



The Hong Kong Polytechnic University
Department of Computing
COMP55422 Multimedia Computing, Systems and Applications
Class of 2011-2012

Group Assignment

Topic: Face Recognition

Date of Submission: 16 April 2012

Group 6:	NG Check Kiu Jay	10629203G
	CHUNG Hau Ying, Janice	11521450G
	LAM Wai Chung	10634631G
	WONG Kwan Shing	10604036G
	QING Pei	11500811G

Abstract

Numbers of general face recognition issues has been a topic which addressed by analyzing the facial feature for many years. It sparked an explosion of interest in the topic of face recognition.

The ultimate challenge of face recognition is to effectively identify human faces. In this project, two methods are adapted to enhance the ability of image reorganization. The characteristics of these techniques might be suitable for many applications. These heuristics might put an effort to predict human face variants behavior under different conditions.

Index	Table of Content	Page
1	Introduction	4
2	Usage of Existing Face Recognition Technology	5
3	Search and planning of face recognition	5
4	System Testing / Result	12
5	Conclusion	14
6	Reference	15

1. Introduction

1.1 Background

Since the past decade, face recognition gained a significant attention and become an active research area of computer vision. Face Recognition is a computer application for automatically identifying or verifying a person by using digital image or video source. This automatic capability can be used in law enforcement, security, biometric authentication and surveillance, etc, whereas it can perform effortlessly and routinely in our daily life. According to Zhao & Chellappa (2002) [5], this is a useful tool to secure our asset and protect our privacy without losing the identity.

According to Li & Jain (2011) [2], Takeo Kanade developed the first automated face recognition system for his Ph.D thesis work in 1973. He demonstrated a simple neural net which can perform face recognition for aligned and normalized face images. By approximating the eigenvectors of face image's autocorrelation matrix, the face description is computed. Such eigenvectors are now known as "Eigenfaces".

In 1989, Kirby and Sirovich [5] worked on low dimensional face representation by using the Karhunen-Loeve transform or Principal Component Analysis (PCA). They planned to make it in an easier way to directly calculate the Eigenfaces.

Regarding to the research of Turk & Pentland (1991), Turk & Pentland reinvigorated face recognition technology. At that time, when using eigenfaces, there was still residual error by detecting the face in cluttered natural imagery and determining the precise location and sale of faces in an image. Therefore, they derived a method to detect and localize faces with eigenface recognition. It could achieve reliable, real-time recognition of faces in a minimally constrained environment. In addition, this is relatively simple, and in real-time pattern.

Fisher face method also created interest to develop face recognition technology. According to the studies of Belhumeur (1997), they applied Linear Discriminant Analysis (LDA) adter a PCA setup. This suggestion can help to achieve higher accuracy. Effective facial features can be achieved by using local filters, for example Gabor jets. This is the milestone of real time face detection.

2. Usage of Existing Face Recognition Technology

Face recognition technology is commercially available as there is a strong demand on using Face Recognition technology as it can increase security in a convenient way. Suggested by Philips, Rauss & Der (1996), there are numbers of security and law enforcement tasks using face recognition technology. In monitoring areas (including airports, border crossings and secure manufacturing facilities and staff only areas), key and identity card are not necessary and the face recognize system only need to scan the users' face. The technology also widely used in verifying identity. For example, users can search the photo record in a database for fraud detection.

3. Search and planning of face recognition

3.1 Recognition

While each input images are different, they are not completely random. Some patterns exist in all facial images, such as the presence of eyes, nose and mouth and the relative position of these organs. By applying Principal Component Analysis (PCA), we may extract these principal components (or eigenfaces in the facial recognition domain) out from the input image data.

Each eigenfaces represents a certain feature of the face. The original face image is a linear expression of some eigenfaces with various weights. Therefore, a facial image can be represented by a weight vector.

By comparing the weight vector of an input image and those of images in the training dataset, it can be determined which training sample is closest to the input in terms of eigenfaces features.

