Gender Classification using Face Recognition

Terishka Bissoon
School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal
Durban, South Africa
Email: terishkabissoon@gmail.com

Serestina Viriri
School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal
Durban, South Africa
Email: viriris@ukzn.ac.za

Abstract— **This paper addresses the issue of gender classification using the method of Principal Component Analysis (PCA) for face recognition and classification of human faces. The use of the PCA algorithm has a maximum success rate of 82%. The gender classification system is then improved by using the Linear Discriminant Analysis (LDA. This algorithm has a machine-learning framework by which it trains on a database and using this trained environment to predict the outcome of other images. The classification is restricted to two classes – male and female. Upon using LDA, the success rate increased to approximately 85%. The database used in this paper for the training and testing of images is called the FERET database.**

Keywords—gender classification, PCA, LDA

I. INTRODUCTION

Gender classification is a binary classification problem, in which one has to predict an image as that of a man or woman [1]. Human beings need to classify gender in order to successfully interact with each other and it proves relatively easy for a human, but it still proves challenging for a computer.

Gender classification has gained a lot of interest in the machine learning and computer vision community. This is because of its importance in the use of Human Computer Interaction technology in order to provide people with a more secure, reliable and convenient way to approach computer vision techniques such as face identification and gender classification [2, 3]. Gender classification could be useful in a number of applications, such as for biometric authentication, high-tech surveillance and security systems, criminology, automatic psycho physiologic inspection, augmented reality etc.

Visual information from human faces provides one of the more important sources of information for gender classification. The aim of this research work is to investigate alternative gender classification techniques using facial features.

Previous studies have broadly classified facial feature extraction in two categories:

- appearance feature-based (global) and
- geometric feature-based (local).

The former finds the decision boundary directly from an image, while the latter is based on geometric features such as eyebrow thickness, nose width etc. [2].

The goal of this project is to determine whether features extracted from computer vision techniques for face recognition can be extended to gender classification and the main aim of this project is to investigate gender classification using learning algorithms that can improve the accuracy rate for gender classification. The computer vision techniques used in this paper include Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).

Figure 1. Example of images from the FERET database

II. LITERATURE REVIEW

There has already been much work done in this field of study. Many people have made significant improvements in this area.

Xu et al. [2] proposed a hybrid approach to gender classification that consists of three modules. The first module normalizes a given face image, where some geometry alignment and gray level normalization is processed in this stage. The second module extracts features from the normalized image to form a feature vector. The third module uses the feature vector as the input, and the output of the module reaches the conclusion. Xu used a Support Vector Machine (SVM) with the radial basis function (RBF) kernel as the classifier. The appearance features are chosen mathematically by using the AdaBoost algorithm and the geometry features are extracted using Active Appearance Model (AAM). These methods achieved a success rate of 92.38% overall for male and female gender classification. Xu states that although the success rate is high, certain aspects still deserve further study. The time consumed for location and feature extraction can be decreased if a further optimization is made [2].

Mannan [3] presents gender classification based on facial images using dimensionality reduction techniques such as

Principal Component Analysis (PCA) and Independent Component Analysis (ICA) along with Support Vector Machine (SVM). The input dataset is divided into training and testing dataset and experiments are performed by varying dataset size. The effect of performing image intensity normalization, histogram equalization, and input scaling are observed.
Using a larger dataset obtained better results, as both PCA and ICA achieved accuracy rates of over 90%. According to Mannan, input scaling does not seem to improve the performance of PCA especially when image intensity or histogram equalization is performed.

Ramesha et al. [4] proposed a Feature Extraction based face Recognition, Gender and Age Classification (FEBFRGAC) algorithm with only small training sets and it yielded good results. Canny edge operator was used and based on the texture and shape information gender classification is done using Posteriori Class Probability with a success rate of around 98%.

III. METHODOLOGY

The gender classification framework consists of a series of phases that helps to produce the final output. This output is a simple result determining whether the image is of male or female gender.

