob Kramer for being on my committee and helping improve the quality of the thesis.

I would like to extend my deepest gratitude to my parents and sister for their constant support throughout my schooling.

Finally I would like to thank my wife for being supportive, for her understanding during this tough period and for encouraging me in pursuing the master degree.

# 1 Introduction

Since the event of 11 September 2001, security is a topic that has received much attention. We keep hearing in the news about the intensified security measures at airports and seaports. Among many security related techniques, there is one solution that utilizes unique biological features: Biometrics. Biometrics refers to the automatic identification of a living person based on physiological or behavioral characteristics. There are many types of biometric technologies available, face recognition, fingerprint recognition, hand geometry recognition, iris recognition, vein recognition, voices and signature recognition.

The biometric identification is preferred over traditional methods because biometrics can identify individuals based on life-long signatures. These other security methods can be lost or stolen and potentially get into the hands of unauthorized users. Biometrics provides a key that cannot be lost or stolen. Biometrics is also convenient because users do not have to carry a card or remember a password. Biometric can eliminate the need for passwords or can be used to consolidate existing passwords and password maintenance which accounts for 40% of corporate help desk calls.

## *1.1 Statement of the problem*

The 3D face recognition has received much attention recently in the biometrics research community. Almost all the studies have shown improvement in recognition accuracy when the 2-D and the 3-D faces are combined. 3D faces are less affected by illumination and pose variation than the 2-D. However, as pointed out by Bowyer *et al* [1], the use of 3D face has a few limitations:

- Current sensors have limited operation ranges, often less than two meters.
- 3D data require much more storage space and long processing time.
- Acquisition is often not fully automated and may need user intervention.

It is not expected that those technical issues will be solved in the near future. Therefore, there is a strong interest to investigate new methods that can be used in the face recognition.

## 1.2 Research methods

This thesis presents a method that can be used in a video sequence in which a subject gradually rotates his or her head from the frontal view. The front view angle is regarded as 0 degree rotation and the profile view angle as 90 degree rotation. Each subject has video sequences under two different illumination conditions, one is the regular lighting and the next is strong shadow lighting. Multiple video frames are then extracted at different pose angles. This video-based method has several advantages:

- The video sequence acquired by the high definition camcorder can provide quality 3D data for face recognition, which means that it can be used in the real world operations such as the check points or the port of entries.

- Not all frames in a rotating head video will be used. The total number of frames can reach a few hundreds in a rotation sequence, and only 10 frames out of the total frames are used by selecting the frontal view as the first frame, the profile view as the last frame and the frames in between are selected based on increasing order of rotation degree.

- Since only 10 frames are selected, we use a very small storage space and much less computational time.

- More importantly, a video sequence of a face with different poses might help reduce the affect of the strong shadow.

## 1.3 Related Works

A literature survey with in-depth discussions of the current developments in 3D methods can be found in [1]. Zhao et al [3] presented a more extensive survey of the existing methods in face recognition. This section gives a brief review of the techniques that are most relevant to this approach.

One motivation of using videos of rotating heads to facilitate face recognition is that, a few recent studies have demonstrated that the multi-sample approach can achieve a performance comparable to that of the multi-modal approach. Using a data set of varying facial expressions and lighting conditions, Bowyer et al [4] reported an improvement in rank one recognition rate from 96.1% with two frames per subject to 100% with four frames per subject. In another study, Thomas et al [5] observed that the recognition rate generally increases as the number of frames per subject increases, regardless of the type of camera being used. They also found that the optimal number of frames per subject is between 12 and 18, given the particular data sets used. However, as noted in [4], “The use of multiple intensity images is of value only if there is some variation between the individual images of a person. And very little is known about how to build the right degree of variation into a multi-sample approach”. This study is an attempt to address certain aspects of the issue raised above with following features:

- Using videos of rotating heads that show continuous pose variations.
- Videos have strong shadows which presents a severe challenge to the recognition algorithm.
- Fusion is performed on the image level.

A large amount of research efforts have been dedicated to video-based face recognition because of the rich temporal information contained in videos. In the very early work [6], the use of a 3D model has been considered important for both tracking and recognition purposes. Various models have been proposed, from geometrical models to more sophisticated deformable models, morphable models and statistical models [7, 8, 9, 10]. Certain 3D models like meshes, point clouds and depth images can be directly used for recognition through either registration minimization or principle component analysis. More frequently, 3D models are used to transform an input 2D image by rendering it so that the face in the resulting image has the desired pose, illumination and expression. The drawbacks of using an explicit 3D model are:

- The accuracy of a reconstructed 3D shape via structure from motion may not be adequate for recognition.
- The computational cost involved in shape rendering, illumination simulation and deformation modeling is high.

Efforts have been made to extract grey level cues (shading, profile curves and silhouette) to aid the 3D model based recognition [11], because intensity variation is often related to an object's shape and its surface reflectance properties, a fact that has been explored in the well known "shape from shading" scheme. A video sequence of a rotating head should contain abundant information about its 3D geometry that, in theory, can be utilized either explicitly or implicitly [12].

## *1.4 Objective*

The main objective of this thesis is to investigate the feasibility of using multiple video frames fused on image level directly for face recognition, without reconstructing a specific 3D model or fitting a generic 3D model.

## *1.5 Reading guide*

Chapter 1 provides a brief introduction about the problem, research method and the related works. Basic terminologies are given in chapter 2. The theory of face recognition is covered in Chapter 3. The experiment setup and results are discussed in Chapter 4 and Chapter 5.