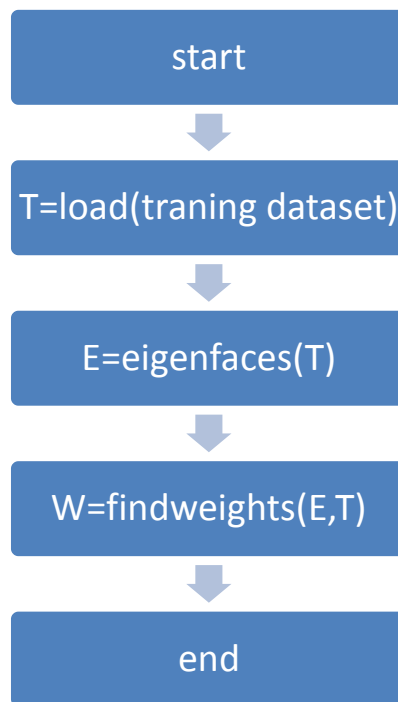


Figure 1 Training

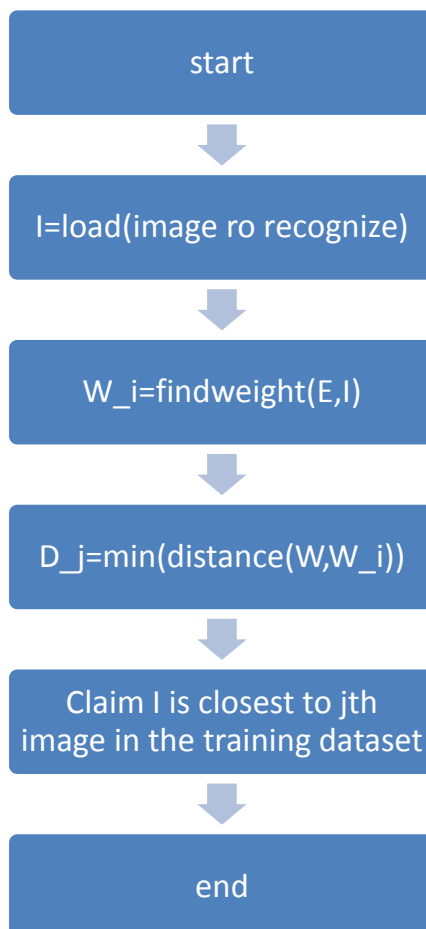


Figure 2 Recognition

To reconstruct the original image, we need all the eigenfaces and weights for each. In order to just recognize a face, some eigenfaces are more important than others. Considering only the most important eigenfaces may reduce the computation and offer a shorter recognition time as a result. In our experiment, we chose the top 20 eigenfaces.

3.2 Database Used

We used the database of faces from AT&T¹. The database consists of face images of 40 distinct subjects, 10 for each. The following aspects are different for different samples of a single subject.

- Time of photo capture
- Lighting condition
- Facial expression
- Accessories (with/without glasses)

All the subjects are facing the camera while some subjects have some side movements.

The file format:

- PGM format
- 92*112 pixels
- 256 grey levels per pixel

3.3 Illumination

Illumination processing is a necessary step for face recognition to be useful. In this section, we will focus on the influence on face recognition caused by illumination.

Illumination variation can be removed: the input face images can be normalized to some state where comparisons are more reliable [9]. We are going to introduce two popular algorithms to do the normalization on the image, they are based on the Retinex theory. The SSR (Single-Scale Retinex) and the MSR (Multi-Scale Retinex) could probably solve the problem of illumination of the images. They will be discussed in details in this section. In fact there are many different kind of algorithms doing similar work such as “adaptive single scale retinex”, “DCT based normalization”, “wavelet based normalization”, “anisotropic diffusion based normalization” etc. We are going to propose the most common algorithms in this report.

¹ <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

3.3.1 SSR (Single-Scale Retinex)

The single scale retinex (SSR) algorithm was proposed by Jobson et al. It is normalization techniques based on retinex theory. Each output value of the function is depended by the corresponding input value and its neighbourhood. The surround function is a Gaussian function. The equation of the single-scale Retinex (SSR) is given by

$$R(x_1, x_2) = \alpha(\log(I(x_1, x_2)) - \log(I(x_1, x_2) * F(x_1, x_2))) - \beta$$

where I is the input image, R is the Retinex output. F is a Gaussian filter (surround or kernel) defined by <http://dragon.larc.nasa.gov/background/pubabs/papers/gspx1.pdf>

$$F(x_1, x_2) = \kappa \exp[-(x_1^2 + x_2^2)/\sigma^2]$$

Below is the example of applying SSR to the images.