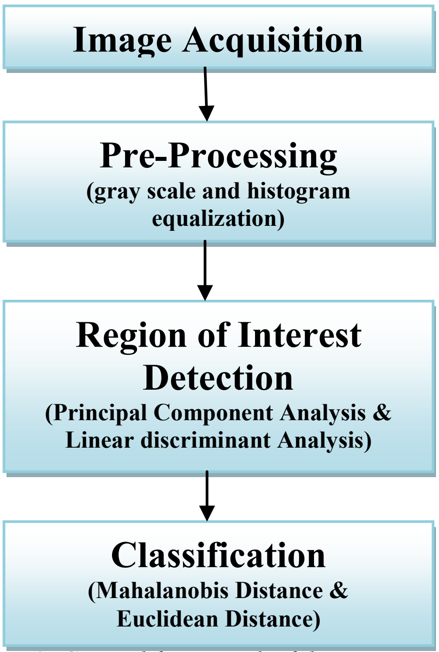

Figure 2. General framework of the system

A. Preprocessing

The preprocessing techniques that were used in this paper consisted of converting the color image to a gray scale image, and then applying histogram equalization on the image. Histogram equalization is used to increase the global contrast on the images. The implementation for histogram equalization is as follows:

Let the grayscale image be {x} and n_i be the number of occurrences of gray level *i*. The probability of an occurrence of a pixel of level *i* in the image is

$$p_x(i) = p(x = i) = \frac{n_i}{n}, 0 \leq i < L,$$

With L being the total number of gray levels in the image, n being the total number of pixels in the image, and $p_x(i)$ being the image's histogram for pixel value *i*, normalized to [0,1].

The cumulative distribution function corresponding to p_x is defined as:

$$cdf_x(i) = \sum_{j=0}^{i} p_x(j)$$

which is also the image's accumulated normalization histogram. A transformation of the form y = T(x) produces a new image {y}, such that its CDF will be linearized across the value range, i.e.

$$cdf_i = iK$$

for some constant K. The new image is defined as

$$y = T(x) = cdf_x(x)$$

The following transformation is applied on the result in order to map the values back into their original range

$$y' = y \cdot \left(\max\{x\} - \min\{x\}\right) + \min\{x\}$$

Figure 3. Example images after applying Histogram Equalization

B. Region of Interest Detection

Principal Component Analysis

Principal Components Analysis is a very well-known method of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences [12].

In order to apply PCA to images the image need to be represented as a column of vectors. A matrix is then formed by combining the column of training set images.

Let this matrix be X:

$X = [x_1 x_2 x_3 \dots x_n]$,where x_i is the i^{th} column vector representing the i^{th} training image.

The mean is calculated and subtracted from each column and the resulting covariance matrix is computed.

Let the mean image be μ:

$$\frac{1}{n}\sum_{i=1}^{n} x_i$$

With $Y = [x_1 - \mu \dots x_n - \mu]$ and the covariance matrix

978-1-4799-3067-8/13/$31.00 ©2013 IEEE

$$Q = cov(Y) == YY^T$$

Once we have the above data calculated we need to find the eigenvectors (a square non-zero matrix that when multiplied by the matrix yields a vector that is parallel to the original). These vectors are known as the principal components. Out of these eigenvectors *m* most significant vectors are chosen [let these vectors be: $e_1, e_2, e_3, \ldots, e_m$]. The value of *m* is the cumulative sum of the eigenvectors.

The features of an image *x* is then computed by projecting it onto the space spanned by the eigenvectors as follows

$$t = [e_1 e_2 \ldots e_m]^T (x - \mu),$$

where t is an *m* dimensional vector of features.

This feature vector *t* is used during training and classification.

Linear Discrimination Analysis

Linear Discriminant Analysis (LDA) is a statistical method which projects the given multidimensional data to a lower dimension [13]. LDA finds a combination of features which characterizes two or classes of objects. In this paper, the two classes are male or female. LDA is closely related to PCA such that they both look for the best linear combinations of variables to explain the data.

Let x_k denote the data matrix of values of an image of *n* number of pixels, where *k* is the number of training images. The pooled within group covariance is defined by,

$$C = \frac{1}{n}\sum_{i=1}^{g} n_i c_i \quad (1)$$

such that *g* is the number of groups (i.e. classes) and c_i is the covariance of group *i* shown in

$$C_i = \frac{(x_i{}^0)^T x_i{}^0}{n_i} \quad (2)$$

where x_i^0 is the corrected mean of the data, $x_i - \mu$ and μ is the global mean of the data.

The above equations are used to train the images in the system.

C. Classification

Principal Component Analysis

For the PCA algorithm, the Euclidean distance was used for classification of images. Euclidean distance is defined as the distance between two points that one would measure with a ruler. The distance between the test image and a training image is shown as

$$distance(x,y) = \sum_{i=1}^{k} (x_i - y_i)^2$$

Principal Component Analysis

After the pooled within group covariance value is calculated, take the inverse of it to calculate the discriminant function values for the different groups. The discriminant function values are based on the Mahalanobis distance. The Mahalanobis distance takes into account the covariance among the variables in calculating the distance, calculating the dissimilarity between groups.