# 2 Biometrics overview

In practice, biometric systems operate in one of three tasks [15, 16]: (1) verification: is the person who they claim to be? (2) Watch-list: is this person in the database? If so, who is he/she? (3) Identification: this person is in the database, but how soon can he/she be found?

*Verification task*: In this case, a user first makes a claim as to his/her identity (e.g., I am John Smith). The biometric system then determines if the user's claim is correct or not. Figure 1 gives a visual example, where the subject on the left makes a claim that he is the person on the right. A classifier can either accept or reject such a claim based on certain recognition criteria.

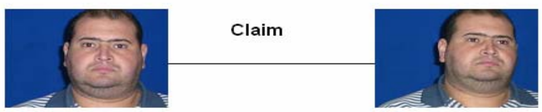



**Figure 1: Correct Verification Claim.**

The performance of a face recogntion system is measured on a scale of 0 to 1, with 1 for the exact match. If the system produces a similarity score of 0.93 for this verification trial, and a preset verification threshold is 0.90, then the face recognition system has correctly determined that the subject in the left picture is the same person in the right picture.

Figure 2 shows a different verification claim. In this example, the subject on the left claims to be the person on the right. The face recognition system returns a similarity score of 0.86 (their facial features are somewhat similar, after all). Again, with a verification threshold of 0.9, the face recognition system determines that the claim is false.

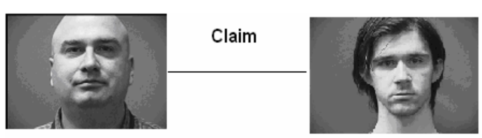



**Figure 2: False Verification Claim.**

*Watch list task*: In this case, the biometric system determines if the individual's biometric signature matches a biometric signature of someone on the watch-list. The individual does not make an identity claim, and in some cases does not even personally interact with the system. An Example could be comparing visitors to the U.S.A against a terrorist database.

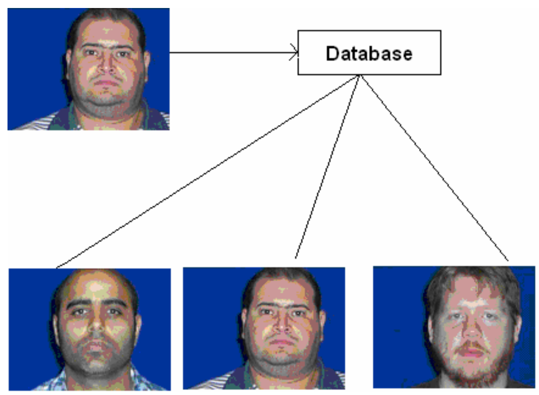



**Figure 3 : The First Watch-list Example.**

Considering the example in Figure 3, where an image (a probe) is an input to the face recognition system database. The system first compares the probe image to each image in the database (gallery). From left to right, if the similarity score for each comparison is 0.86, 0.9 and 0.6, and with a watch-list threshold of 0.85, the recognition system would sound an alarm as one or more of the similarity scores are higher than the threshold. We then look more closely at the subject with the highest similarity score for further analysis.

An alternative setup is that the probe does not have a corresponding match in the database, as illustrated in Figure 4. Again, assuming that the similarity score for each comparison is 0.7, 0.6 and 0.8, from left to right, with the same watch-list threshold of 0.85, no alarm will sound. If the threshold is lowered to 0.75, an alarm would sound the gallery image on the right. Because the match is incorrect, it is called a false alarm.

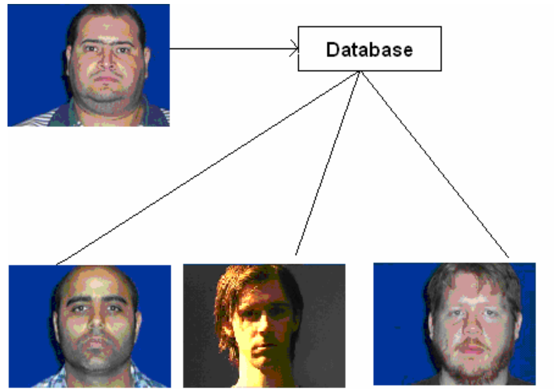



**Figure 4: The Second Watch-list Example.**

*Identification*: Identification is a special case of the watch-list task where we know that every single probe image has a corresponding match in the database. In other words, every single input has already enrolled in the system. So the first question of the watch-list task is answered (is this person in the database?). Identification falls to one question, which one in the system?

First, a biometric signature of an individual is presented to the biometric system as shonw in Figure 5. The recognition system then compares the probe image to each image in the database. If the similarity score for each comparison is 0.86, 0.9 and 0.6, the correct match would be the one with the highest score. If we run many trials, we will know how often the system will return a correct result with the top match.