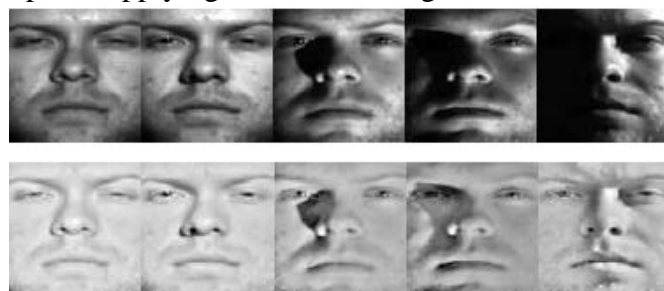


Figure - 3

3.3.2 MSR (Multi-Scale Retinex)

The multi-scale retinex (MSR) algorithm is an extension of the single scale retinex algorithm again proposed by Jobson et al. The mathematical form of the multi-scale Retinex (MSR) is given by where $R_n(x,y)$ are different scale SSR (obtained with different Gaussian functions) and W_n is the weight of each SSR. Usually, these weights are taken to be equal. [7]

$$R_M(x, y) = \sum_{n=1}^N w_n R_n(x, y)$$

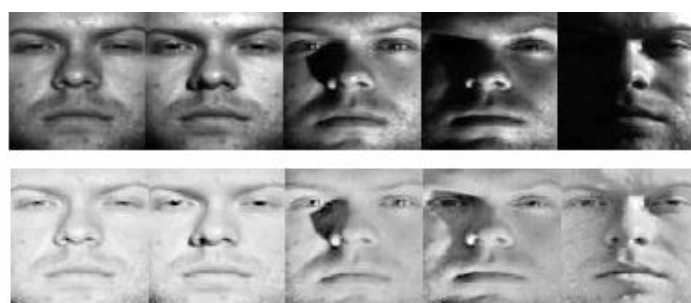


Figure 4 - Example of applying MSR to the images.

3.3.3 Finalize

From the above findings, we could say that MSR should be better but it is more complicated. In our project, we chose to apply the SSR instead of MSR because it is easier to implement and we believed that the processing time would be faster.

The function of SSR in our project would receive an INPUT of a grey-scale image and the OUTPUT should be a grey-scale image processed with the SSR algorithm.

3.4 Head Pose

The proportion of human head varies from person to person and growth change may affect the facial outlook too. In general, it includes general size, shape and position of facial features in front face view. In this section, we mainly focus to head pose variations detection and attempt to correct the face angle if high confidence level can be obtained by using our algorithm.

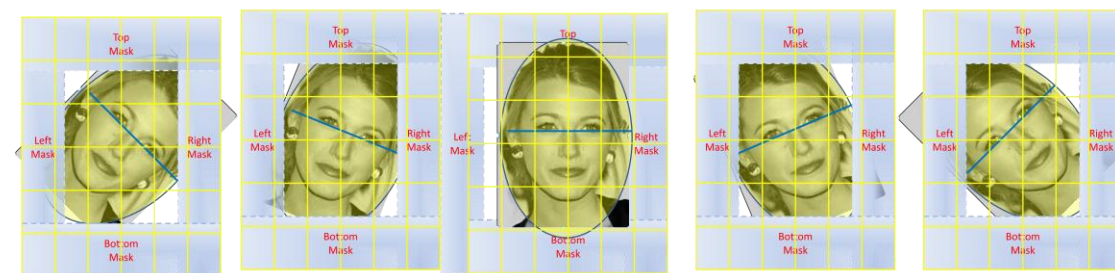


Figure5 – Various head pose conditions to be tracked

In our project, we adopt most common facial features and assigned sample image set of faces² taken by Olivetti Research Laboratory in Cambridge, UK between April 1992 and April 1994. The key workflows are 1)Skin extraction, 2)noise removal, 3)inverse skin for 4)face components detection and 5)eye orientation angle determination.

