Honours Year Project Report

# Altering faces using the Multimodal Discriminant Analysis

By

G Sri Shaila

Department of Computer Science

School of Computing

National University of Singapore

2013/2014

Honours Year Project Report

# Altering faces using the Multimodal Discriminant Analysis

By

G Sri Shaila

Department of Computer Science

School of Computing

National University of Singapore

2013/2014

Project No: H123456
Advisor: Dr. Terence Sim
Deliverables:
Report: 1 Volume
Source Code: 1 DVD

## Abstract

In today's world, we are surrounded by electronic devices that capture our images almost all the time.There has been rising concern over how these images are stored and what happens to these footages after many years. An algorithm which can modify some attributes of the human face while retaining others can be used to preserve the privacy of the individuals whose images are captured by the camera without affecting the usability of the images in classification tasks.

Multimodal Discriminant Analysis is a method which can be used to decompose the variations in a dataset into independent factors that can be controlled separately. This method can be used to alter human faces by changing some attributes while preserving others. In this project, Multimodal Discriminant Analysis has been used to alter the gender, race, age and the identity of human faces. The performance of this method has been evaluated by the use of a survey.

The implementation of Multimodal Discriminant Analysis on human faces and the evaluation of its performance brings us a step closer to our goal of protecting the privacy of individuals in public places.



Subject Descriptors:
- I.4 Image Processing and Computer Vision
- I.3.3 Picture and Image Generation

Keywords:
- Face De-identification,privacy,image alteration


Implementation Software and Hardware:
- Matlab 2009,Open CV 2.4, Visual Studio 2010

## Acknowledgement

I would like to thank my project supervisor, Dr. Terence Sim for giving me his support, guidance and feedback throughout the entire whole project. I would also like to thank the PhD candidate Zhang Li for assisting me in various stages of the project.

I would also like to express my deepest gratitude to my parents for supporting me in my educational endeavours.

# List of Figures


3.1 Shows that MMDA resolves a data vector which is a face into orthogonal vectors corresponding to each mode present in the vector. Each of these modes can then be altered independently. 15
3.2 The steps in the MMDA process. . . . . . . . . . . . . . . . . 17
3.3 Shows a set of 4 images. From left to right, the first one is one of the training images used to create the model, the second one shows the landmark points detected by Stasm, the third one shows the reference face and the last one shows the landmark points that were detected by Stasm. . . . . . . 19
3.4 Shows the result when the training image from the Figure 3 is warped onto the reference face. . . . . . . . . . . . . . . . . 20
3.5 Shows a warped image with the mask to remove the ears and background from the image. . . . . . . . . . . . . . . . . . . . 20
3.6 Users can select the photo they want. . . . . . . . . . . . . . 24
3.7 The user can now choose his desired alteration from the 4 active panels at the bottom. . . . . . . . . . . . . . . . . . . . 25
3.8 The transformed image will be shown after the GO! button is pressed. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 26

4.1 Photos of males in the training set. . . . . . . . . . . . . . . . 29
4.2 Photos of females in the training set. . . . . . . . . . . . . . . 30
4.3 Images in the testing set. . . . . . . . . . . . . . . . . . . . . 31
4.4 Single(Gender) attribute transform,-$G2A3R1\_G1A3R1\_I10$. . 33
4.5 Double(Age, Race) attribute transform,-$G2A2R3\_G2A3R1\_I10$. 33
4.6 Triple attribute transform,-$G1A1R2\_G2A3R1\_I10$. . . . . . . 34
4.7 Changing intensity,-$G2A3R1\_G1A3R1\_I00, 10, 20$. . . . . . . . 34
4.8 Changing identity,-$G2A2R3\_G2A2R3\_I0.5$. . . . . . . . . . . 35
4.9 Questions based on the type of transformation. . . . . . . . . 37
4.10 Questions based on identity transformation. . . . . . . . . . . 38
4.11 Questions based on ranking intensities. . . . . . . . . . . . . . 39
4.12 Results for single attribute transformations from all respondents 42
4.13 Results for single attribute transformations from consistent respondents . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44
4.14 Results for double attribute transformations from all respondents 44

4.15 Results for double attribute transformations from consistent respondents . . . 46
4.16 Results for triple attribute transformations from all respondents 46
4.17 Results for triple attribute transformations from consistent respondents . . . 47
4.18 Results for identity transformations from all respondents . . . 48
4.19 Results for identity transformations from consistent respondents 48
4.20 Results for questions on ranking the intensity for all the classes from all participants . . . 49
4.21 Results for questions on ranking the intensity for all the classes from consistent participants . . . 49
4.22 Proportion of people who got both genders right in a gender transform for all users. . . . 50
4.23 Proportion of people who got both genders right in a gender transform for consistent users. . . . 51
4.24 Proportion of people who got both genders right in a gender transform according to their ages and genders for all users. . 51
4.25 Proportion of people who got both genders right in a gender transform according to their ages and genders for consistent users. . . . 51
4.26 Proportion of people who got both genders right in a gender transform according to their races and genders for all users. . 52
4.27 Proportion of people who got both genders right in a gender transform according to their races and genders for consistent users. . . . 52
4.28 Proportion of people who chose each of the options among all users when the gender was transformed . . . 53
4.29 Proportion of people who chose each of the options among consistent users when the gender was transformed . . . 53
4.30 Proportion of people who chose each of the options among all users when the race was transformed . . . 53
4.31 Proportion of people who chose each of the options among consistent users when the race was transformed . . . 53

# Table of Contents

**Title** i

**Abstract** ii

**Acknowledgement** iii

**List of Figures** iv

**1 Introduction** 1
- 1.1 Overview . . . 1
- 1.2 Motivation . . . 2
- 1.3 Our Approach . . . 2
- 1.4 Contribution . . . 3

**2 Literature Review** 4
- 2.1 Pixelation . . . 4
- 2.2 Blurring . . . 5
- 2.3 BlackOut method . . . 5
- 2.4 k-Same Algorithm . . . 6
- 2.5 k-Same-Select Algorithm . . . 7
- 2.6 De-Identification Camera . . . 7
- 2.7 Person De identification in videos . . . 8
- 2.8 Using Active Appearance Model (AAM) to preserve privacy . 9
- 2.9 Automatic extraction of face identity-subspace . . . 10
- 2.10 Using Principle Component Analysis to encode for different facial attributes . . . 11

**3 MMDA** 12
- 3.1 Mathematical Background . . . 12
- 3.2 Applying MMDA for face alteration . . . 16
- 3.3 Creating Demo GUI . . . 23

**4 Experiments** **28**
4.1 Types of alteration . . . . . . . . . . . . . . . . . . . . . . . . . 28
4.2 Evaluating the performance of MMDA . . . . . . . . . . . . . . 35
4.3 Results and Discussion . . . . . . . . . . . . . . . . . . . . . . . 41

**5 Conclusion and Future Work** **55**

**A References** **A-1**

**B Code** **B-1**

# Chapter 1

# Introduction

## 1.1 Overview

The prevalence of digital cameras in public places has sparked debates and protests about preserving privacy in recent years. For example, some activists in Berlin teamed up to destroy surveillance cameras in public places (The Guardian newspaper,2013). Such incidences highlight need to come up with an effective de-identification algorithm that protects the the privacy of the subjects in video sequences while ensuring the usability of these images. The resultant images of the algorithm should be able to give us some information about the person while preserving his identity. Many previous methods to solve this problem retain few useful information about the original image. We will discuss these methods under the literature review section.

## 1.2 Motivation

The need for an effective de-identification algorithm which retains some information about the original image serves as a motivation for us to come up with a method to alter some attributes of the human face while retaining others. The retained attributes can give us information about the person who is caught on the camera.For example, demographic information about the people in a particular location can be collected by using this method. The information that we can derive about the person after seeing his or her face is defined as attributes. Some examples of attributes are gender, race, age and identity.

## 1.3 Our Approach

The Mutimodal Discriminant Analysis or MMDA can be used to decompose human faces into various modes like gender, race, age and identity which occupy orthogonal spaces. This enables us to alter each of the modes or attributes separately without affecting others. A MMDA model is first obtained by using a training set of images and this model will then be used to alter new faces. We hope that this method will eventually be used to preserve the privacy of the individuals in public places. The performance of this method depends on whether it is able to convince users that a particular alteration has taken place. Surveys were conducted to find out if users are able to perceive the alterations that were made by the MMDA method. These surveys were used to evaluate the performance of the method when it is used to alter faces.

## 1.4 Contribution

MMDA is a method that is used to decompose variations in a dataset into independent factors or modes. In this project, MMDA was used to decompose human face images into different modes like gender,race,age and identity and its performance was evaluated by the use of surveys. The results of the surveys indicate the areas of the method that need to be improved in order for it to be used as a de identification algorithm in future.

# Chapter 2

# Literature Review

In this section, we discuss other methods that are commonly used to protect the privacy of an individual.

## 2.1 Pixelation

A parameter p, called the pixalation factor, p is used to divide the original image into blocks of pixels of size p*p. The pixels in each block are then replaced by the average pixel value of all the pixels in that block. As the value of p is increased, more information is removed from that image.

Higher values of p have to be chosen in order to protect the privacy of the subjects. However, if the value of p is sufficiently high, the resultant image may not even resemble a human face.This will reduce the the usability of these images for classification tasks.(Newton, Sweeney,Torre and Baker, 2003) Therefore, this method is unable to protect the privacy of the subjects while ensuring the usability of the de-identified images.

## 2.2 Blurring

When an image is blurred, each pixel in the image is replaced by a weighted average of the pixels neighbourhood. The Gaussian blur is a commonly used blurring technique. In this technique, the weight of the pixels near the center of the neighbourhood is higher. The size of the neighbourhood is positively correlated to the amount of blurring effect.

When low amounts of blurring is used, the identity of the subject cannot be protected. However, more information about the original image is lost at high levels of blurring. The resultant image may not even look like a human face at high levels of blurring. Hence this method will not be able to protect the privacy of the subjects while ensuring the usability of the de-identified images. (Newton et al, 2003)

## 2.3 BlackOut method

Let H be a person specific face set. The Black out method sets the pixel values of all pixels that represent the faces to 0. This method replaces the faces of people with a black box. If H is a person specific face set, and if a copy of one of the images from the set is de-identified using the Black out method, correct face recognition is limited to guessing with a probability of 1/H.

