Abstract—This paper presents an experimental study on examining the effects of facial and racial features on gender classification. In order to show which facial feature is the most influential for gender classification, parts of several face images, such as, forehead, eyebrows, eyes, nose, lip and chin were masked. For dimension reduction, Principal Component Analysis (PCA) and for determination of gender, Fisher Linear Discriminant (FLD) algorithms were applied to masked face images. Moreover, the effects of racial features on gender classification were studied. Experimental results indicated that the nose is the most influential part for gender classification. Furthermore, the gender of the Asian people is more easily distinguished than that of the people of African origin.

I. INTRODUCTION

Pattern recognition is a research area gradually gaining popularity and importance. It is used for (i) converting hand writing to text, (ii) recognizing human voice, (iii) processing medical data, (iv) recognizing signature and (v) identification from finger print or iris. One of the usage areas of the pattern recognition is gender classification. Mostly, to classify the gender, face images are used, because people can easily cover their bodies with clothes. So determining the gender using body images can be difficult. Determination of features, such as, gender, race, expression of a person, using facial features goes back to 1960 [1-2]. Because of the dynamic structure of the face, this problem has not been completely solved yet. Good results can be obtained using face images with an appropriate pose, from an appropriate distance, with appropriate illumination and face expression.

Security is one of the most important problems nowadays, so recognition of a person using face images has become an interesting research area. Although a lot of research has been done on face recognition, there are very few works on gender classification using face images. There are various usage areas of gender recognition. Primarily gender recognition is used for identification of human faces. For example, a substantial amount of data on a large database may be dramatically reduced using gender classification. It may additionally be used to determine customer profile in shopping centres. Some products may particularly be consumed by men than women. By determining customer profiles and adjusting marketing strategies accordingly may be profitable for both shopping centres and producers.

The first study was done by Jain and Huang in 1991 using artificial neural networks [3]. Using artificial neural networks another study was presented by Golomb, Lawrence and Sejnowski in 1991 [4]. The error rate in this study was 8.1%. Kawano, Kato and Yamamoto separated facial images into special areas. Subsequently, they applied Linear Discriminant Analysis (LDA) and Four Directional Feature Fields (FDF) algorithms to these areas for gender classification. 93.7% accuracy was obtained [5]. Ueki et al. presented a method for gender classification by integrating facial, hairstyle and clothing images. For this purpose the Gaussian Mixture Model (GMM) algorithm was used [6]. Another remarkable work using Support Vector Machine (SVM) was done by Moghaddam and Yang [7]. All of the above works were based on appearance. Furthermore, there are geometric based methods. Burton, Bruce and Dench have determined 73 special points over entire face images. Applying the discriminant analysis, they used the distances between these points for gender classification [8]. Brunelli and Poggio have obtained 16 geometric features over entire face images for gender classification [9]. The appearance based methods are more successful than the geometric based methods. The geometric based methods are limited for applications in gender classification. Therefore, we used PCA and FLD algorithms in our work presented in this paper. K-fold cross validation test method was used for determining the accuracy rate of the above mentioned methods. Experimental results yielded an accuracy rate of 93%. An experimental study on the effects of facial and racial features on gender classification was presented using data from FERET [10-11] and Stanford University databases [12].

II. ALGORITHMS AND TEST METHOD

PCA is one of the most effective methods in data compression and pattern recognition. The aim of the PCA is to reduce the dimension of the data. PCA is used to omit redundant data for feature extraction, data compression and prediction. Since the PCA works in linear domain, it is used in linear applications, such as, signal processing, image
processing, system and control theory and communication. PCA depends on eigenvector method designed to model linear variations in large dimensional data. PCA performs dimension reduction by projecting the original n-dimensional data onto k-dimensional (k<n) linear subspace [13]. Let \( D = \{ x_1, x_2, \ldots, x_n \} \) be samples, each \( x_i \) be \( d \)-dimensional vector and \( \mu \) be the mean of all samples. For reducing dimension from \( n \) to \( k \) using PCA, eigenvectors \( e_1, e_2, \ldots, e_k \) corresponding to the \( k \) largest eigenvalues of scatter matrix \[
S = \sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T
\] (1)
are found. \( E = [e_1, e_2, \ldots, e_k] \) is the best direction to project the data on [14]. PCA tries to find a set of orthogonal basis functions that capture the directions of maximum variance in data, such that,
\[
J_{PCA}(W_{opt}) = \arg \max \frac{W^T S_T W}{W^T S_B W} \quad (2)
\]
is maximized. Here, \( S_T \) is the total scatter matrix of the training set samples. \( W \) is the matrix whose columns are the orthonormal projection vectors [15]. In order to show which facial feature is the most influential for gender classification, several face images were divided into parts, such as, forehead, eyes, nose, lip, chin and were masked. For dimension reduction Principal Component Analysis (PCA) was applied to the masked face images.

The aim of the FLD is to find the optimal linear projection for classification. Original class is not important in PCA. Yet, in FLD, the samples belong to a certain class and this information is used. FLD is used for dimension reduction and for classification. PCA tries to find a set of orthogonal basis functions that capture the directions of maximum variance in data. Unlike PCA which encodes information in an orthogonal linear space, FLD encodes discriminating information in a linearly separable space using bases that are not necessarily orthogonal [16]. Consider a \( c \)-class problem, with the between-class scatter matrix given by
\[
S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (3)
\]
and the within-class scatter matrix given by
\[
S_W = \sum_{i=1}^{c} \sum_{x_i \in x_i} (x_i - \mu_i)(x_i - \mu_i)^T \quad (4)
\]
where \( \mu_i \) is the mean of all samples, \( \mu_i \) is the mean of class \( i \), and \( N_i \) is the number of samples in class \( i \). The optimal projection \( (W_{opt}) \) is the projection matrix which maximizes the ratio of the determinant of the between-class scatter to the determinant of the within-class scatter of the projections
\[
J_{FLD}(W_{opt}) = \arg \max_W \frac{W^T S_B W}{W^T S_W W} = [w_1, w_2, \ldots, w_m] \quad (5)
\]
where \( w_i \) is the set of generalized eigenvectors of \( S_B \) and \( S_W \) corresponding to the \( m \) largest generalized eigenvalues [15].