² The ORL face database - This directory contains a set of faces taken between April 1992 and April 1994 at the Olivetti Research Laboratory in Cambridge, UK. There are 10 different images of 40 distinct subjects. For some of the subjects, the images were taken at different times, varying lighting slightly, facial expressions (open/closed eyes, smiling/non-smiling) and facial details glasses/no-glasses). All the images are taken against a dark homogeneous background and the subjects are in up-right, frontal position (with tolerance for some side movement) The files are in PGM format and can be conveniently viewed using the 'xv' program. The size of each image is 92x112, 8-bit grey levels. The images are organised in 40 directories (one for each subject) named as: sX where X indicates the subject number (between 1 and 40). In each directory there are 10 different images of the selected subject named as: Y.pgm where Y indicates which image for the specific subject (between 1 and 10).

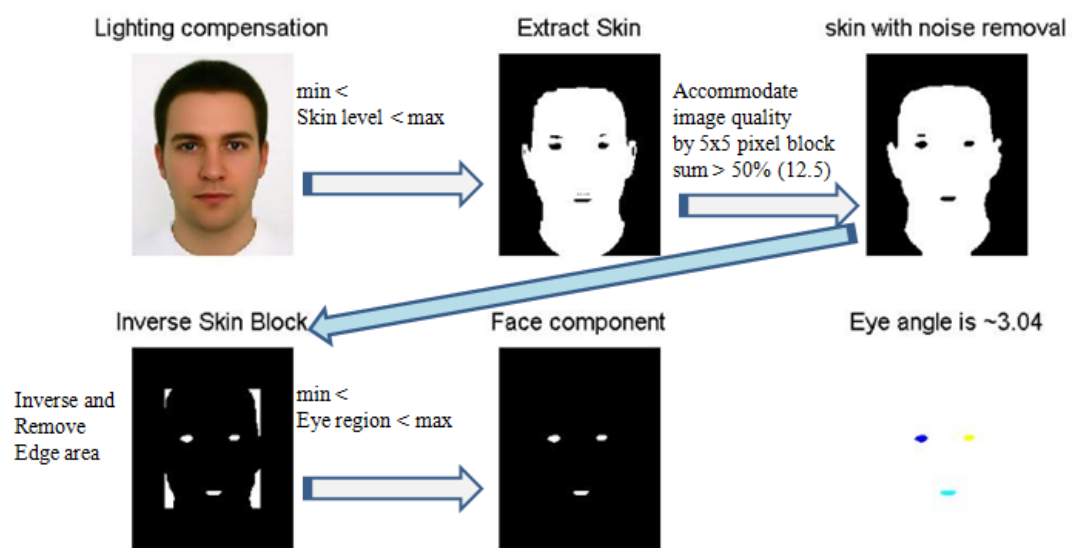


Figure 6 - Workflow of headpose detection

The following outline depicts the human head features and the characteristics are being derived.

- ✓ The width of human head is approximately two thirds of its height.
- ✓ The front view of human face can be divided into three sections, namely the top, middle and the bottom parts. Each section is approximately one third of the face view section.

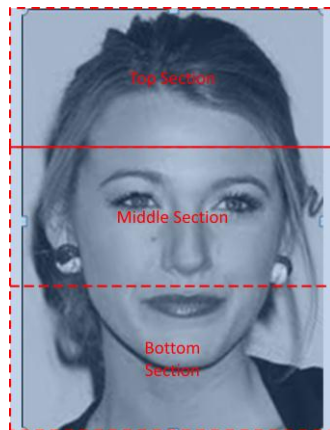


Figure 7

- Top section : cover most of the hair feature (head down to the hairline)
- Middle section: cover eyes, nose and ears. (hairline down to eyes, usually same level with the ears and nose a the lower part aligning the ear lobes.
- Bottom section : include mouth (bottom up from chin to the mouth position)

The proportions will varies if the face view turns, head from above or below. We assume the eyes as the focal point of head image, and the eyebrows are situated approximately from 1/3 down the head without head turn above or below. These distances work out at approximately two third of the width of the face. Also the eye proportion is between 4% and 12.5% image width and between 1.6% and 6.6% image height..