The resulting image of this algorithm is the same regardless of the input facial image. Moreover, the resulting image does not resemble a human face. Hence, while this method is able to guard the privacy of the subject, it is unable to retain any useful information about the person.

## 2.4 k-Same Algorithm

Let H be the person specific face set which contains k faces. The k-Same algorithm (Newton et al, 2003) will replace each of the faces with another image where each of the N pixels has the average computed based on its pixel position and on the images in the face set H.Correct face recognition will again be limited to guessing with a probability of 1/k because all the faces from the faceset will look identical after the de identification process.

If the face images in the face set share very little attributes, the resultant face will not be able to give us the correct demographic information about the faces in the face set. However, this method can give us some information about the faces that make up the face set if the number of faces in the set is small and if most of them have share the same attributes.While this method may preserve some information about all the faces in the face set, this information cannot be used to infer any details about any one person in the face set.

In contrast to the BlackOut method, the resultant image of this method is still a face which may or may not contain useful information about the faces in the face set depending on the images in the face set. Hence, the Average method is preferred over the BlackOut method because the images it generates can give us more information.

## 2.5 k-Same-Select Algorithm

The k-same-Select Algorithm was introduced as an extension of the k-Same algorithm to improve the data utility in the k-Same algorithm. In this technique (Newton et al, 2003) , we divide face set H into smaller groups of k images each. Face images that have been represented by a column vector having N cells can also be represented as a point in an N dimensional space. The closest neighbours of a point can then be identified by using the Euclidian distance function. A point together with its k-1 neighbours can form a cluster. After this, de identification techniques like the k-Same Algorithm can be applied on each cluster.The correct face recognition will be limited to guessing with a probability of 1/k.

In comparison with the k-Same Algorithm, this method can preserve more information about the original face image in the de-identified face because each cluster contains faces that share more attributes.

## 2.6 De-Identification Camera

In this method ( Mrityunjay, Narayanan,2011), the camera identifies the full body image of the people in the scene, tracks his or her motion over a period of time and then preserves his or her privacy by using Gaussian blur or binarizing the intensity values on the parts of the frame that contains the person.

The strength of this method lies in the fact that it tries to protect the identity of the person by altering the image of a person's whole body unlike methods like the k-same and the k-same-select which only focus on altering the facial images of the subject.

However, the de- identification scheme of this method relies on using the Gaussian blur technique. If this technique is used, this method will suffer from the drawbacks that were mentioned in the Blurring method. If the pixel values of the part of the image that contains a person is binarized, some parts of the person's image will appear black while others will appear white. While we may still be able to see an out line of a human being doing an action, we may not be able to derive other information like his gender, race or age from that image.

## 2.7 Person De identification in videos

In this method ( Prachi Agrawal and P. J. Narayanan,2011), humans are identified, segmented and de-identified using the exponential blur technique or the line integral convolution (LIC) technique.The LIC technique distorts the boundaries of a person in order to obfuscate his or her silhouette by using a filter kernel.

In this method, the de-identification techniques are used in such a way that it is able to obscures any other characteristics besides his or her action. This means that the resultant image will not be able to give us any information about the person's gender, race or age. This limits the usability of the resultant images.

## 2.8 Using Active Appearance Model (AAM) to preserve privacy

AAM (Cootes,Edward and Taylor 2001) is a method of matching statistical models of appearance to new images based on a set of model parameters that control modes of shape and grey level variation that were learned from a training set.An iterative matching algorithm that learns the relationship between the changes in the model parameters and induced image errors is used to to match each new image to a suitable model.

This method can be used to protect privacy (Yu and Babaguchi,2007).In the preprocessing procedure, a set of input face images go through the training model to build a statistical Active Appearance Model. During the encoding process model parameters of the input frame is obtained by using the model that was obtained in the preprocessing stage. Based on these values, a mask face is obtained and an anonymous frame is obtained by imposing the mask face onto the original face. Lastly, the privacy information of the person is embedded into the anonymous frame by using the Quantization Index Modulation (QIM) embedding method to provide the frame that preserves the privacy of the person.

This process is reversible. In the decoding process, the AAM parameters are extracted by using the extraction procedure of the QIM data hiding method. A frame will then be synthesized based on the extracted AAM parameters and the AAM model. This frame will then be imposed onto the privacy preserving frame to obtain the original image.

This method is able to protect the privacy of the subject by using the anonymous frame.However, more research needs to be done to find out the amount and type of details that will be preserved by the anonymous face before we can decide on the usability of the anonymous frames.

## 2.9 Automatic extraction of face identity-subspace

Human faces can be divided into functional subspaces which represent the variations in human faces like identity and expression. It is impossible to to separate the different subspaces linearly, so we apportion image weights among initial overlapping estimates of functional spaces in proportion with the subspace variance. This will separate the face into a set of non orthogonal projections. Although these set of spaces overlap with each other, each one codes for a functional space. The memorability information in the face variation describes easily described features like blemishes and this greatly increases the dimensionality of the identity space. Spatial redundancy in small spatially adjacent sub-samples of code are computed and compared to the samples found in the ensemble. This procedure will significantly increase the identification recognition in a test set.(Coston,Cootes,Edwards and Taylor,2002)

However the drawback of this method lies in the fact that the subspaces representing the variation in human faces in this method cannot be used to alter one face attribute without affecting the other attributes. This is due to the non orthogonal nature of the subspaces.

## 2.10 Using Principle Component Analysis to encode for different facial attributes

Principle Component Analysis(PCA) has been widely used to represent facial images at a lower dimension subspace. PCA can also be used to encode for the different properties of human face like gender, race and age.(Buchala,Davey,GaleandFrank,2005) Linear Discriminant Analysis was used to estimate the encoding power of the components that were obtained for each property.

It was found that the classification performances of all the properties was a lot higher than chance levels. Very few components are required to encode for properties such as gender, race and age. These components are predominantly found among the first few components. Hence, they are able to capture a large part of the variance of the data.The identity of the person on the other hand is encoded by a large number of components which are widely distributed.However, some components are found to be encoding for multiple properties. As a result, this method cannot be used to alter some attributes of the human face without affecting others.

# Chapter 3

# MMDA

## 3.1 Mathematical Background

We first need to represent each image by a column vector, $x_i$. Let $X = \{x_1, ....., x_N\}$ with $x_i \epsilon R_D$, which denotes a dataset of D dimensional feature vectors. Each feature vector belongs to one of C classes $\{ L_1, ..., L_c\}$ Let $m_k$ represent the mean of the class$L_k$, and suppose that each class has the same number of vectors, n so that N=nC. Without loss of generality, we can assume that the global mean of X is 0. This means that$(\Sigma_i x_i)/N = m$ =0. If this is not the case, we can always subtract m from each $x_i$.

We need to whiten the data before we can proceed with the method. To whiten the data, we will first compute the total scatter matrix $S_t = \sum_{i=1}^{N} x_i x_i^T$, then eigen-decompose it to get $S_t = UDU^T$. Only non-zero eigenvalues in the diagonal matrix , D and their corresponding eigenvectors in matrix U should be kept. After this we can compute another matrix, $P = UD^-0.5$. The shape of this matrix is $(N-1) \times D$. The P matrix can then be applied to the data to get the $(N-1) \times N$ matrix $\tilde{X} = P^T X . \tilde{X}$ represents the whitened data.

Let the whitened class means be $\tilde{m}_k$. We then proceed to find the between class scatter matrix $\tilde{S}_b$ and the within class scatter matrix $\tilde{S}_w$ of the whitened data.

$$\tilde{S}_b = \sum_{k=1}^{C} n \tilde{m}_k \tilde{m}_k^T$$
$$\tilde{S}_w = \sum_{i=1}^{C} \sum_{\tilde{x}_i \epsilon L_k} (\tilde{x}_i - \tilde{m}_k)(\tilde{x}_i - \tilde{m}_k)^T$$

According to the the Fisher Criterion,

$$J_F(\phi) = trace\{(\phi^T \tilde{S}_w \phi)^{-1} (\phi^T \tilde{S}_b \phi)\}$$

is equal to the ratio $\lambda_b / \lambda_w$ where $\lambda_b$ and $\lambda_w$ are the eigenvectors of $\tilde{S}_b$ and $\tilde{S}_w$ respectively.$\lambda_b + \lambda_w = 1$, so that by keeping the eigenvectors corresponding to $\lambda_b = 1$ in a matrix V, the subspace spanned by V achieves $J_f = \lambda_b / \lambda_w = 1/0 = +\infty$, and is therefore the most discriminative subspace. It has a dimension of C-1.

The property of the subspace spanned by V is termed as the Identity Space. It can be proven that

1. All points from the same class are projected onto the class mean.
2. All within class variation has been projected out.

This means that the identity space shows the class identity and that all classes are separated from each other.

Theorem 1: In WFLD, if V is the set of eigenvectors of $\tilde{S}_w$ associated with $\lambda_w$ =0, then

$$V^T \tilde{x}_i = V^T \tilde{m}_k$$

So far, we have derived the Identity Space based on the classes of one mode. We can extend this concept to for multiple modes like illumination and pose. To do this, we must compute the scatter matrix of each mode p: $\tilde{S}_b^p V^p = V^p$.