K-fold cross validation test is used for determining the accuracy rate of classification algorithms. There are generally two sets called test and training set. The training set is used for training the classifier, the test set is then used for finding error rate of the previously trained classifier. In this method data is divided in \( k \) equal parts. One of the \( k \) subsets is used each time as the test set and the other \( k-1 \) subsets are put together to form a training set. According to this method the test is performed \( k \) times. In this method every sample is used in the training set and the test set. Afterwards the average error \( (\bar{E}) \) across all \( k \) trials is computed.
\[
\bar{E} = \frac{1}{k} \sum_{i=1}^{k} E_i \quad (6)
\]

Generally \( k \) is chosen 10 for tests. Thus, we tested our system by using ten-fold cross validation [17]. A schematic diagram of gender classification system is shown in Fig. 1. According to this figure, firstly face is detected and then PCA and FLD algorithms are applied, lastly gender is determined.

![Fig. 1. A schematic diagram of gender classification system](image)

### III. EXPERIMENTAL RESULTS

In order to show the effects of facial and racial features on gender classification, experiments were done on a database consisting of 480 face images including 240 male and 240 female face images that belong to Stanford and FERET databases. All grayscale front view images were chosen for testing. To detect the face, a program called facedetect of the OpenCV (Open Source Computer Vision) library was applied to all images [18]. So we obtained face images in different dimensions. Then these images were resized to be 128x128 pixels.

Face images to be processed may not include all parts of the face. For example, parts of the face may be blurred/faded or may be missing. An image may still be useful in mentioned conditions. Therefore to determine which part of the face is more influential for gender classification we determined 6 patterns such as forehead, eyebrows, eyes, nose, lips, and chin. We masked some parts of the face images as shown some examples in Figs.2 and 3. To mask different parts of faces, a particular set of pixels of images were made black. Initially upper diagonal triangle of face and subsequently lower diagonal triangle of face were masked and tested as shown in Fig.2(b) and (c), respectively. As can be seen in Fig.2 (b) the image of the masked upper diagonal triangle of the face includes the right eye, the lower part of nose, the lips and the chin. In Fig.2 (c) the image of the masked lower diagonal triangle of the face includes the left eye, the upper part of the nose and the forehead.
The face images were also masked separately as shown in Fig. 3 (forehead, eyebrows, eyes, nose, lips and chin). These masked pictures were then tested.

![Fig. 2. Test pictures (a) Original pictures. Masked are: (b) upper diagonal triangle, (c) lower diagonal triangle, (d) forehead and eyes, (e) lips and chin.]

![Fig. 3. Pictures of some masked parts of face: (a) forehead, (b) eyebrows, (c) eyes, (d) nose, (e) mouth, and (f) chin.]

Results of these tests were presented in TABLE I. Results showed that for gender classification, upper diagonal triangle of the face is more influential than the lower diagonal triangle. To reinforce these results, (i) the forehead and the eyes together, and then, (ii) the lips and the chin together were masked as shown in Fig.2 (d) and (e), respectively. Those masked pictures were tested to determine which parts of the face are effective for gender classification. Results of these tests were presented in TABLE I. Results indicated that forehead and eyes are more influential than lips and chin in gender classification.

![TABLE I

SIGNIFICANCE OF FEATURES IN GENDER CLASSIFICATION: DIAGONAL UPPERSIDE AND DIAGONAL LOWERSIDE OF THE FACE

<table>
<thead>
<tr>
<th></th>
<th>Lower diagonal triangle masked</th>
<th>Upper diagonal triangle masked</th>
<th>Forehead &amp; eyes masked</th>
<th>Lips &amp; chin masked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Rate %</td>
<td>26.75</td>
<td>47.50</td>
<td>27.50</td>
<td>7.50</td>
</tr>
</tbody>
</table>

The results were presented in TABLE II. As you can see in table, the error rate is much more when the nose was masked. The second influential part is forehead and the least influential part is chin. Those results verified our above reported results.

![TABLE II

EFFECTS OF MASKING IN GENDER CLASSIFICATION

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Rate %</td>
<td>35.50</td>
<td>10.75</td>
<td>7.50</td>
<td>45.75</td>
<td>8.75</td>
<td>6.75</td>
</tr>
</tbody>
</table>
In addition to these experiments, we tested the influence of racial features on gender classification. We used Asian and African people’s pictures as shown in Fig. 4 as the test set. While the training set was European and Asian people pictures, the test set was only African people pictures. Similarly, while the training set was European and African people’s pictures the test set was only Asian people pictures. In this way we tested the images belonging to different races.

Finally, we presented the effect of racial features on gender classification. To the best of our knowledge the effects of racial features have not been studied in the literature. The results of our research indicated that in gender classification racial features have not been studied in the literature. The classification. To the best of our knowledge the effects of racial features on gender classification. We used Asian and European people’s pictures the test set was only Asian people pictures.

In addition, the effect of race on the gender classification was tested. The results indicated that, the gender of the Asian people is easily classifiable than the gender of the people of African origin.

ACKNOWLEDGMENT

For parts of the research presented in this paper we used the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office.

Our thanks also go to Dr. Inci Cilesiz for revising the final manuscript.

REFERENCES