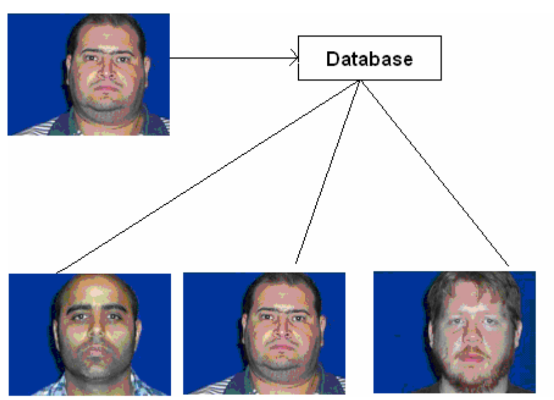



**Figure 5: Identification Example.**

# 3 Face Recognition

## 3.1 Face recognition Steps

Face recognition, like other biometrics, has various implementation details. The process flow in this study has five Steps:

- *Capture an image*: Acquisition can be accomplished by digitally scanning an existing photograph or as a video by using high definition camcorder that can provide 3D information.
- *Find face in image*: This step can be done by using software to detect the location of any faces in the acquired image. This task is difficult and often handled by face detection tools.
- *Image Normalization*: This step represents a way to reduce the influence of various factors such as hair, ear, focus distantce etc.
- *Compare Template*: This step is to compare the images generated at step three with those in a database. In an real application, this process usually gives us a score that indicates how good the match is.
- *Declare Matches*: The final step is determining whether any scores produced are high enough to declare a match. The rules are configurable, so that users can determine how accurate the recognition system is.

## 3.2 Face recognition Methods

There are several methods used in face recognition, some more accurate than others under specific conditions. Recent surveys and reviews on face recognition technologies are provided in [25, 26, 27, 28]. The most popular face recognition methods are eigenface, local feature analysis and elastic graph matching.

- Eigenfaces [29, 30, 25] was developed at Massachusetts Institute of Technology, and was motivated by a technique developed by Sirovich and Kirby in 1987 for efficiently representing pictures of faces using principal component analysis (PCA). Variations of eigenface method are often used as the basic of other face recognition methods. It has been known that this technique does not help a lot at the recognize process and measure similarity between faces. However, according to Woodward et al. [31], the mathematical properties of the eigenface representation and matching process have been demonstrated to achieve a better result in certain environments. Like all facial recognition methods, the eigenface method works best at the frontal image capture situations [24].

- Local feature analysis [30, 24] is another widely used method, and can work well for some changes in the face expression and aging. Local feature analysis refers to a class of algorithms that extract a set of geometrical metrics and distances from facial images and uses those features as the basic for representation and comparison. The features used are typically the eyes, mouth, nose, jaw line, eyebrows and cheeks. These features are represented with their position, size and general outline. The good performance compared with other techniques, are among the factors that has made this technique popular. One drawback for this method is that it is dependent on a relatively constant environment.

- Elastic graph matching [32, 33, 34] has a main advantage that it is invariant to affine transformations and localized changes in facial expressions [35]. Features are extracted at specified locations of face. The distances between these nodes are recorded. Some features are more reliable and important for recognition than others, and hence an approach with the use of weights has been introduced. An extension of the approach uses several images of the same individual, typically from different angels, which is called elastic bunch graph matching [32, 33, 34]. Each node on the graph then contains several values. This improves the recognition because it will be more robust to differences in posture and expressions.

# 4 Experiment Setup

## 4.1 Video Collection

Videos were acquired in two collection groups, the first collection has 101 subjects participated in the video collection process, the second collection being carried out 20 days after the first one (see Table 1). The second collection has 47 subjects participated. The videos of the 47 subjects who enrolled in both collections will be used as gallery and probe sets, while the videos of the remaining 54 subjects who appeared only in the first collection will be used as the training set.

Some subjects changed their appearance between the two groups, such as beards, mustaches, piercing and glasses.

**Table 1: DATA COLLECTIONS AND LIGHTING CONDITIONS**

| | **First Collection** | **Second Collection** |
|---|---|---|
| Subjects | 101 subjects. | 47 subjects, |
| Condition One | **Regular indoor light**.<br>Rotation: 90 degrees.<br>Expression: neutral, smile, angry, surprise. | **Regular indoor light**.<br>Rotation: 90 degrees.<br>Expression: neutral, smile, angry, surprise. |
| Condition Two | **Strong Shadow.**<br>Rotation: 90 degrees.<br>Expression: neutral, smile, angry, surprise. | **Strong Shadow.**<br>Rotation: 90 degrees.<br>Expression: neutral, smile, angry, surprise. |

In each collection session, a subject sat on a rotating chair in front of a camcorder against a blue background curtain. The subject slowly turned his/her body by 90 degrees (from the frontal view to the profile view). The turning process was done twice, first with the regular indoor light, and then with strong shadows cast by a headlight. A few samples obtained under the two lighting conditions are shown in Fig. 6.

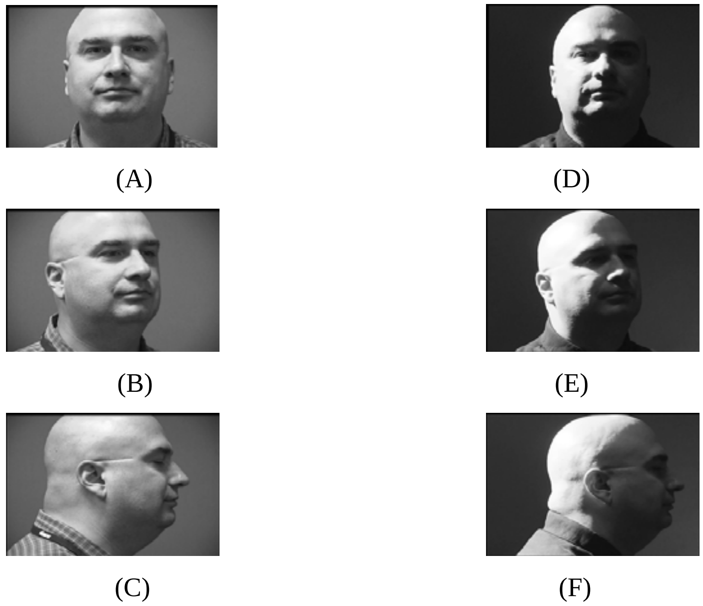



**Figure 6: Sample frames from two video sequences**

## 4.2 Setup

The videos are captured in an indoor setting. The camera is placed at a distance of 2-3 feet from the subject (Figure 7). The videos only capture the rotating face of the subjects from degree 0 to degree 90.

