Gender Classification using Single Frontal Image per Person: Combination of Appearance and Geometric based Features

Saeed Mozaffari, Hamid Behravan
ECE Department, Semnan University
Semnan, Iran.
mozaffari@semnan.ac.ir

Rohollah Akbari
Islamic Azad University, Qazvin Branch
Qazvin, Iran.

Abstract—Today, many social interactions and services depend on gender. In this paper, we introduce a single image gender classification algorithm using combination of appearance-based and geometric-based features. These include Discrete Cosine Transform (DCT), and Local Binary Pattern (LBP), and geometrical distance feature (GDF). The novel feature, GDF proposed in this paper, is inspired from physiological differences between male and female faces. Combination of appearance-based features (DCT and LBP) with geometric-based feature (GDF) leads to higher gender classification accuracy. Our system estimates gender of the input image based on the majority rule. If the results of DCT and LBP features are not identical, gender classification will be based on GDF feature. The proposed method was evaluated on two databases: AR and ethnic. Experimental results show that the novel geometric feature improves the gender classification accuracy by 13%.

Keywords-component: Gender classification, sex recognition, single image, appearance features, geometric features,

I. INTRODUCTION

Gender classification is an important task in today technological world with a variety of applications such as commercial profiling, surveillance purposes, monitoring applications, and human-computer interaction. Gender classification can also be used as a filtering mechanism to enhance performance of face recognition both in terms of speed and accuracy [1].

So far, different methods have been proposed for gender classification using gait [2], iris [3], and hand shape [4]. However, the majority of studies on gender classification are based on face information. A comparison study of different gender classification approaches using face information can be found in [5]. Recently, the problem of sex classification was studied on the faces with arbitrary viewpoints and under occlusion [6].

Most of these methods for gender classification are based on training processes using several samples for each person. They include Support Vector Machines, Neural Networks, and Principal Component Analysis [7]. Preparing multiple training image samples from different point of views or under different lightening conditions is usually difficult or even impossible in some applications like video surveillance or criminal applications. There is a trend towards single image per person face recognition and understanding.

Previous studies for facial gender classification can be classified into two categories [7]: geometric-based and appearance-based. In the former category, the distance between different points in the facial image is used, while the latter category is based on the pixel-values of the face image. Each of these approaches has its own advantages and disadvantages. Methods of the appearance-base category can be considered holistic, in which the feature vector is extracted from the whole face image. Such a representation has the advantage over geometric-based category that information is extracted from all regions of the face. A disadvantage of the holistic appearance-based representation, however, is its sensitivity to local appearance variations, which can be caused by variability in position, pose, expression and illumination. One solution to this problem, which we proposed in this paper, is to use general faces obtained by averaging over facial images of the same sex in the database. Comparing the input image to the male and female general images can considerably reduce local appearance variations.

In this paper we proposed a new method for gender classification which is based on single frontal image per person. To increase gender classification accuracy, our system utilizes two appearance-based features (DCT and LBP) and one geometric-based feature (GDF). The gender of the input image is based on the majority rule.

II. THE PROPOSED GENDER ESTIMATION METHOD

To be compatible with a single image per person face recognition system, the proposed algorithm uses a non-training procedure for gender estimation. This strategy tackles any inconveniences concerning training data collection or retraining process after new class (person) addition. Block diagram of the proposed algorithm is shown in Fig.1.

A. Feature extraction

Our single image gender classification algorithm is based on two appearance-based features and one geometric-based feature. The final decision is made by majority rule, which
means if results of DCT and LBP features are not identical, gender classification will be based on GDF feature. Details of the feature extraction methods are presented in this section.

1) **Discrete Cosine Transform:**

DCT is a well-known transformation technique used for image compression. Our idea is to use the most significant DCT coefficients to represent and classify the image. DCT transform for a 2D image, \( f(i,j) \), with \( N \) columns and \( M \) rows is defined as:

\[
F(U,V) = \left( \frac{2}{N} \right)^{\frac{1}{2}} \left( \frac{2}{N} \right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} A(i)A(j) f(i,j) \cos \left( \frac{\pi}{2N} (2i + 1) \right) \cos \left( \frac{\pi}{2M} (2j + 1) \right)
\]

(1)

In which \( A(i) = \frac{1}{\sqrt{2}} \) for \( i = 0 \), and 1 otherwise.

2) **Local Binary Pattern**

LBP feature is used frequently for gender classification [8]. LBP operator labels the pixels of an image by thresholding the \( 3 \times 3 \) neighborhood of each pixel with the center value and considering the result as a binary number:

\[
S(f_c - f_p) = \begin{cases} 
1 & f_c \geq f_p \\
0 & f_c < f_p 
\end{cases}
\]

(2)

Where \( f_c \) is the center value and \( f_p \) are the neighboring values around \( f_c \). Then the LBP value at the center pixel is:

\[
LBP(f_c) = \sum_{p=0}^{7} S(f_c - f_p)2^p
\]

(3)

The notation \( \text{LBP}^{p,R} \) is used for the uniform LBP operator [8] which means using the LBP operator in a neighborhood of \( P \) sampling points on a circle of radius \( R \). Figure 2 describes LBP operator for a \((8,1)\) neighborhood. Figure 3 shows the result of LBP for \( p=8 \) and \( R=1 \) on facial images.