$\text{Eyemin} = \text{round}(\text{ImageWidth}/25 * \text{ImageHeight}/60);$

$\text{Eyemax} = \text{round}(\text{ImageWidth}/8 * \text{ImageHeight}/15);$

As you can see from the illustration below, the two eyes focus areas bisects the head of the vertical line.

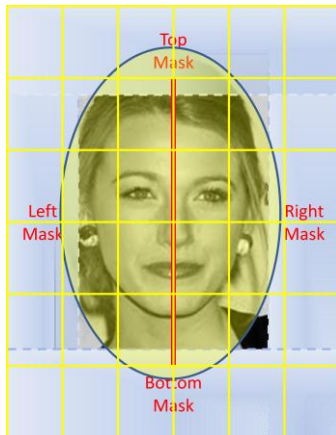


Figure 8 -

Head pose with vertical line

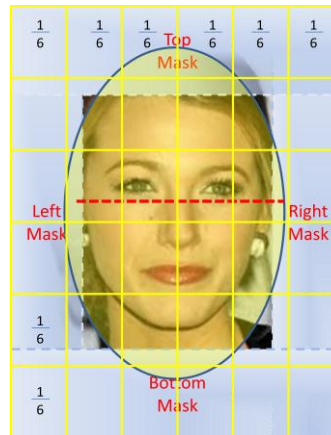


Figure 9 -

Image layout with grid lines

In many cases, eye is the only facial feature which is easily and effectively identifiable from facial skin. We detect eyes depending basically on the intensity level variance between Skin level sampled. The Skin intensity level sampling are taken by introducing three circular line (CL[1:3,1:12]) tracks starting from centre of the image, each with approximately one third of the image width.

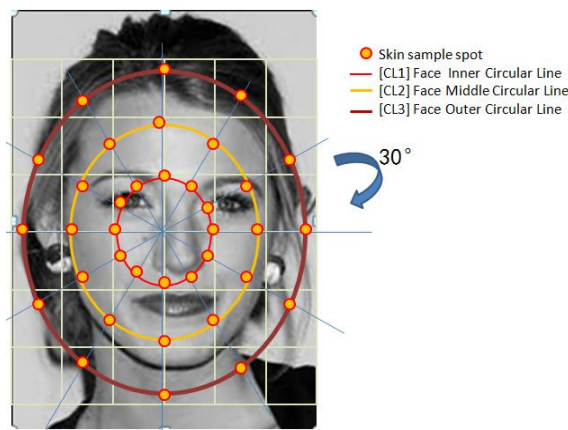


Figure10 - Circular-Line tracking of skin level

The skin intensity level samples are taken in 30 degree increment clockwise starting from Y axis and our observation tends to use the mean value taken from the two inner circular lines for our setup. By making an assumption of intensity variance of face skin level with a variations upto 50%.

$$\text{skinlevelmin} = \min(\text{mean}(\text{CL1}(1,:)), \text{mean}(\text{CL1}(2,:))) * 0.5;$$

$$\text{skinlevelmax} = \max(\text{mean}(\text{CL1}(1,:)), \text{mean}(\text{CL1}(2,:))) * 1.5;$$

However, we do not consider complex conditions such as different head view projections as they will fail under our imaging conditions.

```
[SkinIndexRow,SkinIndexCol] =find(skinlevelmin<I & I<skinlevelmax);
for i=1:length(SkinIndexRow)
```

S(SkinIndexRow(i),SkinIndexCol(i))=1;

end

After the extraction of facial skin region, a simple noise removal routine by examining a 5x5 pixels block each with 50% intensity level '1'. Then by using Matlab function "regionprops(CC,'Area');" and "L = labelmatrix(CC);", we attempt to keep area with our pre-determined eye area ranges, and also apply a peripheral mask so as to focus our eye components inside face Skin area.