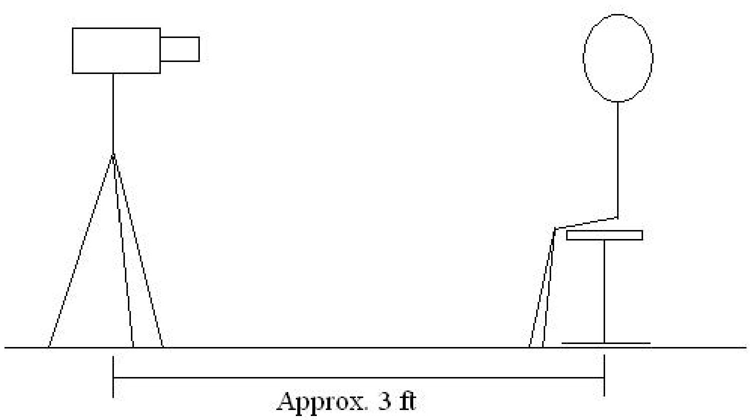



**Figure 7: Geometry of camera setup.**

Two lighting conditions are considered in this data collection process: regular light and shadow. Figure 8 shows the setup with regular lighting condition. Regular light depicts the normal indoor ceiling tube light. This setup allows equal amount of light to fall on the face without creating severe shadow effect (Figure 9 (a) and Figure 9 (b)).

**Figure 8 Regular lighting.**

**Figure 9: Regular Light in Frontal and Profile view**

**(a) Regular light + frontal view** **(b) Regular light + profile view**

The second setup, as shown in Figure 10, has a head light on the right side of the subject in order to create a strong and distinctive shadow on the face. The effect of shadow can be seen in Figure 11 (a). Only the left half of the face is clearly visible whereas the right half of the face has a shadow which makes it difficult for human eyes to see. In Figure 11 (b), the back of the head (especially those with long hairs) is illuminated with strong reflections whereas the facial parts are not as bright as the hairs.

**Figure 10: Shadow lighting.**

**Figure 11: Shadow Light in Frontal and Profile view**

**(a) Shadow + frontal view.** **(b) Shadow + profile view**

## 4.3 Subjects

The people who participated in the video collection phase are volunteers on YSU campus which include students, staff and faculty members. According to the policies of YSU IRB committee, the significance of the research and related implementation issues are explained to each participant. A video that clearly demonstrates the activities involved in the data collection process is presented and viewed by each subject. Each volunteer must sign a consent form approved by the IRB committee. The consent form summarizes the procedure, states the purpose of the research, and indicates that no private information will be collected about subject's name, age and sex.

## 4.4 Procedure

A subject is asked to sit on a chair facing the camera, in a relaxed position. Two steps are involved in each session:

- The subject turns slowly towards his/her left for 90 degrees to capture the smooth transition from frontal view to profile view. The process is repeated twice, one at the regular light and the next time at the strong shadow lighting.
- The subject is asked to turn back to face the camera. The subject can be asked to turn back slowly if the previous transition has not been captured. Then the head light is turned on to created shadow effect, steps 1 and 2 are repeated.

Once the video is captured it is important to ensure that the video is saved to the destination directory, before the subject leaves. In case the video was not saved properly the subject can be asked to pose for another session.

## 4.5 Frame Selection Criteria

To select ten frames of different pose angles, certain criteria are designed to facilitate the process:

- As described in Table 2. The ground truth of ten pose angles is established using ten frames that were acquired with measured angles (chair orientation).

- Chair orientation is done through ten ground angles from the frontal view (0 degree) to the profile view (90 degree) in 10 degree increasing order.

- This procedure ensures the quality of selected frames. Sample of three subjects of ground truth data are shown Table 3.

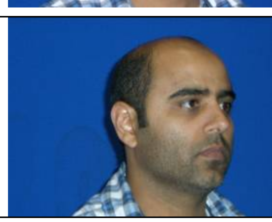
- In this thesis, ten frames were *manually* selected for each subject, in order to keep a high level of data consistency; frame selection was done by the same person.