For a dataset $\tilde{X}$ , with M modes , with the pth mode having $C^p$ classes and each class in that mode having $n^p$ datapoints. Suppose that the dataset contains the full Cartesian product of the modes. This means that the face dataset contains $C^1$ people with each person under $C^2$ poses and each pose under $C^3$ illuminations. The total number of data points can be found by taking the product of the number of classes of all modes:

$$\Pi_{p=1}^{M} C^p$$

Theorem 2: If $V_p$ and $V_q$ are the Identity Spaces for modes p and q where $p \neq q$ then

$$(V^p)^T V^q = 0$$

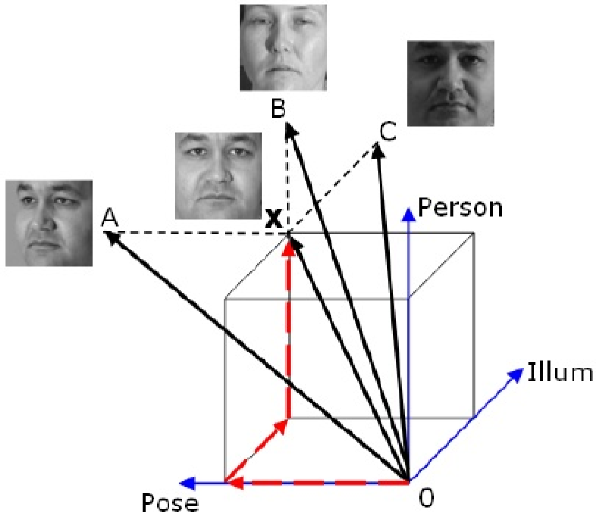


Figure 3.1: Shows that MMDA resolves a data vector which is a face into orthogonal vectors corresponding to each mode present in the vector. Each of these modes can then be altered independently.

In the figure above, each orthogonal axis is a subspace and its dimension is $C^p - 1$. Moving along each axis only alters that particular mode and all other modes will not be affected. In Figure 1, face X is decomposed into three modes by using the MMDA method. Face A shows the same person in the same illumination but different pose while face B shows a different person in the same illumination and pose while face C shows the same person in the same pose but different illumination.

All the identity vectors taken together will have a dimension of $\Sigma_{p=1}^{M} C^p - M$. The whitened data vector on the other hand has a dimension of N-1. The remaining subspace therefore has a dimension of $r_0 = N - \Sigma_{p=1}^{M} C^p + M - 1$ and this is termed as Residual Space.

Residual space is actually the intersection of the within class scatter matrix of all modes. It contains any residual variations that are found outside of all the Identity Space.

## 3.2 Applying MMDA for face alteration

MMDA can decompose an images into its numerous modes which are independent of all the other modes. This process requires the images to be processed before and after the MMDA technique has been applied. In this project, we use datasets containing images of human faces. Each step in the whole process will be discussed in this section. In this report, the entire process is termed as the MMDA process.

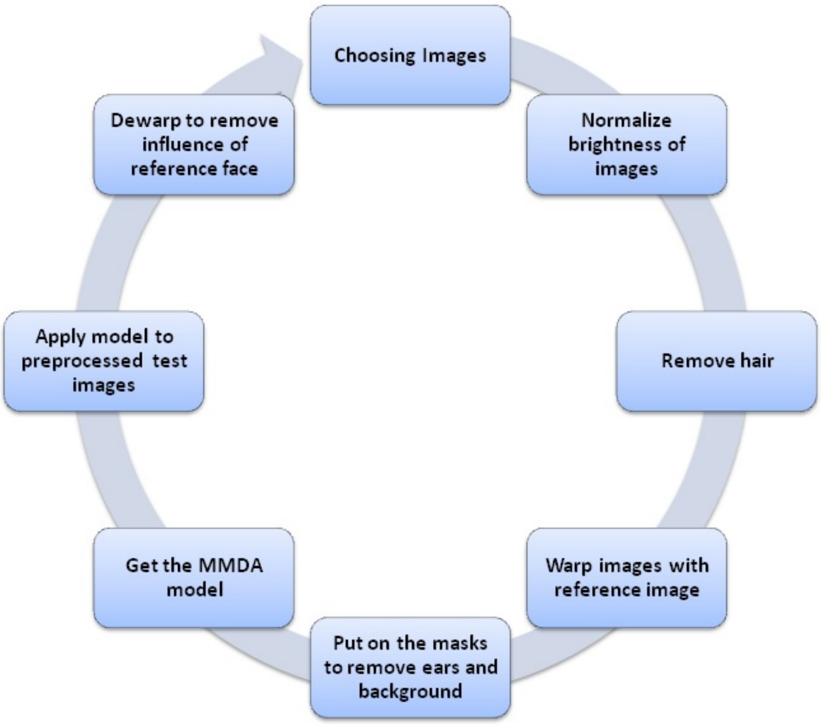


Figure 3.2: The steps in the MMDA process.

**Choosing the images**

In this project, we have decided to consider three modes. They are the gender, race and age. The first mode, gender has two classes which are male and female and the second mode, race has three classes which are Caucasian, Indian and Oriental and the third mode, age has three classes which are the ages between 20-30, 40-50 and above 50.

The training dataset had a total of 18 people. There are 9 males and 9 females. In each gender group, there are 3 people who belong to the Caucasian, Indian and Oriental race. Of the three people who belong to each of the races, one of them is aged between 20 - 30 while another person is aged between 40 - 50 while the third person is above 50 years of age.

**Normalize the brightness of images**

Some images have uneven brightness. This causes the resulting image after the MMDA process has been applied on these images to have varying brightness as well. Programs which assess the effectiveness of the MMDA algorithm may be adversely affected by the uneven brightness in the image.

Hence we used a simple program to normalize the brightness of the image before putting it through the MMDA process. Since the images from the database are in color, we consider the red array, green array and the blue array separately. We find the minimum and the maximum value in this array. Let the array be represented by X and let the minimum and the maximum value be represented by min and max respectively. We then normalize the image by recalculating the values based on the difference between the minimum and the maximum value. This will ensure that all parts of the image are of equal brightness. (see Appendix B -partA1)

**Removing hair from the images**

These modes can be easily observed from the frontal side of the face. All the landmark points used in this technique are also found on the face. Hence, it is necessary for us to crop the image such that the hair is no longer visible on the image. The function cropAndAlign was used to align all the faces as well as to remove part of the forehead. (see Appendix B- partA2)

**Warping**

The goal of the warping processes is to deform an image such that it matches another one. To do this, the users will first have to provide some matching points on the two images. The warping process is completed when the points in the first image is interpolated to corresponding points on the second image. In this project, Stasm, which is an opensource C++ library for finding features in face has been used to find the matching points in the two images. Images are warped to a reference face so that slight tilts in the original image can be rectified. The reference image is usually an image which shows a face that looks straight at the camera. The warping process is done by using the ThinPlateSpline(Lombaert,2006) code that is available online.

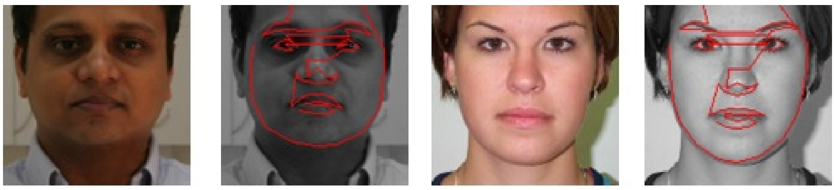

Figure 3.3: Shows a set of 4 images. From left to right, the first one is one of the training images used to create the model, the second one shows the landmark points detected by Stasm, the third one shows the reference face and the last one shows the landmark points that were detected by Stasm.

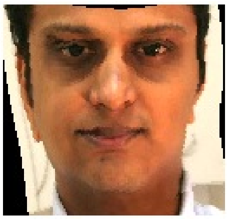

Figure 3.4: Shows the result when the training image from the Figure 3 is warped onto the reference face.

**Putting on the mask**

A mask has to be placed on the warped image to remove the ears and the background of the image. This was done by using Photoshop. Skipping this step will result in the production of images with many artifacts because the MMDA technique will try to separate all parts of the image including the ears and the background into the different modes.

Figure 3.5: Shows a warped image with the mask to remove the ears and background from the image.

**Implementing the MMDA method to get the model**

The training dataset of 18 images is stored under the matrix obs_u of the size of the size $15758 \times 18$ where the first 15 616 elements represent the pixel values of the image and the next 142 elements are the landmarks of the image.

The labels of gender, race and age are stored in a matrix called labels. The shape of this matrix is $3 \times 18$. The first row of the matrix stores the information about the gender. The number 1 represents male and 2 represents female. The second row of the matrix stores information about the race. The number 1 represents the Caucasian race while the number 2 represents the Indian race while the number 3 represents the oriental race. The last row represents the age group. Three age groups are considered in this project. The first age group is between 20-30 years of age and the second age group is between the 40-50 years of age and the third age group is the age group of people above the age of 50.

The data in obs_u is first whitened by using the WFLD method described earlier. The data is whitened and stored in a matrix X with a size of $17 \times 18$. Each image in the dataset is represented by 17 principle components. (see Appendix B-part B1)

We then project the whitened data that is contained in the matrix X into orthogonal spaces. This will help us to create the structure m, which is the MMDA model. (see Appendix B-part B2)

**Using the model to alter test images**

Let M be the model that was trained on the 18 training images. Let the single vector X represent the testing image. The top elements of the vector represent the pixel values of the image from left to right and top to bottom while the bottom values represent the landmark points of the test image. These landmark points were obtained by using Stasm programme. Vector t should then be calculated according to the following equation.

$$t = M.Q' \times M.P' \times (X - M.org)$$

In this equation, Q' refers to the transpose of the vector Q while P' refers to the transpose of vector P.

The t vector is a 17 x 1 vector which contains the coefficients for gender,race and age in the top 5 elements. The first element contains the coefficient for gender while the second and third elements contain the coefficients for race while the fourth and the fifth element contains the coefficient for age. The other 12 elements are the coefficients in the Residual space. When we want to alter a certain attribute of the test face, appropriate portions of the t vector should be replaced with the class means from the structure, m.lowobs which contains the identity vector for each class.

After modifying vector t to get t2,it has to be reconstructed according to the following equation.

$$Xr = M.Pr \times M.Q \times t2 + M.org$$

Now Xr is a vector containing the color and landmark points of the altered face. However, this face still warped over the reference face. Hence, this vector has to be reshaped and de warped from the reference face to obtain the altered face which is not affected by the reference face. (see Appendix B-part B3)

**Normalize the brightness**

The pixels which represent the new face are contained in the top part of the Xr matrix. The MMDA process may have caused some pixel values to exceed 255 while other values may be lower than 0. Hence, these values have to be normalized again before we can save the information as an image. In this step, we use the same equation that we used in the previous normalization step.

**Dewarping**

The new face that we have attained through MMDA implicitly contains the shape of the reference face. So now we dewarp the face using the landmarks in the new face and the landmarks in the reference face. We interpolate the reference landmarks to the landmarks in the new face.

