GENDER CLASSIFICATION FROM FACE IMAGES WITH LOCAL TEXTURE PATTERN

YI-JUI LI, CHIH-CHIN LAI, CHIH-HUNG WU, SHING-TAI PAN, SHIE-JUE LEE

1,2,3Department of Electrical Engineering, National University of Kaohsiung, 811 Kaohsiung, Taiwan
3Department of Computer Science and Information Engineering, National University of Kaohsiung, 811 Kaohsiung, Taiwan
E-mail: steve lié0817@gmail.com, {clai, johnw, stpan}@nuk.edu.tw, leesj@mail.ee.nsysu.edu.tw

Abstract - Recognizing human gender automatically by a computer is a challenging problem. It has been attracting research attention due to its wide real-life applications. Gender classification can be viewed as an essential preprocessing step in face recognition. Because human faces contain a lot of really useful information, many approaches based on facial features have been investigated for gender classification. In this paper, we present a novel texture pattern as feature descriptor to identify the gender from the facial images. The classification is performed by using a support vector machine. Experimental results on the FERET database are provided to illustrate the proposed approach is an effective method, compared to other similar methods.

Keywords - Gender Classification, Local Texture Pattern, Support Vector Machine

I. INTRODUCTION

Automatic gender classification from face images has been receiving much research interest for many years due to its various practical and successful applications, such as security control, visual surveillance, and other face analysis tasks. Most of the existing methods for gender classification focused on the application of feature extraction and classification techniques on frontal face images [1].

The framework of a gender classification is similar to one of a facial expression recognition system [2]. It consists of three modules: face acquisition, facial feature extraction, and facial features classification, respectively. The feature extraction is crucial to the whole classification process. If inadequate features are used, even the best classifier could fail to achieve accurate recognition. Generally, there are two main approaches to extract facial features: geometric feature-based methods and appearance-based methods [3]. The geometric feature-based approaches require extra computation to localize different facial components. In addition, they usually require accurate and reliable facial feature detection and tracking, which are difficult to accommodate in many situations. In contrast to geometric features, the appearance-based methods consider a whole face image. These methods are simple to implement and work on the facial images directly to represent facial textures without explicitly extracting any geometrical features. Lu et al. [4] proposed a pixel-pattern-based texture feature for real-time gender recognition. This feature is insensitive to illumination and relatively time saving. They adopt Adaboost to select the most discriminate features subset and support vector machine is used for classification. The face images are acquired under controlled conditions in many studies on gender recognition. However, face images may be acquired in unconstrained conditions in real-life applications. Shan [5] considered gender recognition on real-life faces. He uses local binary patterns (LBP) to represent the facial features, and then adopts Adaboost to learn the most discriminative LBP features. Jabid et al. [6] presented an appearance-based texture descriptor, local directional pattern (LDP), to represent facial image for gender classification. The LDP features are computed from the edge response values which encode the local texture of facial appearance. Experimental results show that their method can achieve higher accuracy on the images collected from FERET face database. In order to overcome the limitations of existing local texture patterns, a robust appearance-based feature descriptor, directional ternary pattern (DTP), is proposed for gender classification [7]. Higashi et al. [8] presented a hybrid method which combines Gabor filters and LDP for age and gender classification. The facial features extracted by their method are robust to noise and illumination variations. In this paper, we present an appearance-based texture descriptor constructed with the local zigzag pattern (LZP) to represent the facial features. The recognition performance is evaluated using the FERET database with a support vector machine (SVM) classifier. The Face Recognition Technology (FERET) database [9], which is the best known public database for face recognition, contains 14,126 images of 1,199 individuals. Experimental results are provided to illustrate that our approach yields improved recognition rate against other similar methods. The rest of the paper is organized as follows. Firstly, the concept of LZP is introduced. Secondly, our proposed approach is described in detail. Thirdly, experimental results and comparisons with other methods are presented. Finally, a conclusion is given.

Gender Classification From Face Images With Local Texture Pattern
II. THE PROPOSED APPROACH

2.1. Local Zigzag Pattern
Local pattern features (such as LBP [10] and its different variants [11]) play an important role in various applications of computer vision, especially in the facial image analysis task. The basic concept behind these local pattern features is that the pixels of an image are labeled by considering the properties of the neighborhood surrounding each pixel. In this study, we propose a novel local texture descriptor, LZP, to extract the texture in the image. The LZP descriptor is an eight-bit binary code assigned to each pixel of an image. For a 3 × 3 neighborhood of pixels, the code is generated by comparing its adjacent pixel according to the zigzag scanning order as shown in Figure 1. The LZP value is derived by

\[
LZP = \sum_{i=0}^{9} s(g_i - g_{i+1}) \times 2^i,
\]

where \(g_i\) denotes the gray value of the pixel \(p_i\), \(Th\) is a threshold specified by the user.