## Table 2: Frame Selection Criteria

| Angle | Ears | Eyes | Nose | Cheeks | Hair above ears | eyebrows |
|---|---|---|---|---|---|---|
| **0°** | Both ears shown | Both eyes are shown | Directly facing | Both cheeks shown equally | Both sides shown | Both are shown |
| **10°** | One ear disappears and the other shows more details | Both eyes are shown | One side of the nose shown more than the other | One cheek is shown more details, the other | One side shows better detail | Both are shown |
| **20°** | One ear disappears and the other shows more details | Both eyes are shown | One side of the nose shown more than the other | One cheek is shown more details, the other | One side disappears | Both are shown |
| **30°** | Ear shown more details such as curves and some of the inside curves. | Both eyes are shown | One side of the nose disappears | One side of the cheeks almost disappears | Only one side shown | Both are shown |
| **40°** | Again more details shown | One eye partially disappears | More details shown for one side of the nose; such as area between nose and face | One side of the cheeks almost disappears | Only one side shown; Also some hair above and behind the ear is shown | Both are shown |
| **50°** | Ear shown better and hair behind it starts to show | One eye partially disappears ; still see eyelashes and eyelid | More details shown for one side of the nose; such as area between nose and face | One side of the cheeks disappears | Only one side shown; Also some hair above and behind the ear is shown | One eyebrow starts to disappear |
| **60°** | Ear shown better and hair behind it starts to show | One eye partially disappears ; still see eyelashes and eyelid | More details shown for one side of the nose; such as area between nose and face | One side of the cheeks disappears | Only one side shown; Also some hair above and behind the ear is shown | One eyebrow starts to disappear |
| **70°** | Ear shown almost facing you; hair behind ear shows | One eye almost disappears ; still see eyelashes | More details shown for one side of the nose; such as area between | One side of the cheeks disappears | Only one side shown; Also some hair above and behind the ear is shown | One eyebrow starts to disappear (almost 50% disappears) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | nose and face | | | |
| **80°** | Ear shown almost facing you; hair behind ear shows; back side of the head shows | One eye disappears ; | Side of the nose almost facing you | One side of the cheeks disappears | Only one side shown; Also some hair above and behind the ear is shown | One eyebrow starts to disappear (almost 70% disappears) |
| **90°** | Complete side shot of the face | One eye disappears ; side shot for the other | Side of the nose facing you | One side of the cheeks disappears | Only one side shown; Also some hair above and behind the ear is shown | One eyebrow disappears |

**Table 3: Ground Truth for Frame Selection.**

| Angle | Subject 1 | Subject 2 | Subject 3 |
|---|---|---|---|
| **0°** |  |  |  |
| **10°** |  |  |  |
| **20°** |  |  |  |
| **30°** |  |  |  |
| **40°** |  |  |  |

| | | | |
|---|---|---|---|
| **50°** |  |  |  |
| **60°** |  |  |  |
| **70°** |  |  |  |
| **80°** |  |  |  |
| **90°** |  |  |  |

## 4.6 Fusion on Image Level

The majority of biometric fusion was done on the score or rank level [13]. Only a few studies have used image (sensor) level fusion. For example, Chang et al [54] evaluated the performance of a multi-biometrics system by concatenating a face image and an ear image. Fusion on the image level has the advantage that information in raw data is preserved, and is suited for the multi-sample approach as long as the number of samples per subject is reasonably small. This study use image level fusion to integrate as many as ten frames per subject. Image fusion represented here in three steps:

- Ten frames were chosen for each subject with following rotation degrees: 0, 10, 20, 30, 40, 50, 60, 70, 80 and 90.
- Each frame was normalized using the coordinates of two facial markers:
    - If a face rotated by 0, 10, 20 or 30 degrees, the centers of eyes were used.
    - If a face rotated by 40 degrees, the left corners of eyes were used.
    - If a face rotated by 50, 60, 70, 80 or 90 degrees, the top of the nose and the middle point between the center of ear and the top of nose were used (Figure 12).
- The normalized images were then aligned vertically to create a fused image. Figure 13 shows samples of fused images.