## 3.3 Creating Demo GUI

A graphical user interface in Matlab was created based on the MMDA process. The user can browse the files in his computer and choose any photo that contains a person's face. This photo should show the frontal part of the person's face.

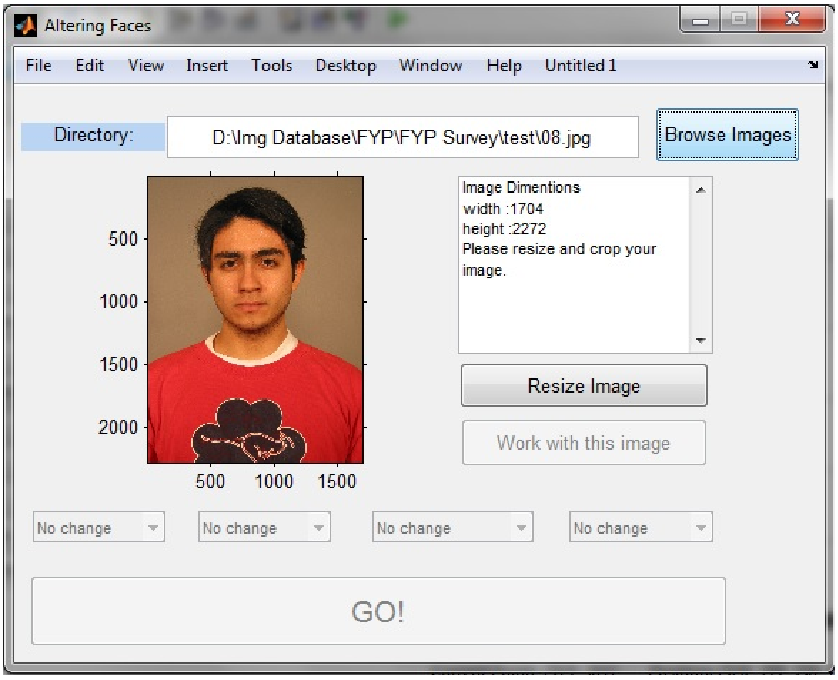


Figure 3.6: Users can select the photo they want.

After the user has chosen the photo he wants to alter, he has to resize the photo by clicking on the resize button. He should resize the original photo in such a way that the face covers more than 60% of the image panel. Then user has to use the movable window frame to choose the face of the photo. The face should fit inside the window frame. The user should then make his selection by clicking inside the window frame. The window frame which consists of the face will now fill the image panel. The 'Work with this image' button will now be activated. If the user is satisfied with the presentation of the photo, he should click on this button. Once he has clicked on this button, the 4 panels that are found at the bottom of the GUI will become active.

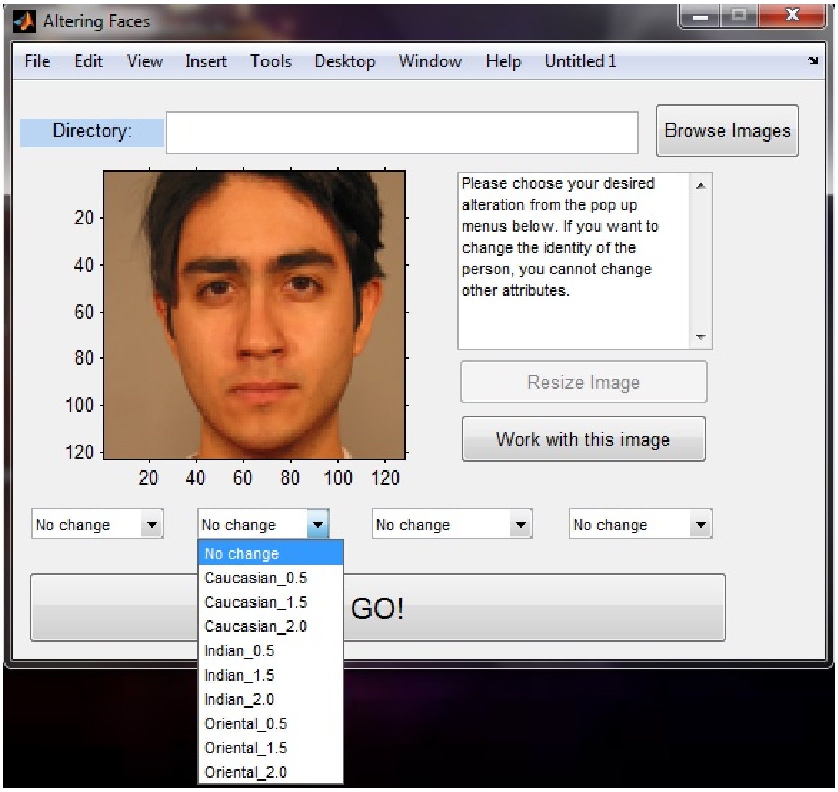


Figure 3.7: The user can now choose his desired alteration from the 4 active panels at the bottom.

The 4 active panels at the bottom control the type and the intensity of the desired alteration. The first panel shows the alteration options for gender, the second one shows the alteration options for race, the third on shows the alteration options for age and the last one shows the options for identity. The number beside the type of alteration represent the magnitude of the intensity of the alteration. If the original image shows a person who belongs to the Indian race and if the chosen alteration is Caucasian_1.5, the elements in the t matrix that represent control the race will be substituted with the elements that represent the Caucasian race from the model M. The elements that represent the Caucasian race will first be multiplied by 1.5 before they are substituted inside the t matrix.

After selecting his desired alterations, the user should press on the GO! button. The transformed image will be shown in the image panel.

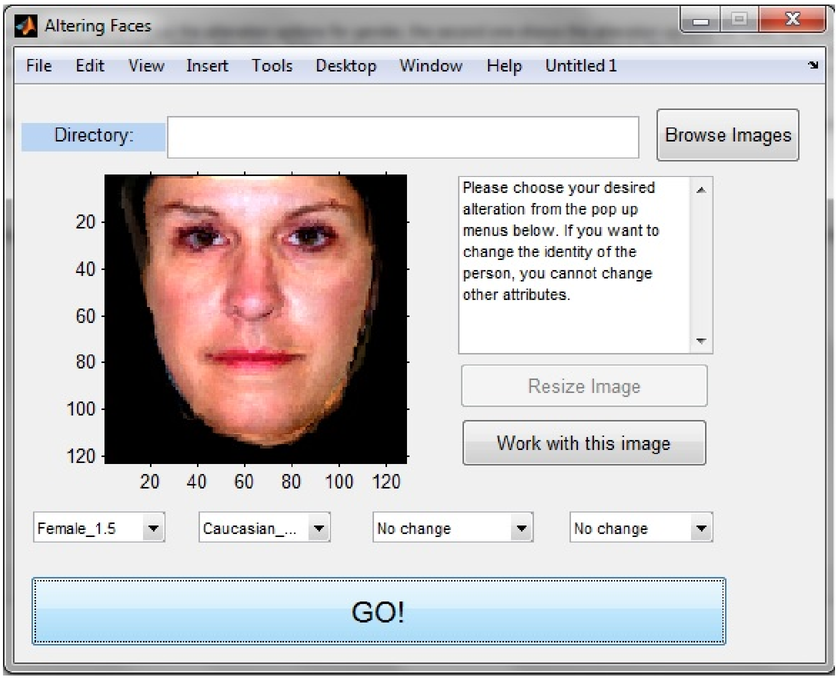


Figure 3.8: The transformed image will be shown after the GO! button is pressed.

Figure 3.8 shows the transformed image when the Indian man in the figure 3.7 is transformed into a female with an intensity of 1.5 and into a person who belongs to the Caucasian race with an intensity of 1.5.

# Chapter 4

# Experiments

## 4.1 Types of alteration

In this project, we have tried to alter the gender, race, age and combinations of these modes in the images in the testing set by using the MMDA model we obtained from the training set. The training set contained 18 people. There were 9 males and 9 females. In each gender, 3 people people belonged to each of the three races. The races are Caucasian, Indian and Oriental. Of these three people, one person was aged between 20-30, another was aged between 40-50 and the third person is above 50 years old. All training images were of high resolution. Spectacle frames and the presence of facial hair may result in artefacts in the final images. Hence, care was taken to make sure that none of them wore spectacles and that all males were clean shaven.

The following figure shows the photos that were used in the training set. The string of letters and numbers below each image describes the attributes of the person. G1 refers to the male gender, G2 refers to the female gender, R1 refers to the Caucasian race, R2 refers to the Indian race , R3 refers to the Oriental race , A1 refers to the 20-30 age group ,A2 refers to the 30-40 age group and A3 refers to the above 50 group.


Figure 4.1: Photos of males in the training set.

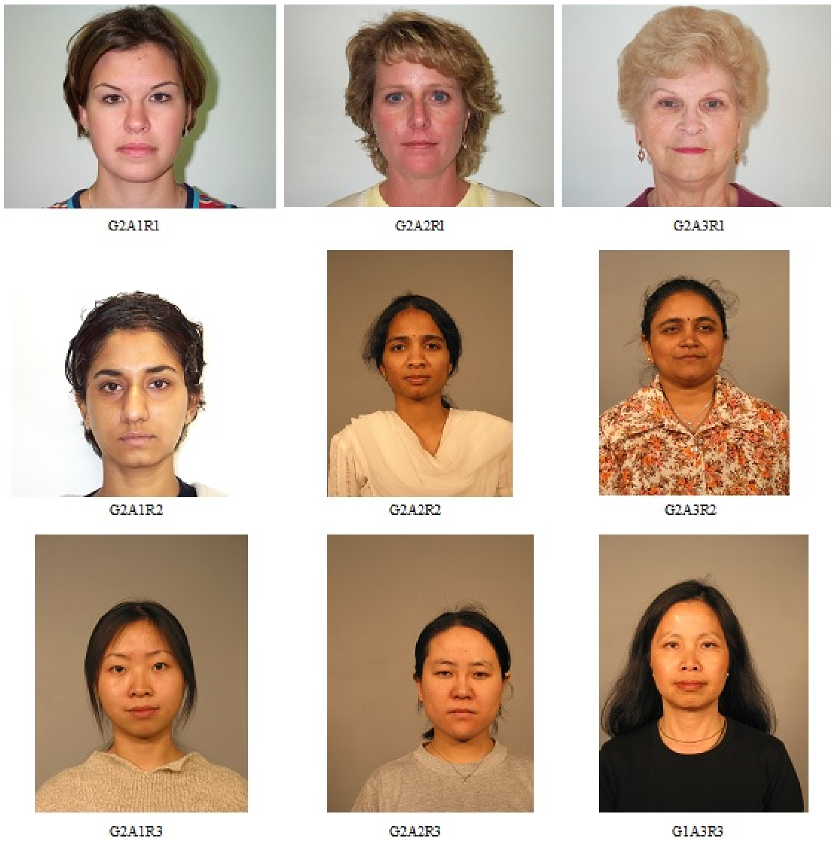


Figure 4.2: Photos of females in the training set.