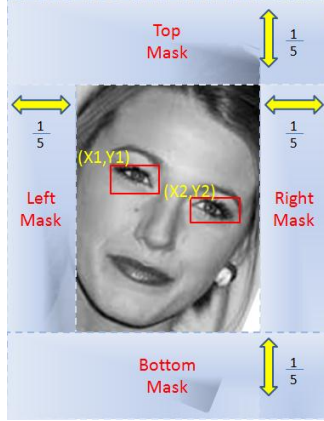


Figure 11 -

Region mask layout

TopArea => between 0 and 20% Height

BottomArea=> between 80-100% Height

LeftArea => between 0-20% Width

RightArea => between 80-100% Width

By utilizing the Matlab function regionprops (Measure properties of image regions) with 'BoundingBox' option, the coordinate of the two eye components, if it can be detected, and the associated head pose angle turn can be determined.

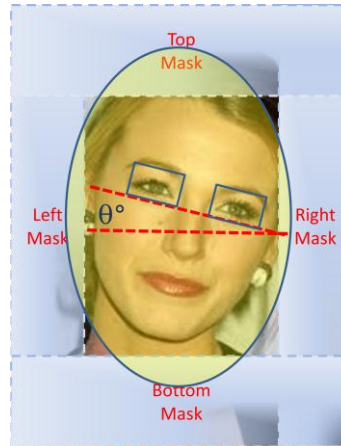


Figure 12 – head pose turn angle determination

Finally, we attempt to rotate the head angle and extract the two eye regions, its similarity of both sides, for further comparison by checking the mean difference of the intensity level of each eye regions. We suggest applying the head pose correction if the confidence level is greater than 90%.

Because of noise in the images due to lighting condition, facial expression or wearing, the successful detection rate (i.e. images with correctly detected eye/nose region relative to the whole set of facial database) is not so high, but it is reasonable.

The efficiency of the method is evaluated with ORL databases (40 sets variety of

images). It has been shown by experimental results that 18 out of 400 images can be detected by the proposed method. Also 7 out of 11 additional images can be detected and we consider the outcomes are acceptable.

The followings are the seven samples by implementing *rgb2gray transformation* from colour images.

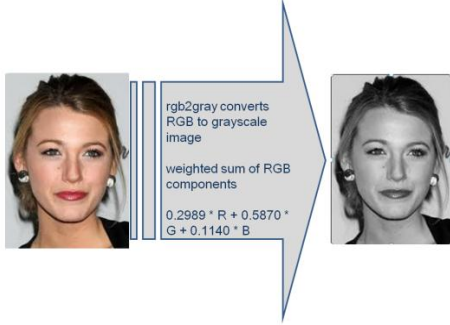






















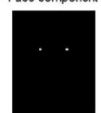



















 <p>rgb2gray converts RGB to grayscale image weighted sum of RGB components $0.2989 * R + 0.5870 * G + 0.1140 * B$</p> <p>Rgb2gray transformation</p>	<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-2.94</p>  <p>Sample #1</p>
<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-24.5</p>  <p>Sample #2</p>	<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-3.04</p>  <p>Sample #3</p>
<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-0</p>  <p>Sample #4</p>	<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-39.8</p>  <p>Sample #5</p>
<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-40.4</p>  <p>Sample #6</p>	<p>Lighting compensation</p>  <p>Extract Skin</p>  <p>skin with noise removal</p>  <p>Inverse Skin Block</p>  <p>Face component</p>  <p>Eye angle is ~-31.8</p>  <p>Sample #7</p>

Table 13

4. System Testing / Result

The Facial recognition method was implemented in MATLAB and examines with set of different images by using ORL image set. To proceed testing with MATLAB, the folder structure shall be retained by extracting the programming package onto local folder, then open “face_cognition.m” matlab to start the testing.

Our project will complete the testing in three phases, 1) Ready Images sample baseline and training set building for program image library, 2) prepare test image for program test and 3) search for the best match face image from the program image library.