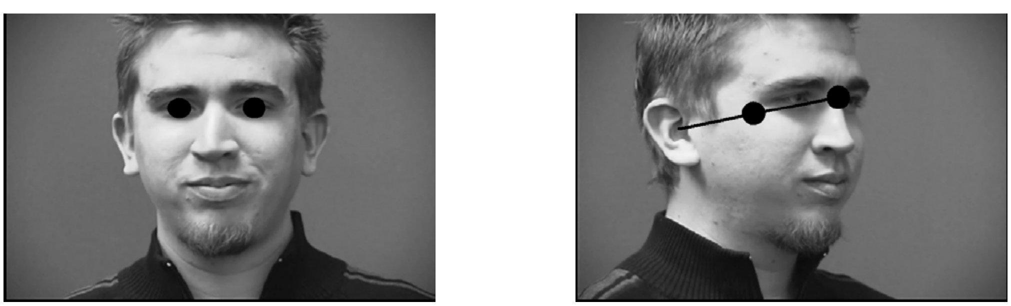


**Figure 12: Selection of two facial markers for image normalization**

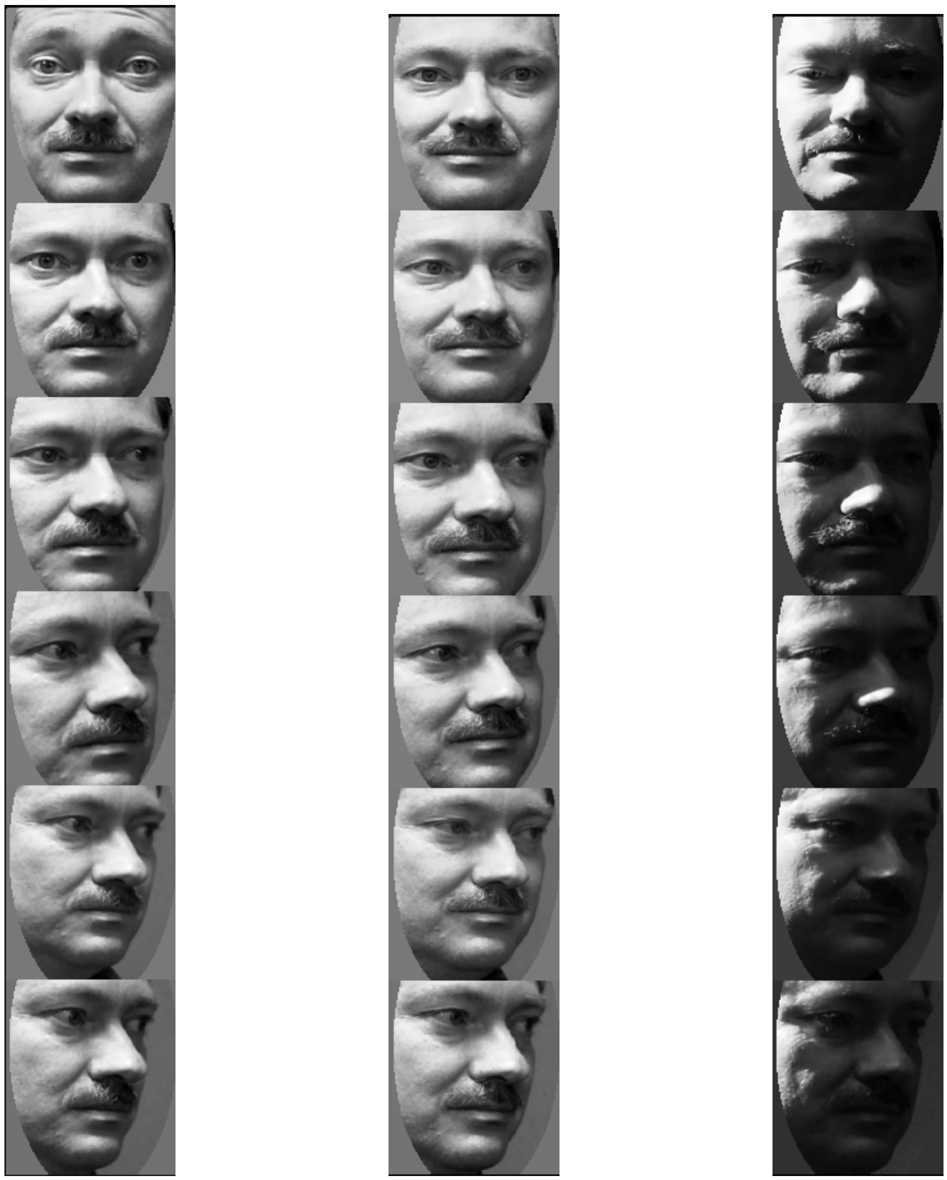


**Figure 13 : Samples of fused images.**

# 5 Experiment Results

## 5.1 First Condition Test:

The purpose of the first condition Test-1 is to examine the performance of multi-sample fusion using different numbers of frames – Approximate number of 200 frames– per subject under the regular light. Started with one frame per subject (0 degree), and then fused it with the next frame (10 degrees), until integrated all ten frames. For example, a 10-frame fusion concatenates frames in an increasing order of rotation degrees: 0, 10, 20, 30, 40, 50, 60, 70, 80 and 90. Using the data set, an increasing improvement in the recognition rate has been observed. About 10 percent performance increase was observed using the seven frame fusion. The same level of performance improvement is expected with the ten-frame fusion data set.

## 5.2 Second Condition Test:

One way to assess the robustness and effectiveness of a recognition method is to apply it to images of severe illumination changes. So, Test-2 that has the same data set as Test-1, except that its probe set consists of images from the second collection with shadows on the faces. As we can see in the sample images, the shadows almost black out half of the faces, the testing condition is more challenging than the first test, about 20 percent performance increase was observed using the seven frame fusion. The same level of performance improvement is expected with the ten-frame fusion data set.

## 5.3 Discussion

It should be stressed that not all fusions will result in positive outcomes. The overall trend is that the performance improves as the number of frames per subject increases, but their relationship is not strictly monotonic. The recognition rate actually dropped in certain cases. There are several possible explanations:

- Some subjects blinked their eyes because of the headlight.
- Sometimes subjects rotated relatively fast leading to blurred images.
- There might be a more fundamental issue of multi-sample fusion that is related to the interplay of sample sets and their combined effect

.

As suggested in [13], if two sets of samples are positively correlated, the noise in the samples could negate any performance gain from their fusion. In the 2-frame fusion case, the performance drop may be explained by the lack of variations between the 0-degree frames and the 10-degree frames. In other words, the faces in those two sets are so similar that their fusion provides little complementary benefit.

## 5.4 Summary

To achieve a significant increase in face recognition rate under challenging conditions necessitates the development of new techniques such as 3D scans, high resolution images, multi-sample and multi-modal methods [53]. Using videos that capture the continuous pose and illumination changes of moving faces to provide implicit 3D information is one possible solution. This thesis presents some preliminary experimental results of recognizing faces that rotated up to 90 degrees using multi-frame fusion. Based on the two tests with videos taken under two lighting conditions, several observations can be made:

- Recognition rate shows large improvements in both tests, about 10% under the regular lighting condition and about 20% under the strong shadow condition. This performance increase can be, to a large degree,

- A linear function seems inadequate to describe the relationship between the recognition rate and the number of frames used in fusion. It is likely that finding an optimal number of frames to achieve the maximum performance increase will be task-dependent. I will conduct more experiments using 20 to 90 frames per subject with rotation intervals of 1 to 5 degrees. I will also investigate this issue in the framework of scale-space aspect graph so as to find the minimum number of frames that can provide sufficient 3D information for face recognition [14].