The figure on the next page shows the test images that were chosen for the experiment.


Figure 4.3: Images in the testing set.

Each of the test faces was transformed into each possible combination of gender,race and age. Since we have 2 classes under the gender mode and three classes under the race and age mode, the maximum number of different combinations of these modes is $2 \times 3 \times 3 = 18$. Transformations in which only a single attribute was changed are called 'single attribute transformations' while transformations in which double attributes are changes are called 'double attribute transformations' while transformations in which triple attributes are changed are called as 'triple attribute transformations'.

For single attribute transformations, the intensity of the transform was varied. The elements from the model M that were used to replace the corresponding elements in the t matrix were multiplied by 0,1 and 2 before they were substituted into the t matrix.

In this report transformations are described in the following format, G2A3R1_G1A3R1_I10. The attributes are described by using the same symbols as mentioned previously in this section. The portion of the string of characters before _ describe the original person while the second set of characters after the _ describe the attributes of the transformed image.For example, G2A3R1_G1A3R1_I10 means that the original person is a female who belongs to the Caucasian race and is above 50 years of age. A gender transformation took place and the transformed image should show a male who belongs to the Caucasian race and whose age is above 50. In this transformation, all other attributes besides gender should be preserved. The letter followed by the number after the second _ shows the intensity of the transform. The substituted elements in the t vector were multiplied by value that is obtained by dividing this number by 10. In this case the substituted elements are multiplied by 1.

In the following images,the image on the left shows the original image while the one on the right shows the transformed image.

Figure 4.4: Single(Gender) attribute transform,-$G2A3R1\_G1A3R1\_I10$.

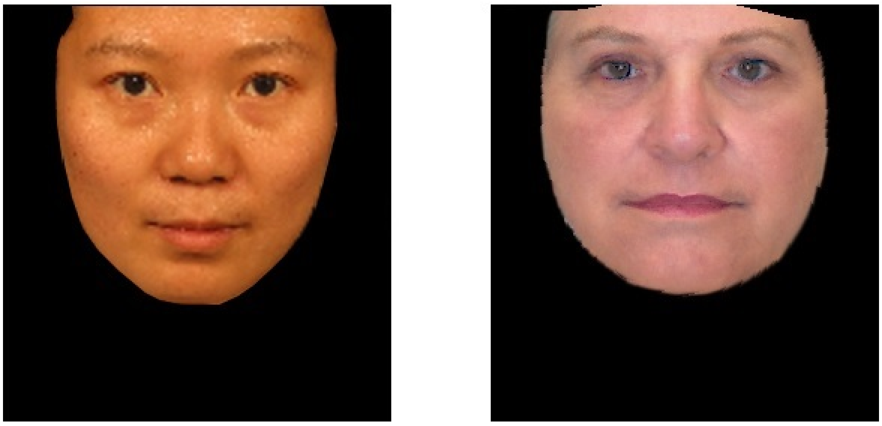

Figure 4.5: Double(Age, Race) attribute transform,-$G2A2R3\_G2A3R1\_I10$.

Figure 4.6: Triple attribute transform,-$G1A1R2\_G2A3R1\_I10$.

For single attribute transformations, the intensity of the transform was varied. The elements from the model M that were used to replace the corresponding elements in the t matrix were multiplied by 0,1 and 2 before they were substituted into the t matrix.

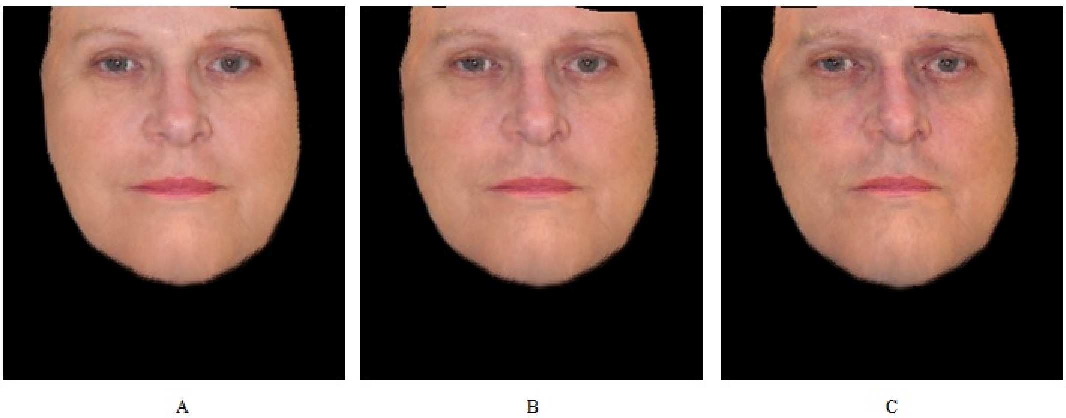


Figure 4.7: Changing intensity,-$G2A3R1\_G1A3R1\_I00, 10, 20$.

In figure 4.7, the original person is a female who belongs to the Caucasian race and is above 50 years of age. This image went through a gender transformation with in three different intensities 00,10 and 20. The resultant images of these three transformations are A,B and C respectively.

In addition to all these transformations, each of the test images also went through an identity transformation. In this transformation, the elements in the residual space of the t vector is multiplied by 0.5. While the identity space contains elements that control certain attributes like gender, race and age, the residual space contains elements that control all other attributes of the face. This implies that changing the magnitude of the elements in the residual space should not affect the gender, race or age of the original image. However, the identity of the person should change. This means that transformed image should show another person who has the same gender,race and age as the original person.

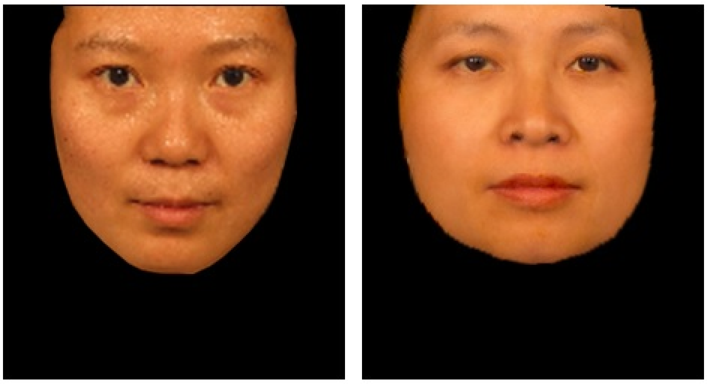

Figure 4.8: Changing identity,-$G2A2R3\_G2A2R3\_I0.5$.

## 4.2 Evaluating the performance of MMDA

In order to evaluate the performance of MMDA on these 10 test faces, a survey was designed based on each of the 10 images. 30 participants took part in each of these surveys.

### Objectives of the survey

The survey was designed to answer the following questions.

1. When the MMDA algorithm is used to alter an attribute in the test image, is the change perceived by the user?

2. When the MMDA algorithm is used to alter one attribute, does the user perceive a change in another attribute?

3. When the MMDA algorithm changes the intensity of one attribute, is this change perceived by the user?

### Survey Design

The survey has three main parts.

1.Questions based on the 18 transformations for that particular face. The aim of these questions is to check if the participant can identify the correct transformation that has taken place.

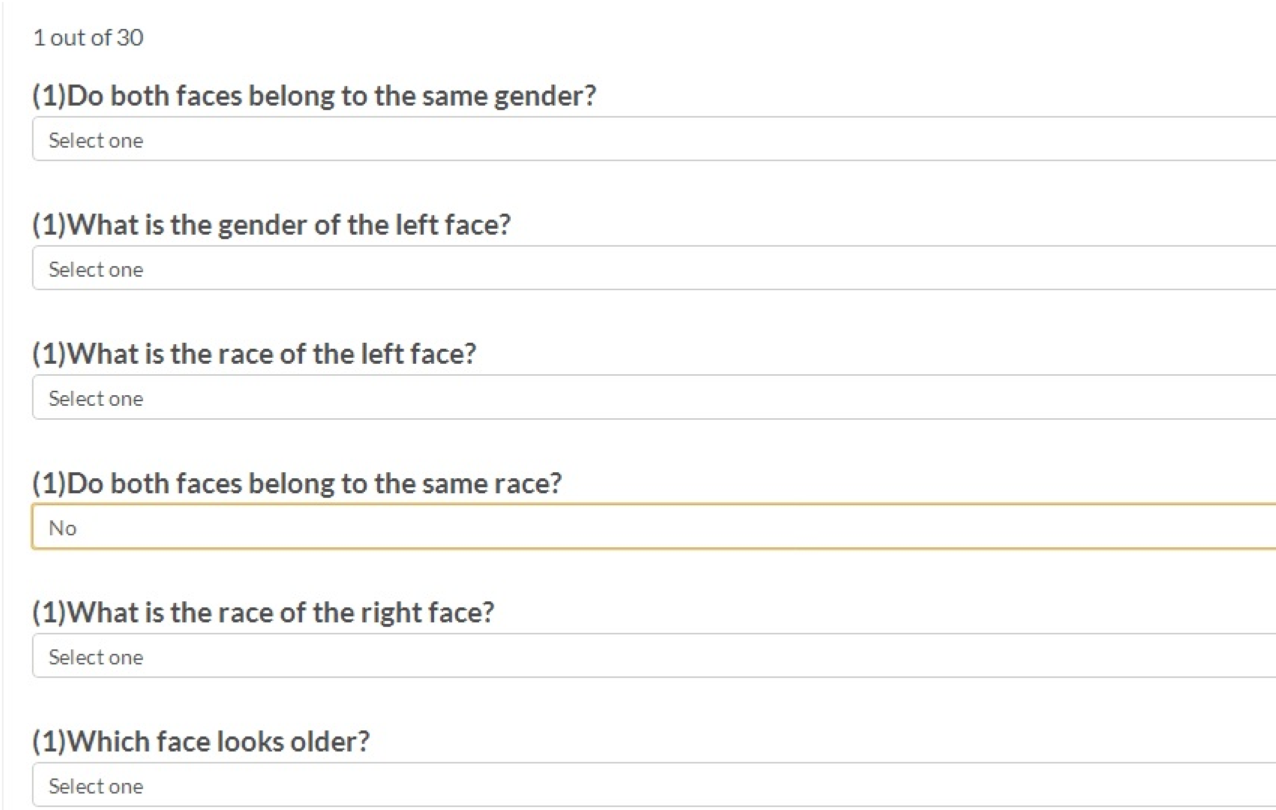


Figure 4.9: Questions based on the type of transformation.

