Pattern Recognition 46 (2013) 519–528

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Robust gender classification using a precise patch histogram ☆

Huang-Chia Shih *

Human–Computer Interaction Multimedia Laboratory, Department of Electrical Engineering, Yuan Ze University, Chungli, Taoyuan, Taiwan, ROC

ARTICLE INFO

Article history:
Received 17 April 2012
Received in revised form
12 July 2012
Accepted 3 August 2012
Available online 14 August 2012

Keywords:
Biometric analysis
Gender classification
Active appearance model
Local binary patch
Bayesian classifier
Face recognition
Human–computer interaction

ABSTRACT

This study proposed a precise facial feature extraction method to improve the accuracy of gender classification under pose and illumination variations. We used the active appearance model (AAM) to align the face image. Images were modeled by the patches around the coordinates of certain landmarks. Using the proposed precise patch histogram (PPH) enabled us to improve the accuracy of the global facial features. The system is composed of three phases. In the training phase, non-parametric statistics were used to describe the characteristics of the training images and to construct the patch library. In the inference phase, the choice of feature patch from the library needed to approximate the patch of the testing image was based on the maximum a posteriori estimation. In the estimation phase, a Bayesian framework with portion-oriented posteriori fine-tuning was employed to determine the classification decision. In addition, we developed the dynamic weight adaptation to obtain a more convincing performance. The experimental results demonstrated the robustness of the proposed method.

1. Introduction

Recently, the biometric analysis of the human face has been shown to reveal a large amount of physical and psychological information. The use of biometrics in gender classification has resulted in the field of biometrics expanding at a rapid rate. Biometrics can reveal a substantial amount of high-level semantic information, including gender, age, ethnicity, and emotion. Generally speaking, gender classification is divided into two main categories: geometry-based and appearance-based. The geometry-based category is focused on extracting the geometric feature points from the facial image and describes the shape structure of the face. Saatci et al. [1] presented an algorithm to determine the gender and expression of facial images by using active appearance models (AAMs) [2,3] for feature extraction and support vector machines (SVMs) for classification. Mäkinen et al. [4] illustrated a systematic evaluation on gender classification, and showed how face alignment influences the accuracy of gender classification using AAM. Based on the AAM, a pose and shape-independent texture feature extraction for face recognition is proposed in [5].

The appearance-based gender classification methods can be divided into two categories: texture-oriented and statistics-oriented. The texture-oriented approach utilizes different texture descriptors to characterize the gender of a facial image, and utilizes a machine learning strategy to recognize the gender. Many texture characteristics have been applied in gender classification, such as the local binary pattern (LBP) [6,7], local Gabor binary mapping pattern (LGBMP) [8], edge histogram [9], and wavelet transform [10,11]. Baluja et al. [12] demonstrated a feature describing the relationship between the gray-scale values of two pixels using five different types of pixel comparison operators. The Adaboost algorithm [13] was applied to identify the sex of a person from a low resolution facial image.

Among the all-texture features, LBP can be treated as a general approach to the conventionally divergent structural and statistical models of texture analysis. The methodology of the LBP-based face description approach is well-established in both face analysis and its applications. It is robust to monotonic gray-scale changes such as in illumination variations. The basic methodology for the LBP-based face description is as proposed by Ahonen et al. [14]. A notable example is the illumination-invariant face recognition system proposed by Li et al. [15], which combines the LBP features of near-infrared images and Adaboost learning. Hadid and Pietikäinen [16] proposed the spatiotemporal LBPs description for gender recognition from video sequences. Zhang et al. [17] used the LBP-based feature in terms of 40 Gabor filters with different scales and orientations for face recognition. Zhao et al. [18] adopted the spatiotemporal LBP descriptors to represent and recognize the mouth regions and visual lipreading, respectively.

The statistics-based approach usually acquires satisfactory results for the classification scheme. It focuses on using different features that are quantified into a probability to characterize a

☆ This work was supported in part by the National Science Council (NSC) of Taiwan, under grant NSC 100-2628-E-155-007-MY2.

* Tel.: +886 3 4638800x7122; fax: +886 3 4639355.
E-mail address: hcshih@saturn.yzu.edu.tw

0031-3203/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.08.003

facial image as to gender using its visual characteristics. Aghajanian et al. [19] proposed a patch-based framework to determine the ambiguous inside of an object and roughly replace each patch from the pre-defined library and frequency parameters in order to provide the Bayesian posteriori probability. Toews et al. [20] presented the combination of the local scale-invariant feature transform (SIFT) and the object class invariant model for detecting, localizing and classifying the visual gender specific traits. Li et al. [21] used another patch-based feature representation called Spatial Gaussian Mixture Models to describe the image spatial information, while taking the local and global scales into consideration for image misalignment.