Since the motivation here is to utilize implicit 3D information in videos via multi-frame fusion, it would be interesting to compare its performance with those of using explicit 3D data such as range images, so that its efficacy can be benchmarked. This requires a data set that includes both range images and rotating head videos of the same subject. This data set will also allow us to have a better understanding of the interaction of individual frames, which is of fundamental significant to the success of multi-biometric systems.

My suggestion is to collect those data and double the size of our database to 200 subjects. One related issue is whether image level fusion and score level fusion would yield the same performance, because score level fusion is more computationally attractive if a large number of frames are needed for fusion.

Finally, I would like to emphasize that, although a complete video sequence of 90 degrees head rotation is rare in real situations, this kind of data is an ideal test bed that allows us to examine various factors that influence the performance of the multi-sample method. Moreover, in certain realistic scenarios such as video surveillance, even a short video segment that captures partial head rotation could be valuable for recognition.

# Bibliography


[1] K. W. Bowyer, K. Chang, and P. J. Flynn. "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," Computer Vision and Image Understanding, vol. 101, no. 1, pp. 1-15, 2006.

[2] I. A. Kakadiaris, G. Passalis, G. Toderick, M. N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis. "Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach," IEEE Trans. on Pattern Analysis and Machine Intelligence., vol. 29, no. 4, pp. 640-649, 2007.

[3] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399-458, 2003.

[4] K. W. Bowyer, K. Chang, P. J. Flynn, and X. Chen, "Face recognition using 2-D, 3-D and Infrared: Is multimodal better than multisample?" Proceed of the IEEE, vol. 94, no. 11, pp. 2000-2012, 2006.

[5] D. Thomas, K. W. Bowyer, and P. J. Flynn, "Multi-frame approaches to improve face recognition," IEEE Workshop on Motion and Video Computing, pp. 19-19, Austin, TX, 2007.

[6] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," Proceedings of the IEEE, vol. 83, no. 5, pp. 705-740, 1995.

[7] V. Blanz, and T. Vetter, "Face recognition based on fitting a 3D morphable model," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 9, pp. 1063-1074, 2003.

[8] D. DeCarlo, and D. Metaxas, "Optical flow constraints on deformable models with applications to face tracking," Inter. Jour. of Computer Vision, vol. 38, no. 2, pp. 99-127, 2000.

[9] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Three-dimensional face recognition," International Journal of Computer Vision, vol. 64, no. 1, pp. 5-30, 2005.

[10] K. Lee, J. Ho, M. Yang, and D. Kriegman, "Video-based face recognition using probabilistic appearance manifolds," Proc. of IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 313-320, 2003.

[11] C. Beumier, and M. Acheroy, "Face verification from 3D and grey-level clues", Pattern Recognition Letters, vol. 22, pp. 1321-1329, 2001.

[12] M. Husken, M. Brauckmann, S. Gehlen, K. Okada, C. V. Malsburg, “Evaluation of implicit 3D modeling for pose-invariant face recognition,” Proceedings of SPIE, vol. 5404, pp. 328-338, 2004.

[13] A. A. Ross, K Nandakumar, and A. K. Jain, Handbook of Multibiometrics, Springer, 2006.

[14] D. W. Eggert, K. W. Bowyer, C. R. Dyer, H. I. Christensen, D. B. Goldgof, “The scale space aspect graph”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 1, pp. 1114-1130, 1993.

[15] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. June 2003. Handbook of Fingerprint recognition. Springer Verlag, New York, USA.

[16] A. K. Jain, A. Ross, and S. Prabhakar. January 2004. An Introduction to Biometric Recognition. IEEE Transactions on circuits and systems for video technology, 14(1).

[17] D. M. Blackburn. March 2004. Biometrics 101, version 3.1. Federal Bureau of Investigation.

[18] A. K. Jain, S. Pankanti, and S. Prabhakar. 2003. Biometric Recognition: Security and Privacy Concerns. IEEE Security & Privacy, The IEEE Computer Society.

[19] P. J. Phillips, P. Grother, R. J. Michaels, D. Blackburn, E. Tabassi, and M. Bone. March 2003. Face Recognition Vendor Test 2002: Evaluation Report. http://www.frvt.org.

[20] A. J. Mansfield and J. L. Wayman. August 2002. Best Practices in Testing and Reporting Performance of Biometric Devices, Version 2.01. Centre for Mathematics and Scientific Computing, National Physical Laboratory, Queens Road, Teddington, Middlesex, TW11 OLW.

[21] T. Mansfield, G. Kelly, D. Chandler, and J. Kane. March 19, 2001. Biometric Product Testing Final Report Issue 1.0. Centre for Mathematics and Scientific Computing, National Physical Laboratory, Queens Road, Teddington, Middlesex, TW11 OLW.

[22] A. Ross and A. Jain. 2003. Information fusion in biometrics. Pattern Recognition Letters, 24, pages 2115–2125.