4.1 Test result summary :

Result-Set(A) - Image matching Tested without adjustment

Result	Image matching Tested without adjustment	Ratio
Fail	28	7%
Success	372	93%
	400	

Result-Set(B) – Enable Headpose detection and correction module

Result	Image matching Tested with Headpose detection and correction	Ratio
Fail	27	6.75%
Success	373	93.25%
	400	

3 x Image set ID matching improved :

- 257, 278 & 346

2 x Image set ID mis-matching against Result-Set(A) :

- 215 & 242

Result-Set(C) – Enable Illumination adjustment module

Result	Image matching Tested with Illumination adjustment	Ratio
Fail	36	9%
Success	364	91%
	400	

10 x Image set ID matching improved :

- 62, 212, 259, 326, 352, 359, 360, 376, 378 & 379

18 x Image set ID mis-matching against Result-Set(A) :

- 7,10,50,139,148,190,216,235,252,271,280,298,321,341,353,358,374,394

Result-Set(D) – Enable both Headpose detection and correction, and also Illumination adjustment modules

Result	Image matching Tested with both features	Ratio
Fail	36	9%
Success	364	91%
	400	

10 x Image set ID matching improved :

- 62, 212, 259, 326, 352, 359, 360, 376, 378 & 379

18 x Image set ID mis-matching against Result-Set(A) :

- 7,10,50,139,148,190,216,235,252,271,280,298,321,341,353,358,374,394

Above table shown the details success rate in different situation in 400 images, the comparison including “Image matching Tested without adjustment”, “Image matching Tested with Headpose detection and correction”, “Image matching Tested with Illumination adjustments” and “Image matching Tested with both features”.

4.2 Test result analysis

From the result, the success rate of processing with “headpose” feature is a little bit higher than without any features. This is all depends on the training dataset as we have rotated some of the images to do the headpose testing. For example, ImageID 257 is one of them; it has been successfully recognized with headpose feature. It is meaning that the more images has been rotated, the more success rate of “headpose”.

5. Conclusion

As our initial goal of this research is improving the effectiveness of face recognition technology, it is recommended to focus on improving illumination and detects the head pose variations. As view the previous history of face recognition development, we suggested different algorithms to predict human face and improve the accuracy of face recognition while conditions change. From the testing result, it has shown that adjusting the rotation angle and illumination can improve the success rate.by implementing our suggested face recognition algorithm in MATLAB with numbers of images.

6. Reference

- [1.] Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces (1997): recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(7), 711-720
- [2.] Li, S. Z. & Jain, A. K. (2011). *Hand Book of Face Recognition*. London: Springer-Verlag London Limited.
- [3.] Philips, P. J., Rauss, P. J. & Deng, S. Z. (October 1996). FERET (Face Recognition Technology) Recognition Algorithm Development and Test Results. *U.S. Army Research Laboratory*, 73.
- [4.] Turk, M.A., Pentland, A.P. (1991): Eigenfaces for Recognition. *J. Cogn. Neurosci.* 3(1), 71-86
- [5.] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève procedure for the characterization of human faces," *IEEE Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108, 1990.
- [6.] Zhao, W., and Chellappa, R. (2002), "Image-based Face Recognition: Issues and Methods", in B. Javidi (ed.), *Image Recognition and Classification*, Marcel Dekker eBook ISBN: 978-0-203-91096-2
- [7.] Improving the Global Impression of Brightness of the Multi Scale Retinex Algorithm for Wide Dynamic Range Pictures Marius Herscovitz, Evgeny Artyomov, Orly Yadid-Pecht The VLSI Systems Center, Ben-Gurion University.
http://www.ee.bgu.ac.il/~Orly_lab/Publications/Marius_paper1_SPIE.pdf
- [8.] A Study of Improved Enhancement Method for Aerial Imagery by Yu Su, Department of CS&IE, Yuanpei University, 16-Dec-2009.
<http://203.72.2.115/Ejournal/AC01001602.pdf>
- [9.] Illumination Processing in Face Recognition Yongping Li, Chao Wang and Xinyu Ao Shanghai Institute of Applied Physics, Chinese Academy of Sciences China.
http://cdn.intechopen.com/pdfs/10200/InTech-Illumination_processing_in_face_recognition.pdf
- [10.] The INface toolbox v2.0 for illumination invariant face recognition by Vitomir Struc, 29 Jan 2010
<http://www.mathworks.com/matlabcentral/fileexchange/26523>
- [11.] Efficiency of Face recognition using Sparse representation by Neeraj Pai 29045615 UFL.
<http://www.mathworks.com/matlabcentral/fileexchange/authors/213487>