The discriminant function is calculated as follows:

$$f_i = \mu_i C^{-1} z^T - \frac{1}{2}\mu_i{}^T + \ln(p_i),$$

and the function f_i is used for prediction:

$$f_i = d^T C^{-1} d,$$

Where $d = z - \mu_i$

IV. RESULTS AND DISCUSSIONS

The images used for the training and testing of this paper were taken from the FERET database [1]. A total of 600 images were used in this project. The gender classification system was tested 3 times with 200 images each time. For each round 100 images were used for training and 100 images were used for testing. Each set of training images contained equivalent numbers of male and female images.

The gender was classified into two classes: male and female. Once the training was complete, the system was able to accept a test image to test the system. Eigenvectors of this test image were calculated according to above and was compared to the training dataset. The results from the training sets are used as input to the testing set.

Figure4. Female images that were correctly classified

Figure5 . Male images that were correctly classified

Round 1: The system had an initial success rate of 82% for face detection with PCA, which then increased to 85% with the LDA algorithm for classification.

Round 2: The system had an initial success rate of 78% for face detection with PCA, which then increased to 84% with the LDA algorithm for classification.

Round 3: The system had an initial success rate of 83% for face detection with PCA, which then increased to 85% with the LDA algorithm for classification.

The following bar graph is a visual representation of the results achieved in the gender classification system.

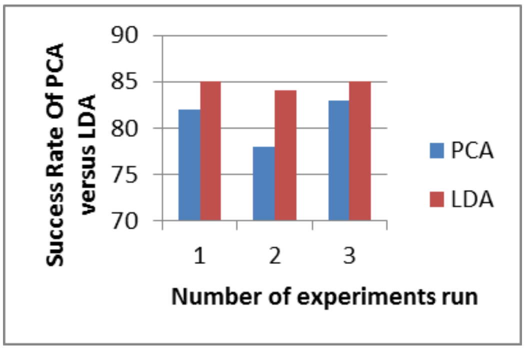

Figure 5. Results of PCA with LDA

As can be seen from the results, the results were improved when the LDA algorithm was used after the PCA algorithm. The explanation for these results could be because LDA models the differences between classes.

V. Conclusion

This paper has used principal components analysis with linear discriminant analysis to classify gender. The LDA based system was implemented for the gender classification between two classes. The results achieved seem to correlate with similar works discussed in the Literature Review.

The gender classification system achieved consistent results, with differences between different sets of data a few percent or less.

The most successful round of testing included a success rate of 83% when the PCA algorithm was applied, with an overall success rate of 85% reached when the LDA algorithm was applied to the system.

Acknowledgment

The authors would like to give thanks to the Facial Recognition Technology (FERET) Project for the use of their database in this paper.

References

[1] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, The FERET evaluation methodology for face recognition algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, October 2000.

[2] Ziyi Xu, Li Lu and Pengfei Shi, A Hybrid Approach to Gender Classification from Face Images

[3] Fahim Mannan, Classification of Face Images Based on Gender using Dimensionality Reduction Techniques and SVM

[4] Ramesha K, K B Raja, Venugopal K R and L M Patnaik, Feature Extraction based Face Recognition, Gender and Age Classification, 2010.

[5] Philips, P. J., H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithm", Image and Vision Computing, 1998, vol. 16, no. 5, pp. 295-306.

[6] Wikipedia < http://en.wikipedia.org/wiki/Principal_component_analysis > (accessed on 08 April 2012)

[7] Viola P., Jones M. J. "Robust Real-Time Face Detection", International Journal of Computer Vision , 2004, 57(2) pp: 137-154

[8] M. D. Malkauthekar and S. D. Sapkal, "Experimental Analysis of Classification of Facial Images," IEEE International Advance Computing Conference, pp.1093-1098, 6-7 March 2009.

[9] Arnulf B. A. Graf and Felix A. Wichmann, Gender Classification of Human Faces

[10] E. Makinen, R. Raisamo, "An experimental comparison of gender classification methods", Pattern Recognition Letters 29 (2008) 1544-1556

[11] M. Castrillon-Santana and Q. C. Vuong, "An Analysis of Automatic Gender Classification"

[12] Lindsay I Smith, "A tutorial on Principal Components Analysis", 26 February 2002

[13] Xiaoguang Lu and Anil K. Jain , Ethnicity Identification from Face Images, Proceedings of SPIE International Symposium on Defense and Security: Biometric Technology for Human Identification, 2004

[14] K. Fukunaga, Inroduction to Statistical Patterm Recognition, Academic Press, San Diego, California, 1990