[23] L. Hong and A. K. Jain. December 1998. Integrating Faces and Fingerprints for Personal Identification. IEEE transactions on pattern analysis and machine intelligence, 20(12), pages 1295–1307.

[24] 2002. Facial Scan Technology: How it works. Technical Report, International Biometrics Group.

[25] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. December 2003. Face Recognition: A Literature Survey. ACM Computing Surveys, 35(4).

[26] S. G. Kong, J. Heo, B. R. Abidi, J. Paik, and M. A. Abidi. 2005. Recent advances in visual and infrared face recognition: a review. Computer Vision and Image Understanding, volume 97, page 1.

[27] S.Z. Li and J. Lu. 2005. Face detection, alignment, and recognition. In G. Medioni and S.B Kang, editors, Emerging Topics in Computer Vision, chapter 9. Prentice Hall PTR.

[28] S.Z. Li and A. Jain. Handbook of Face Recognition. Springer, New York, 2005.

[29] M. A. Turk and A. P. Pentland. 1991. Face recognition Using Eigenfaces. Proc. IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii.

[30] J. Chirillo and S. Blaul. June 2003. Implementing biometric Security. Wiley Publishing, Inc.

[31] J. D. Woodward Jr., N. M. Orlans, and P. T. Higgins. 2003. Biometrics. McGraw-Hill/Osborne, New York.

[32] J. Zhang, Y. Yan, and M. Lades. September 1997. Face Recognition: Eigenface, Elastic Matching, and Neural Nets. Proceedings of the IEEE, 85(9), pages 1423– 1435.

[33] L. Wiskott, J. M. Fellous, N. Kruger, and C. von der Malsburg. July 1997. Face Recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), pages 775–779.

[34] D. S. Bolme. Elastic Bunch Graph Matching. MSc thesis, Master of Science, Colorado State University, Fort Collins, Colorado, May 2003.

[35] J. Wayman, A. Jain, D. Maltoni, and D. Maio. 2005. Biometric Systems: Technology, Design and Performance Evaluation. Springer-Verlag London Limited.

[36] J. Phillips. The Face Recognition Grand Challenge (FRGC). http://www.frvt.com/FRGC/.

[37] Aurora Computer Services Ltd. http://www.facerec.com/.

[38] G. Pan, Y. Wu, Z. Wu, and W. Liu. July 2003. 3D face recognition by profile and surface matching. Neural Networks, 2003, IEEE, volume 3, pages 2169–2174.

[39] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. October 2000. The FERET evaluating methodology for face recognition algorithm. IEEE Transactions on pattern analysis and machine intelligence, 22(10).

[40] M. Krechel and N. Tekampe. Common Criteria Protection Profile For Biometric Verification Mechanisms, Version 1.0. Technical report, TUV Informationstechnik GmbH, Germany, October 2004.

[41] R. Beveridge, D. Bolme, M. Teixeira, and B. Draper. May 1, 2003. The CSU Face Identification Evaluation System User's Guide: Version 5.0. Computer Science Department Colorado State University.

[42] February 2000. Best practices in testing and reporting performance of biometric devices, Issue 1. Report for CESG and Biometrics Working Group.

[43] P. J. Phillips, A. Martin, C.L. Wilson, and M. Przybocki. February 2000. An Introduction to Evaluating Biometric Systems. National Institute of Standards and Technology.

[44] V. Bruce and A. Young. 1998. In the eye of the beholder. The science of face perception. Oxford University Press Inc., New York.

[45] V. Bruce, P. J. B. Hancock, and A. M. Burton. 1998. Comparison between Human and Computer Recognition of Faces. ICAFGR, Nara, Japan, (408-413).

[46] Y. Adini, Y. Moses, and S. Ullman. July 1997. Face Recognition: The Problem of Compensating for Changes in Illumination Direction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), pages 721–732.

[47] R. Cross, J. Shi, and J. F. Cohn. December 2001. Quo vadis Face Recognition? In Third Workshop on Empirical Evaluation Methods in Computer Vision, Kauai, Hawaii, pages 119–132.

[48] J. C. Brigham. 1986. The influence of race on face recognition. Aspects of face processing, Nijhoff, pages 170–177.

[49] P. Chiroro and T. Valentine. 1995. An Investigation of the Contact Hypothesis of the Own-race Bias in Face Recognition. The Quarterly Journal of Experimental Psychology, 48A(4), pages 879–894.

[50] D. T. Levin. 2000. Race as a Visual Feature: Using Visual Search and Perceptual Discrimination Tasks to Understand Face Categories and the Cross-Race Recognition Deficit. Journal of Experimental Psychology: General, 129(4), pages 559–574.

[51] N. Furl, P. J. Phillips, and A. J. O'Toole. 2002. Face recognition algorithms and the other-race effect: computational mechanisms for a developmental contact hypothesis. Cognitive Science, volume 26, pages 797–815.

[52] J. Sadr, I. Jarudi, and P. Sinha. 2003. The role of eyebrows in face recognition.
Perception, volume 32, pages 285–293.

[53] P. J Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Margues, J. Min, and W. Worek, "Overview of the face recognition grand challenge," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR2005), pp. 947-954, Washington DC, 2005.

[54] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and combination of ear and face images in appearance-based biometrics," IEEE Trans. on Pattern Anal. and Machine Intell., vol. 25, no. 9, pp. 1160-1165, 2003.