Figure 4.9 shows the questions that follow after a pair of images like figure 4.6 is shown.The first question in the question set asks "Do both faces belong to the same gender?". The options given are "Yes" and "No". The second question asks "What is the gender of the left face?". The options given are "Male" and "Female". The third questions asks "What is the race of the left face?". The options given are "Caucasian","Indian" and "Oriental". The fourth question asks "Do both faces belong to the same race?" If the user answers 'no' the question, another questions which asks for the race of the right face will appear, If not, this question will not appear.The last question asks "Which face looks older?". The options given are "The face on the right" , "The face on the left" and "Both of them are of the same age". In all the pair of images that show the transformations, the image on the left shows the original image while the one on the right shows the transformed image.

2.Questions based on the identity transform. The main aim of this question is to check if the user can perceive that the transformed face has the same gender,race and age as the original image while knowing that both images are showing two different persons.


20 out of 30

(20)Are both images showing the same person?

No

(20)Do both faces belong to the same gender?

Select one

(20)What is the gender of the left face?

Select one

(20)What is the race of the left face?

Select one

(20)Which face looks older?

Select one

(20)Do both faces belong to the same race?

No

(20)What is the race of the right face?

Select one

Figure 4.10: Questions based on identity transformation.

Figure 4.10 shows the set of questions that follow a pair of images which show that a transformation in identity has taken place like Figure 4.8. The first question asks the participant if both images show the same person. If the participant answers 'no', all the other questions will appear. Otherwise, he has to move on to the next question based on another transform.

3.Questions based on the different intensities of the transformations. The aim of this type of questions is to check if the users can perceive the change in intensities of the transformations correctly.

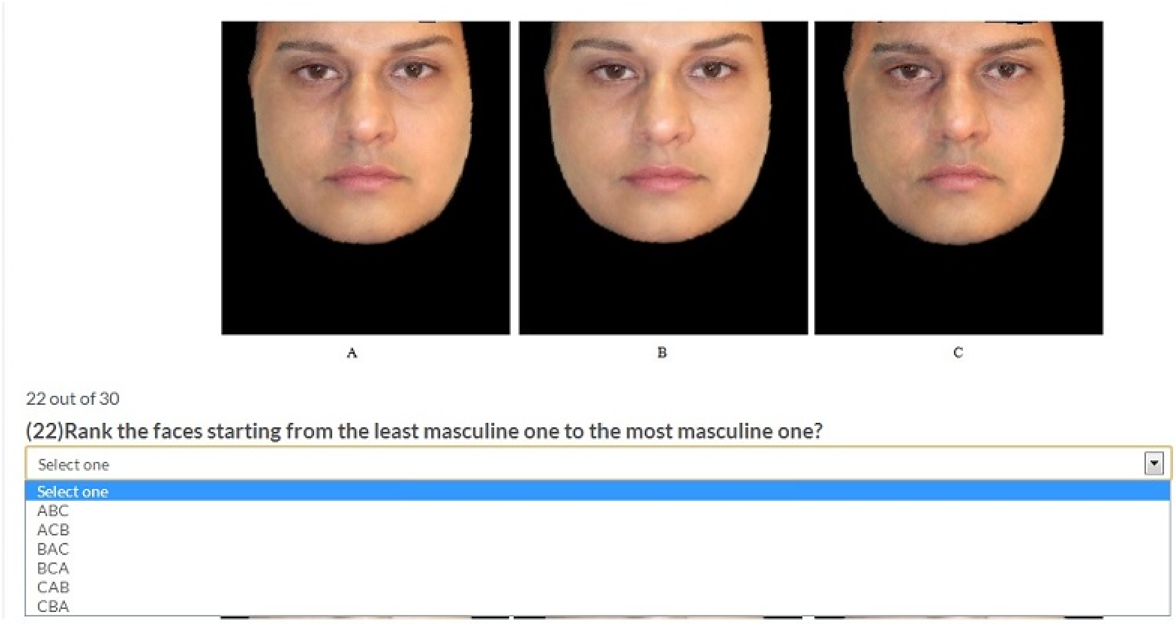


Figure 4.11: Questions based on ranking intensities.

Similar questions like the one in Figure 4.11 were asked for each of the single attribute transformations.

In addition to these questions 3 questions were repeated. Participants who gave the same answers to 2 or more of the three questions were considered to be 'consistent' survey respondents. The results from the group of 'consistent' users is considered to be more reliable than the results from all the participants.

**How the results were calculated**

In the first part of the survey, the users answered the questions based on single, double and triple attribute transformations. The answers that were given to each type of transformation was analysed separately. The proportion of respondents who perceived a change in gender,race and age correctly for each type of transformation was computed in the following way.

To calculate the proportion of respondents who perceived a change in gender correctly in a survey, we count the number of participants who perceived the gender of the left or the original image correctly. We then calculate the number of participants who got the gender of both the original and the transformed person correct. To calculate the proportion, we divide the number of people of people who got both genders right by the number of people who perceived the gender of the original image correctly in all the surveys.

To calculate the proportion of respondents who perceived a change in race correctly in a survey, we count the number of participants who perceived the race of the left or the original person correctly. Then, we calculate the number of participants who got the race of both the original and the transformed person correct. To calculate the proportion, we divide the number of people who got both races correct by the number of people who perceived the race of the original image correctly in all the surveys.

In the survey, the participants are made to compare the ages of the original images and the age of the transformed image. The question "Which face looks older?" asks them to do so.The user is given three options to choose from. The first option is "The face on the right" and the second option is "The face on the left" and the last option is "Both of them are of the same age". To calculate the proportion of people who perceived the change in age correctly, we divide the number of participants who choose the correct option by the total number of people who took the surveys.

In this method, we ignore people who are not able to perceive the gender or race of the original image properly. This is a reasonable way to compute the proportion because the users should be able to identify the gender and the race of the person in the original image correctly. If he or she is not able to do this correctly, it is highly likely that the user is not trying his best to answer the survey questions. Hence, his or her answers have to be ignored.

## 4.3 Results and Discussion

In this section, the results of the survey will be presented in a tabular format. Types of alterations which were not perceived correctly to a large extent will be analysed in detail.

| Alterations | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right |
|---|---|---|---|
| Gender Change | 95/260 =0.365 | 213/222=0.959 | 58/300 =0.1933 |
| Race Change | 335/521=0.643 | 341/426=0.800 | 111/600=0.185 |
| Age Change | 356/515=0.691 | 435/450=0.966 | 448/600=0.747 |

Figure 4.12: Results for single attribute transformations from all respondents

Figure 4.12 shows the proportion of survey respondents who perceived the correct change based on the images. For example, the first row of the table shows the proportion of respondents who perceived the gender, race and age correctly when the gender of the original image was transformed. This means that an if the original image was male, it will be transformed to a female face and that if the original image is female, it will be transformed to a male. The proportions were calculated by using the method mentioned in the previous section. The numerator of the fractions in the cells reflect the total number of people who got both genders or races right when each of the 10 test faces went through a gender transform. The denominator of the fractions represent the total number of people in all the surveys who got the gender or race of the original or left image correct. When we calculate the proportion for the age question, the numerator represents the total number of people who answered the question right in all the surveys when the gender transform took place. The denominator represents the total number of participants who took all the surveys.

From Figure 4.12 we can see that, the proportion of people who got both genders right when the images went through a gender transform is 0.365. In our analysis, if the proportion is greater than 0.5, we will consider that the transformation has performed well in terms of preserving or changing gender, race and age. Hence, we know that our algorithm is not able to convince people that a gender change has occurred when their gender is altered However, the proportion of people who got both genders right when the race or the age of the person is transformed, is 0.643 and 0.691 and this is higher than 0.5. This shows that our algorithm is able to preserve genders but it is not good at altering them.

The proportion of people who got both races right when there was a gender, race and age change is high. Since all these proportions are higher that 0.8, we can safely say that our algorithm is good at preserving and changing races.

The proportion of people who got both ages right is only high when the images went through an age transform. This shows that while our algorithm is good at altering ages, it is not so good at preserving ages.

| Alterations | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right |
|---|---|---|---|
| Gender Change | 42/101=0.416 | 90/90=1.00 | 15/107=0.140 |
| Race Change | 122/201=0.607 | 157/170=0.924 | 38/214=0.178 |
| Age Change | 134/203=0.660 | 178/181=0.983 | 167/214=0.780 |

Figure 4.13: Results for single attribute transformations from consistent respondents

Figure 4.13 show the proportions for the consistent survey respondents. These results are similar to the one that we obtained from all the survey users.

The following figure shows the results that were obtained for the double attribute transformations.

| Alterations | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right |
|---|---|---|---|
| Gender and Age Change | 162/517=0.313 | 428/447=0.957 | 433/600=0.722 |
| Gender and Race Change | 181/519=0.349 | 342/426=0.803 | 117/600=0.195 |
| Age and Race Change | 669/1039=0.644 | 690/850=0.812 | 848/1200=0.707 |

Figure 4.14: Results for double attribute transformations from all respondents

When the images went through the gender and age transformation or the gender and race transformation, the proportion of people who got both genders right is low. However, when gender was not one of the attributes that was changed, the proportion of people who got both genders right is high.This result about the gender perception agrees with the result that we obtained in the single attribute transform. The results regarding the gender change in double attribute transformations show that our algorithm is good at preserving the gender of the person but it does a poor job in altering the gender of the person.