3) **Geometric-based feature extraction**

Unlike appearance-based features (DCT and LBP) which use pixel-values of the whole image, geometric feature (GDF) is based on distances between some certain points of the facial image. As mentioned before, the advantage of the geometric-based method is its robustness in the presence of noise and changes in illumination.

For this type of feature, the connected components are determined by applying a region growing algorithm. Then, for each connected component with a given minimum size, the best-fit ellipse is computed using the properties of the geometric moments [9].

Fig.4 shows the result of ellipse fitting on some images of database. In this paper we use two types of information from fitted ellipse. The ratio between the major axes to the minor axes is used as the first component of the geometric-based feature. For the second one, RMS distance between ellipse and face contour is utilized.

The former criterion indicates the face elongation which is superior for females than males. According to physiological study, the female faces are more rounded and their cheeks tend to be fuller and more rounded than male ones. So, the second component of the proposed geometric-based feature has higher values for males.

B. **General face image construction**

As the proposed algorithm is not based on training process, a reference, which is called general image, is needed to determine gender of the input image. The general
male and female images are obtained by a linear combination of the same sex images (Fig.5).

C. Partitioning

Segmenting into areas allows extraction of facial features which are hypothesized to be generally distinct between male and females. In this paper, all images including the input image, image samples in the database, and general face images are partitioned in the same manner. According to Fig.6, each image is segmented into 6 partitions, each including at least one face feature (eye, eyebrow, nose, mouth). Regarding the original frontal image, we have 7 partitions.

D. Face features highlighting

All face features do not have the same importance in gender classification or face recognition. To highlight important face features, the input image is filtered with Sobel and Laplacian of Gaussian (LOG) operators. The two output images have different information that is combined by wavelet image fusion. Those face features like eyes and mouth which are more important for gender classification are intensified in the highlighted image. Fig.7 shows the result of face feature highlighting.

E. Weights calculation

We used image entropy to determine which partition of the highlighted image has higher weight in the gender classification process. Those partitions with higher entropy (weight) are considered as more important blocks.

F. Distance measurement

Some different distance criteria such as Euclidean (4), Manhattan (5), and Chebyshev (6) distances have been used in this paper to measure the distance between partitions in the input image and their corresponding in the general faces. By considering weight of each face partition, the input image is classified according to minimum distance. Experimental results show that Euclidean distance has the best accuracy.

\[
d(X, Y) = \|X - Y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}
\]

\[
d(X, Y) = \sum_{i=1}^{n} |x_i - y_i|
\]

\[
d(X, Y) = \max_{i=1}^{n} |x_i - y_i|
\]

III. EXPERIMENTAL RESULTS

To evaluate the proposed method, we used two facial databases: AR and ethnic. AR database consists of different facial images of 56 female and 70 male persons (total 126 persons) [10]. However in this paper, we only used neutral (frontal facial image without any expression) images of AR database. We also built an ethnic database (with the same distribution as AR) consists of facial images of Iranian persons. The most difference between AR and ethnic databases is the use of scarf (Hijab) for females. Comparing with AR, more men in the ethnic database have mustache and beard. Scarf and beard in ethnic database considerably separate male and female classes (Fig.8). This is mainly due to help of these factors for ellipse fitting during localization process.

As the first experiment, accuracy of appearance-based features is studied (Table.1). Fusion of DCT and LBP features yields accuracy of 84.6% and 80.3% on ethnic and AR databases, respectively. To explore robustness of the proposed system, general faces of one database is used for image classification of another database. As Table.1 shows, in this scenario, classification accuracy is degraded at most by 6%, which means our gender estimation system is not sensitive to image acquisition condition.

The aim of second experiment is to investigate the role of appearance-based and geometric-based features combination in the overall classification accuracy. GDF (geometric-based feature) has accuracy of 88.9% and 62.7% on ethnic and AR databases, respectively. The proposed geometric-based feature classifies ethnic male and female images with higher accuracy compared with AR database.
IV. CONCLUSION

We investigate the problem of gender classification using single frontal image per person. Two types of features have been used for this purpose. First, we utilized two different appearance-based features (DCT and LBP). Second, a new geometric feature, called GDF, is proposed based on physiological differences between male and female faces. This feature had very promising results on ethnic database. Since appearance-based and geometric-based features are intrinsically different, their combination yields very high accuracy for both ethnic and AR databases.

REFERENCES