Fig. 1. The zigzag scanning pattern.

2.2. Feature Representation using LZP
The combination of local structure information and spatial relationships provides a better facial feature representation and describes the image content more accurately [12]. Therefore, in order to capture the local information of the micro-patterns in the face image, the facial image is divided into M small non-overlapping regions \(\{R_1, R_2, ..., R_M\}\), where each region has the same size commonly. There are two steps to represent the face using the proposed texture descriptor. First, the histogram of LZP within each region is independently generated, respectively. Second, the LZP histograms extracted from each region are then combined yielding a single enhanced histogram. Finally, the enhanced histogram is used as the facial feature vector to represent the global description of the face. Figure 2 illustrates the proposed feature extraction framework.

Fig. 2. The proposed facial features extraction framework.

2.3. Gender Classification with SVM
After feature extraction, the next task is to classify the different input patterns into distinct defined classes with a proper classifier. In the study of facial expression recognition, a comparative analysis of four machine learning techniques are examined, and SVM performed the best [13]. Accordingly, we also consider the SVM to verify the effectiveness of the proposed approach. SVM is a very popular technique for data classification in the machine learning community. The concepts of it are Statistical Learning Theory and Structural Minimization Principle [14]. SVM has been shown to be very effective because it has the ability to look for a hyperplane meet for the requirement of classification, which is a best support vector to distinguish two different classes, under the condition of limited information based on small samples. The support vector can maximize the gap between classes and ensure the accuracy of classification at the same time.

\[
f(x) = \text{sign} \left( \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + c \right),
\]

where \(\alpha_i\) are Lagrange multipliers of the dual optimization problem, \(c\) is a bias or threshold parameter, and \(K(\cdot, \cdot)\) is a kernel function. More and more researchers pay attention to SVM-based classifiers for the facial image analysis task, since their demonstrated robustness and ability to handle large feature spaces makes them particularly attractive for this work.

III. RESULTS AND DISCUSSION

The proposed method is evaluated on the FERET (Face Recognition Technology) database [15]. The FERET face database is a widely used open dataset for evaluation of face recognition algorithms, and has also been used by many researchers for gender recognition. There are 2,718 face images, from which 1,712 belong to male and 1,006 belong to female subjects from FERET database, selected to evaluate the performance of our approach. Some images are selected from the FERET database as shown in Figure 3. The LIBSVM [16] is a most widely used tool for classification. In our work, we used it in all experiments and the classification results are estimated by a 10-fold cross validation scheme.

Fig. 3. Sample images from the FERET face database.
3.1. Influence of Threshold Value
In our approach, the threshold value used to obtain the LZP code is an important factor that influences the recognition performance. In this experiment, we want to show that the impact of the threshold value. For example, we divide the image as $3 \times 3$ non-overlapping regions. The results are presented in Figure 4. It is clear that the use of the threshold value can achieve good results. On the other hand, the recognition rate is not improved as the threshold values are increased. From our observation, the better recognition performance can be obtained when the threshold is set to 5.

![Fig. 4. Recognition rates (%) with different threshold value.](image)

3.2. Comparison with Other Methods
In order to show the effectiveness of the proposed approach, we carry out experiment comparing with other existing methods in [4], [6], and [17]. Note that the results of different methods may not be directly comparable because of differences in experimental setups and pre-processing procedures, but they can still indicate the discriminative performance of every approach. The experimental results are shown in Table 1. We observe that the recognition rate of our approach is substantially higher than the rest.

<table>
<thead>
<tr>
<th>Method</th>
<th>Gender recognition rate (%)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Overall</td>
</tr>
<tr>
<td>PPPTF [4]</td>
<td>96.14</td>
<td>79.84</td>
<td>91.76</td>
</tr>
<tr>
<td>LDP [6]</td>
<td>94.81</td>
<td>95.33</td>
<td>95.05</td>
</tr>
<tr>
<td>LBP + VAR [17]</td>
<td>95.38</td>
<td>87.69</td>
<td>91.54</td>
</tr>
<tr>
<td>Ours</td>
<td>97.59</td>
<td>93.72</td>
<td>96.16</td>
</tr>
</tbody>
</table>

### REFERENCES