In this paper we proposed a robust facial feature description method with a statistical classifier to determine the gender, which represents facial characteristics locally and globally to provide a posteriori probability with more confidence. The active appearance model (AAM) is used to align face images. Facial images are modeled by the patches around the coordinates of landmarks. A so-called precise patch histogram (PPH) will be extracted after the AAM landmark points are determined. We applied non-parametric statistics to describe the characteristics of the training images and construct the patch library in the training phase. We also exploited the relationship between the PPH features of the testing images and those of the library-images to predict the gender in the testing phase. In the present study we proposed a Bayesian framework in which we marginalized the feature patches to determine the classification. It was evident that the accuracy of the global facial features was improved using the proposed PPH feature.

The major contributions of this paper include (1) The proposed patch-based feature acquisition, which provides more precise local and global facial features (i.e., PPH) for gender classification. (2) The flexible library selection approach based on the eigenface with k-means clustering provides a huge range of patch choices for model estimation in both the training and the testing process. (3) The Bayesian-oriented gender determination framework with portion-oriented posteriori fine-tuning.

The remainder of the paper is organized as follows: Section 2 provides a brief overview of the system. In Section 3, we illustrate how to encode the PPH based on the AAM facial feature points. In Section 4, the algorithms of the library selection and the patch-based gender classification are explained. Section 5 presents the experimental results of the system, and finally, we draw our conclusions and provide a discussion in Section 6.

2. System overview

The flow-diagram of the gender classification system is shown in Fig. 1 as follows: (1) apply the characteristics of the eigenface with k-means clustering for the library selection, (2) employ AAM to extract the landmarks from which the feature patches can be extracted, (3) choose the representative samples from the training image database to be used as a patch library, (4) classify the gender by using the inference procedure and the Bayesian estimation to fine tune the classification.

In the training phase, the posteriori probability of each feature patch of the library-images is obtained by the training inference. The patch is then ranked using the values of the posteriori probabilities. The library selection procedure is the critical procedure of the proposed gender classification system, and is based on the number of assigned libraries with patch ranks. In the testing phase, the posteriori probability of each feature patch of the input testing image is determined using the on-line inference process. Finally, a well-defined Bayesian estimation algorithm is used to marginalize the overall posteriori probabilities of the feature patches to make the gender decision. In addition, the portion-oriented posteriori fine-tuning method is used to enhance the results of the classification.

The eigenface [22] determination is based on the principle component analysis (PCA) for finding an appropriate feature representation in a low-dimensional space. There are two primacy procedures: (a) *Eigenspace generation*: given I face normalization image patches $\mathbf{x}_i$ which converted into the column vector type. (b) *Projection onto the eigenspace*: each of the training image patches $\mathbf{x}_i$ is projected onto the eigenspace as a weighting vector which calculated using the eigenspace and the average patch vector with inner product. After the k-means clustering, we select the centroid of each cluster and the feature point with the minimum Euclidean distance as the qualified library image.

3. Precise facial feature extraction

The patch-based gender classification is mainly composed of three components: testing images, patch library and training images. In the training phase, a statistical model is built to describe the characteristics of the training images and is used to construct the patch library. In the testing phase, all patches of the testing images are approximated by the patches in the library. We then compute the posteriori probability for each patch of the approximate image and apply it for ranking in the training phase.

3.1. Patch feature extraction using AAM

To extract more precise facial features, we need to align the facial images in advance. The AAM algorithm provides a good fit for locating the pre-defined facial model. Thus, we used the AAM to fit the facial images. In this paper there are 28 AAM landmark points that need to be determined and utilized. Each patch will be

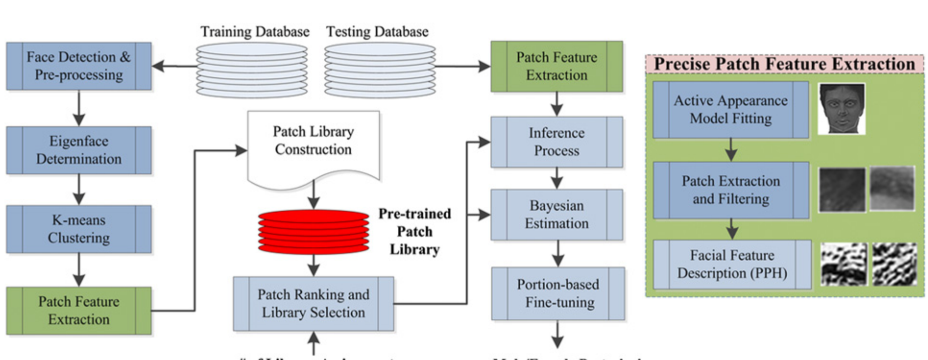

Fig. 1. The flow-diagram of the gender classification system.