Figure 4.14 also shows that the proportion of people who got both races right is high in all three double attribute transformations. This shows that our algorithm is good at preserving and altering the races of the test faces. This agrees with the results that we from in the single attribute transformations.

This figure also shows that the proportion of people who got both ages right is high only when the age attribute is altered. If the age attribute is not altered as in the alteration where gender and race are changed, the proportion of people who got both ages right is low. This shows that while our algorithm is able to alter the ages well, it does poorly when it comes to preserving the age.This agrees with the results that we obtained in the single attribute transformation.

| Alterations | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right |
|---|---|---|---|
| Gender and Age Change | 67/202=0.332 | 179/181=0.989 | 166/214=0.776 |
| Gender and Race Change | 79/201=0.393 | 152/167=0.910 | 42/214=0.196 |
| Age and Race Change | 246/407=0.604 | 309/338=0.914 | 330/428=0.771 |

Figure 4.15: Results for double attribute transformations from consistent respondents

Figure 4.15 show the proportions for the consistent survey respondents. These results are similar to the one that we obtained from all the survey users.

The following figure shows the results that were obtained for the triple attribute transformations.

| Alterations | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right |
|---|---|---|---|
| Gender , Race and Age Change | 351/1033=0.340 | 685/866=0.791 | 844/1200=0.703 |

Figure 4.16: Results for triple attribute transformations from all respondents

Figure 4.16 shows that the proportion of people who got both genders right is low. However, the proportion of people who got both races and ages right in the triple attribute alterations in high. These trends agree with the previous results which show us that our algorithm is able to alter the race and age of the users well but it is not able to alter genders well.

| Alterations | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right |
|---|---|---|---|
| Gender , Race and Age Change | 246/407=0.604 | 309/338=0.914 | 330/428=0.771 |

Figure 4.17: Results for triple attribute transformations from consistent respondents

Figure 4.17 show the proportions for the consistent survey respondents. The proportion of consistent participants who got both genders right is high. This disagrees with all the results regarding gender obtained in all the previous types of transformations. This is probably because a large proportion of people who belonged to the consistent participants group happened to get both genders right. Hence, we should still stick to the view that our algorithm is good at preserving genders but does poorly when altering them.

This figure also shows that a high proportion of people in the 'consistent' users group got both races and ages right when these attributes were altered. This is in line with our previous results.

The following figure shows the results that were obtained for the triple attribute transformations.

| Identity Question | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right | Proportion that got both identities right |
|---|---|---|---|---|
| Identity Change | 166/221=0.751 | 154/193=0.798 | 43/243=0.177 | 243/300=0.81 |

Figure 4.18: Results for identity transformations from all respondents

Figure 4.18 shows that the portion of participants who thought that the genders and race were preserved during the identity transform is high. A large proportion also thought that the two images are showing two different people.However, the proportion that got both ages right is low. This means that a majority of them thought that the age was not preserved during the transformation.

| Identity Question | Proportion that got both genders right | Proportion that got both races right | Proportion that got both ages right | Proportion that got both identities right |
|---|---|---|---|---|
| Identity Change | 50/100=0.50 | 83/86=0.965 | 13/104=0.125 | 104/107=0.972 |

Figure 4.19: Results for identity transformations from consistent respondents

Figure 4.19 shows the proportions for the consistent survey respondents. These results are similar to the one that we obtained from all the survey users.

The following figures show the results that were obtained from the questions on intensity. It shows the proportion of survey participants who managed to rank the images according to their intensities correctly for each of the different classes in gender, age and race.

| Ranking Intensity | Proportion that got rankings right |
|---|---|
| Male | 192/300=0.640 |
| Female | 143/300=0.477 |
| Age1 | 42/300=0.14 |
| Age2 | 175/300=0.583 |
| Age3 | 227/300=0.757 |
| Caucasian | 154/300=0.513 |
| Indian | 220/300=0.733 |
| Oriental | 174/300=0.580 |

Figure 4.20: Results for questions on ranking the intensity for all the classes from all participants

| Ranking Intensity | Proportion that got rankings right |
|---|---|
| Male | 80/107=0.748 |
| Female | 49/107=0.458 |
| Age1 | 19/107=0.178 |
| Age2 | 77/107=0.720 |
| Age3 | 100/107=0.935 |
| Caucasian | 64/107=0.598 |
| Indian | 89/107=0.832 |
| Oriental | 82/107=0.766 |

Figure 4.21: Results for questions on ranking the intensity for all the classes from consistent participants

Figures 4.20 and 4.21 show that a large proportion of the participants in both the groups are able to perceive the difference in intensities for all the classes correctly except for Age1.

From the previous results, it is apparent that our algorithm suffers from two main weaknesses. They are

1. Gender alteration does not perform well.
2. Age is not preserved when the alteration does not change the age of the original person.

In this part of this section, we try to find out the attributes of the original person for which the gender transformation works well . In the following figures, we get the proportion of all users and consistent users who got both genders right when when a gender alteration took place. The images only went through a gender transform and all other attributes were kept constant. These two groups of participants are categorised according to the classes of attributes that they belong to.

| Gender | | Ages | | | Race | | |
|---|---|---|---|---|---|---|---|
| Male | Female | Age1 | Age2 | Age3 | Caucasian | Indian | Oriental |
| 70/144 =0.486 | 25/116 =0.216 | 58/140 =0.414 | 15/73 =0.205 | 22/47 =0.468 | 47/122 =0.385 | 25/57 =0.439 | 23/81 =0.284 |

Figure 4.22: Proportion of people who got both genders right in a gender transform for all users.

| Gender | | Ages | | | Race | | |
|---|---|---|---|---|---|---|---|
| Male | Female | Age1 | Age2 | Age3 | Caucasian | Indian | Oriental |
| 28/46 =0.609 | 14/55 =0.255 | 16/37 =0.432 | 6/32 =0.188 | 20/32 =0.625 | 22/55 =0.4 | 9/14 =0.643 | 11/32 =0.344 |

Figure 4.23: Proportion of people who got both genders right in a gender transform for consistent users.

Figures 4.22 and 4.23 show that the gender transformation tends to work better for males than for females. It also tends to work better for people who belong to the Age3 group and for people who belong to the Indian race compared to other age or race groups. In the following figures we farther analyse to see if the original race of the person affects the performance of this transformation within each race and age group.

| Age1 | | Age2 | | Age3 | |
|---|---|---|---|---|---|
| Male | Female | Male | Female | Male | Female |
| 53/88=0.602 | 5/52=0.0962 | 4/29=0.139 | 11/44=0.25 | 9/20=0.45 | 13/27=0.481 |

Figure 4.24: Proportion of people who got both genders right in a gender transform according to their ages and genders for all users.

| Age1 | | Age2 | | Age3 | |
|---|---|---|---|---|---|
| Male | Female | Male | Female | Male | Female |
| 16/24=0.667 | 0/13=0 | 1/8=0.125 | 5/24=0.208 | 9/18=0.5 | 11/14=0.786 |

Figure 4.25: Proportion of people who got both genders right in a gender transform according to their ages and genders for consistent users.

| Caucasian | | Indian | | Oriental | |
|---|---|---|---|---|---|
| Male | Female | Male | Female | Male | Female |
| 32/55=0.582 | 15/67=0.224 | 20/30=0.667 | 5/27=0.185 | 18/59=0.305 | 5/22=0.227 |

Figure 4.26: Proportion of people who got both genders right in a gender transform according to their races and genders for all users.

| Caucasian | | Indian | | Oriental | |
|---|---|---|---|---|---|
| Male | Female | Male | Female | Male | Female |
| 11/14=0.786 | 11/41=0.268 | 9/10=0.9 | 0/4=0 | 8/22=0.364 | 3/10=0.3 |

Figure 4.27: Proportion of people who got both genders right in a gender transform according to their races and genders for consistent users.

From the four figures above, we can see that the gender transform works better for males that females for all race groups and for individuals who belong to Age1. The face that this transform works best for males of all races compared to females is probably the reason why this transform works better for males than females.

In the following figures, we get the proportion of all users and consistent users who chose each of the options to the age question when the original image went through the gender and race transformation separately. The images only went through one of these transforms at any one time. These two groups of participants are categorised according to the age group they belong to. The question that asks users to compare the ages of the original and the transformed image reads ”Which face looks older?”. The first option is ”The face on the right” and the second option is ”The face on the left” and the last option is ”Both of them are of the same age”. In the figures below, the letter R represents the the first option, while the letters L and S represent the second and the third option respectively.

| Age1 | | | Age2 | | | Age3 | | |
|---|---|---|---|---|---|---|---|---|
| R | L | S | R | L | S | R | L | S |
| 45/150 =0.3 | 69/150 =0.46 | 36/150 =0.24 | 33/90 =0.367 | 37/90 =0.411 | 19/90 =0.211 | 5/60 =0.0833 | 52/60 =0.867 | 3/60 =0.05 |

Figure 4.28: Proportion of people who chose each of the options among all users when the gender was transformed

| Age1 | | | Age2 | | | Age3 | | |
|---|---|---|---|---|---|---|---|---|
| R | L | S | R | L | S | R | L | S |
| 11/38 =0.289 | 20/38 =0.526 | 7/38 =0.184 | 10/34 =0.294 | 16/34 =0.471 | 8/34 =0.235 | 1/35 =0.0286 | 34/35 =0.971 | 0/35 =0 |

Figure 4.29: Proportion of people who chose each of the options among consistent users when the gender was transformed

| Age1 | | | Age2 | | | Age3 | | |
|---|---|---|---|---|---|---|---|---|
| R | L | S | R | L | S | R | L | S |
| 97/300 =0.323 | 128/300 =0.427 | 75/300 =0.25 | 68/180 =0.377 | 77/180 =0.428 | 33/180 =0.183 | 12/120 =0.1 | 105/120 =0.875 | 3/120 =0.025 |

Figure 4.30: Proportion of people who chose each of the options among all users when the race was transformed

| Age1 | | | Age2 | | | Age3 | | |
|---|---|---|---|---|---|---|---|---|
| R | L | S | R | L | S | R | L | S |
| 16/76 =0.211 | 35/76 =0.461 | 25/76 =0.329 | 23/68 =0.338 | 32/68 =0.471 | 13/68 =0.191 | 3/70 =0.0429 | 67/70 =0.957 | 0/70 =0 |

Figure 4.31: Proportion of people who chose each of the options among consistent users when the race was transformed

Figure 4.28 to figure 4.31 show that most of the participants thought that the original image or the image on the left was older even when the transformation did not involve altering the age. This implies that the algorithm is making the test faces look younger. This may be because the wrinkles and other facial blemishes in the test faces may be smoothed out in the first few steps of the MMDA process before the mask was placed on the test images. These steps may have altered the shape of the faces and the features and this may be the reason as to why the gender transformation works better on male faces than on female faces.

# Chapter 5

# Conclusion and Future Work

MMDA is a method that is able to decompose an image into numerous modes that are independent of other modes. In this project, MMDA was used to decompose human faces into different modes like gender, face,race and identity. The performance of this algorithm was evaluated by the use of the survey. It was found out that the algorithm is good at preserving and changing the race, preserving the gender and altering the ages of the test faces.

More work needs to be done to find out why this method is not able to alter the gender and preserve the ages of the original image well. Analysing how the first few steps in the MMDA process alter the original image may give some clues as to what needs to be done in order to improve the algorithm's performance in these areas.

# Appendix A

# References


Buchala, Davey, Gale, and Frank (2005),Principal Component Analysis of Gender, Ethnicity, Age and Identity of Face Images, Department of Computer Science, University of Hertfordshire, College Lane, Hatfield, AL10 9AB, UK and Department of Psychiatry, QEII Hospital, Welwyn Garden City, AL7 4HQ, UK

Costen, Cootes, Edwards and Taylor (2002), Automatic extraction of the face identity subspace, Department of Medical Biophysics, Imaging Science and Biomedical Engineering, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT,UK,Image and Vision Computing 20 (2002) 319-329

E. Newton, L. Sweeney, and B. Malin. Preserving Privacy by De-identifying Facial Images,Carnegie Mellon University, School of Computer Science, Technical Report, CMU-CS-03-119 Pittsburgh: March 2003.

Lombaert, H. (2006, June 19). Manual Registration with Thin Plates. In Herve Lombaert. Retrieved October 29, 2013, from Thin plate spline website: http://step.polymtl.ca/ rv101/thinplates/

Mrityunjay, P. J. Narayanan. The De-identification camera, International Institution of Information Technology Hyderabad, 2011 Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics

Prachi Agrawal and P. J. Narayanan. Person De-Identification in videos, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 21, NO. 3, MARCH 2011

Stallwood, O. (2013, January 25). Game to destroy CCTV cameras: vandalism or valid protest? In The Guardian. Retrieved November 28, 2013, from The Guardian website: http://www.theguardian.com/theguardian/shortcuts/2013/jan/25/game-destroy-

T.Sim, S. Zhang, J. Li and Y. Chen. Simultaneous and Orthogonal Decomposition of Data using Multimodal Discriminant Analysis. In IEEE 12th International Conference on Computer Vision, September 2009.

Xiaoyi Yu, Kenta Chinomi, Takashi Koshimizu, Naoko Nitta, Yoshimichi Ito, and Noboru Babaguchi. PRIVACY PROTECTING VISUAL PROCESSING FOR SECURE VIDEO SURVEILLANCE. Graduate School of Engineering, Osaka University, Japan

# Appendix B

# Code

A1

```
[a,b]      = size (X)

minval  = min (X)

maxval = max(X)

Y=ceil(((X - minval*ones(a,b))./(maxval*(ones(a,b))-minval*(ones(a,b))))*255);
```

A2

```
function  [img tformAlign] = cropAndAlign(im,left_eye,right_eye)

eyepos = [-1,0;1,0];

xrange = [-3,3];

yrange = [-2.5,4.5];

scale = (xrange(2)-xrange(1))/127;

tformAlign = cp2tform( [left_eye(1),left_eye(2);right_eye(1),right_eye(2)],eyepos,'n

img = imtransform( im,tformAlign,'bilinear', 'XData',xrange,'YData',yrange,'XYScale
```

B1

```
[ D, N ] = size( obs );

[ M, NN ] = size( labels );

PCA_model = PCA_make( double(obs));

m.org = PCA_model.org;

m.eigVal = PCA_model.val * N;

m.eigVct = PCA_model.vct;

X = diag(m.eigVal.^(-0.5)) * PCA_project( double(obs), PCA_model );
```

Where:

Helper function PCA_make is defined as following:

function m = PCA_make( obs) is used to make an eigenspace model for observations wi

```
% pre-condition:

% obs: an (nxN) matrix; N observations, each id n dimensional.

% post-condition:

% m: an eigenmodel; a structure comprising

% N: the number of observations input to make the model.

% org: the mean of the observations.

% vct: a (n x p) matrix of eigenvectors (columns) p is decided

% val: a (p x 1) matrix (column vector) of eigenvalues.
```

Helper function PCA_project (obs, m) is used to project obs onto each discriminant s

given an MMDA model m.

B2

```
[PCA_model, XX, m] = MMDA_make( obs_landmark, labels );

function [PCA_model, XX , m]= MMDA_make(obs,labels)

[ D, N ] = size( obs );
[ M, NN ] = size( labels );

%D is the number of dimentions , N is the number of people
%M is the number of modes , NN is the number of people

m.dim = D;
m.num_data = N;
m.num_mode = M;

% 1. Make Eigen model
PCA_model = PCA_make( double(obs), 'keept', 1e-4 );
m.org = PCA_model.org;
m.eigVal = PCA_model.val * N;
m.eigVct = PCA_model.vct;
X = diag(m.eigVal.^(-0.5)) * PCA_project( double(obs), PCA_model );
m.P = m.eigVct * diag(m.eigVal.^(-0.5));
m.Pr = m.eigVct * diag(m.eigVal.^(0.5));
XX = X;
%% The XX here is the whitened data. (Only principal components of data.)
m.disVct = cell(M,1);
```

```
m.disVal = cell(M,1);
m.lowObs = cell(M,1);
m.C = [];
% Loop for each mode
for kk = 1 : M
    label = labels( kk, : );
    Hb = [];
    C = max( label ); %% finding maximum no of classes in the mode
    m.C = [ m.C C ];
    % 2. Build Hb for each mode
    K = [];
    for ii = 1 : C
        [ row, col ] = find( label==ii );
        k = length( col );
        Xmu = mean( X(:, col), 2 );
        Hb = [ Hb k^0.5 * Xmu ]; %%summation of Hb
        K = [ K 1/(k^0.5) ];     %%summation of Hb
    end

% 3. Make V1 for each mode
    [ Vct, Val ] = MMDA_eig( Hb );
    m.disVct{kk} = Vct;
    m.disVal{kk} = Val;


   % 4. Project means onto low dim space
```

```
    m.lowObs{kk} = Vct'*Hb*diag(K);  % C-1 X C

    clear Hb Vct Val;

end


% 5. Compute the residual space

for kk = 1 : M

    U = m.disVct{kk};

    X = X - U * U' * X;

    clear U;

end

[ m.resVct, m.resVal ] = MMDA_eig( X );


% 6. Compute Q matrix, which combines all V1

Q = [];

for kk = 1 : M

    Q = [Q, m.disVct{kk}];

end

m.Q = [Q, m.resVct];

return;


Where:

Helper function PCA_make is defined as following:

function m = PCA_make( obs) is used to make an eigenspace model for observations wi

% pre-condition:

% obs: an (nxN) matrix; N observations, each id n dimensional.
```

```
% post-condition:
% m: an eigenmodel; a structure comprising
% N: the number of observations input to make the model.
% org: the mean of the observations.
% vct: a (n x p) matrix of eigenvectors (columns) p is decided
% val: a (p x 1) matrix (column vector) of eigenvalues.
%lowObs : a 3*1 cell which contains the vectors which represents the different clas
                    mode
```

Helper function PCA_project (obs, m) is used to project obs onto each discriminant s given an MMDA model m.

[Appendix B3]

```
for ageno=2:3
for magnitude=1:3
for person_no=1:10

X(:,1) = obs_landmark (:,person_no);
t(:,1) = m.Q' * m.P' * (X - m.org);

if (ageno==1)
    t(4,1) = m.lowObs{3,1}(1,1)*magnitude;
    t(5,1) = m.lowObs{3,1}(2,1)*magnitude;
elseif (ageno==2)
    t(4,1) = m.lowObs{3,1}(1,2)*magnitude;
```

```
    t(5,1) = m.lowObs{3,1}(2,2)*magnitude;
elseif (ageno==3)
    t(4,1) = m.lowObs{3,1}(1,3)*magnitude;
    t(5,1) = m.lowObs{3,1}(2,3)*magnitude;
end

Xr = m.Pr * m.Q * t + m.org;
normalise_Xr = Xr(1:15616,1);
result=ScaleImage( normalise_Xr, 1 );
result=result*255;
RecImage = reshape (result (1:15616,1),122,128);
%RecImage = reshape (Xr(1:123510,person_no),230,179,3);
RecImage = round (RecImage);
G = RecImage;
G=uint8(RecImage);
%imshow (G)
imwrite(G,'MMDA_demo_result.bmp','bmp');
image = imread('MMDA_demo_result.bmp');
cd ../
imwrite(G,'MMDA_demo_result.bmp','bmp');
newface_landmarksxy =zeros(71,2);
counter=15617;
%newlm = Xr(15617:15770)
for points=1:71
    xval = Xr(counter,1);
```

```
    yval = Xr(counter+1,1);

    counter=counter+2;

    newface_landmarksxy(points,:) =[xval,yval];

end



cd ThinPlateSplice

[wimg, w, a] = TPS_im_warp(G,newface_landmarksxy,ref,person_no);%prev used

imshow(wimg)
```