FAST GENDER RECOGNITION BY USING A SHARED-INTEGRAL-IMAGE APPROACH

Bau-Cheng Shen[1], *Chu-Song Chen*[1], *Hui-Huang Hsu*[2]

[1]Institute of Information Science, Academia Sinica, Taipei, Taiwan
[2]Department of Computer Science and Information Engineering, Tamkang University, Taipei, Taiwan

ABSTRACT

We develop a new approach for gender recognition. In this paper, our approach uses the rectangle feature vector (RFV) as a representation to identify humans' gender from their faces. The RFV is computationally fast and effective to encode intensity variations of local regions of human face. By only using few rectangle features learned by AdaBoost, we present a gender identifier. We then use nonlinear support vector machines for classification, and obtain more accurate identification results.

Index Terms— Gender Recognition, AdaBoost, Real AdaBoost, Support Vector Machine, Integral Image

1. INTRODUCTION

Human's faces reveal various information including gender, age and ethnicity. They provide important cues for many applications, such as biometric authentication and intelligent human-computer interface. In this paper, we present a new method that can identify humans' genders from their face images (Figure 1).

Fig. 1. The center image is a male face. The left and right images are female faces.

In the past, many researches devote to finding good image features for gender recognition. Among them, Adaboost [1] [2] [3] [4] is a good tool for feature selection. There are many gender-recognition algorithms constructed based on AdaBoost. Wu and Ai [5] proposed a boosting method based on the look-up-table (LUT) weak classifiers. In their work, they train three different detectors (gender, ethnicity, and age) to acquire demographic information by using the two-class Real AdaBoost algorithm [2]. Shakhnarovich and Viola [6] used Viola and Jones' cascaded AdaBoost method [1] to train a face detector, which is a linear combination of the weak classifiers selected from rectangle features. They then selected a set of weak classifiers from the face detector for gender and ethnicity recognition. Instead of rectangle features, Shummet and Henry [7] designed a new set of weak classifiers which use the relationship between two pixels' intensities as features, and show that the recognition rate can be further improved. Instead of AdaBoost, Moghaddam and Yang [8] investigated the use of nonlinear support vector machines (SVMs) to classify gender from face images. With the Gaussian radial basis function (RBF) kernel, a high recognition rate can be achieved. More recently, Fox and Abdesselam [9] used the shunting inhibitory convolutional neural network to extract features and lessen the affection of distortions and variations, and achieve an impressive recognition performance. However, the computational loads of these two approaches are high. They are thus not suitable for real-time applications.

Among the above, the Shakhnarovich and Viola's method [6] is probably the most efficient one in terms of the computational cost for real applications. It is because that some of the rectangle features used for face detection are re-used for gender recognition. Hence, in this method, we do not need to re-compute the features once the face has been detected, and so the entire computational time (including both face detection and gender recognition) can be reduced. In addition, it is also well known that the evaluation of rectangle features can be considerably speeded up by using the integral-image technique [1]. However, the recognition rate of [6] still needs a considerable improvement.

In this paper, we develop a fast gender recognition algorithm based on rectangle features too. Rectangle features can be used to describe sub-regions of a human face, and hence pixel-wise data can be transformed into component-wise data. We cascade several rectangle features into a feature vector called RFV. The RFV is then served as a descriptor for faces to identify the gender. We did a comparative study by employing RFV in various kinds of classifiers, such as the nearest-neighbor (NN), principle component analysis (PCA), and nonlinear SVMs, and suggest an effective detector for gender recognition.

Unlike the method of Shakhnarovich and Viola [6], we

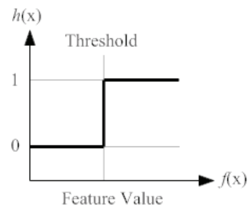

Fig. 2. Binary-valued weak classifier.

Fig. 3. The prototypes of the employed rectangle features.

do not restrict ourselves to select the rectangle features only from those used to construct the face detector. In other words, we allow the rectangle features to be selected arbitrarily in a large feature pool. In this way, the rectangle features selection can be more discriminative, and hence our approach is more accurate for gender recognition. Although in our approach, the rectangle features used for gender recognition and face detection may not be the same, they still share the same integral image. Hence, we can still re-use the integral image, originally built for face detection, to efficiently evaluate the rectangle features for gender recognition. The total computational time (including both face detection and gender recognition) can thus be saved as well.

This paper is organized as follows. Section 2 introduces our approach. Section 3 shows some experimental results. Section 4 gives a conclusion.

2. OUR APPROACH

We employ the AdaBoost algorithm for feature selection. Typically, the AdaBoost alrogithm selects weak learners of binary-valued outputs obtained by thresholding the feature values as shown in Figure 2. However, a disadvantage of the thresholded-type weak learners is that it is too crude to discriminate the complex distributions of the positive and negative training data. To deal with the problem, Schapire et al. [2] suggested the use of the Real AdaBoost algorithm. In the following, we introduce the feature pool used in our work in Section 2.1, and the Real AdaBoost algorithm in Section 2.2. Then, we present the construction of RFV in Section 2.3.

2.1. Feature Pool

The feature pool employed in our approach for Real AdaBoost learning consists of rectangle features. Rectangle features are intensity-based feature. They can be fast evaluated by a pre-constructed integral image, and is very efficient

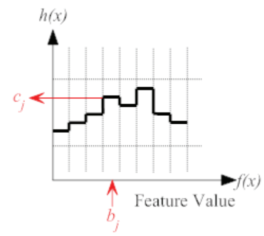

Fig. 4. Real-valued weak classifier.

Fig. 5. Male and female face sample images in Feret database.

to compute [1]. In our work, five kinds of rectangle features are employed, as shown in Figure 3. Given a detection window $\mathcal{W}$, rectangle features with different types, positions and sizes can be produced. Suppose there are K rectangle features in the detection window. We use $g_k(\mathcal{W})$ to represent one of rectangle features. The feature pool can be constructed as

$$\mathcal{F} = \{g_k(\mathcal{W})\},$$

where $k = 1 \sim K$.

In this work, the size of detection window is 24×24. We totally have $32,620$ rectangle features in the feature pool $\mathcal{F}$.

2.2. Feature-Selection Algorithm

Assume that the training data are $\{x_i\}_{i=1}^{m} \in \Re^n$ and the corresponding labels are $\{y_i \in \{-1, 1\}\}_{i=1}^{m}$, where 1 and -1 represent the male and female, respectively. To represent the distributions of positive (male) and negative (female) data, the domain space of the feature value is evenly partitioned into N disjoint bins, denoted as $\{b_j\}_{j=1}^{N}$, as shown in Figure 4. Each bin b_j has a real-valued output c_j, which is calculated according to the ratio of the training data input to the bin. Given input data x and its feature value $f(x)$ evaluated by applying some rectangle feature g_k, the weak learner output $h(x)$ is a mapping $h : x \to \{c_1, \ldots, c_N\}$; if x is quantized to the bin b_j, then $h(x) = c_j$. After selecting T weak classifiers, the strong classifier of Real AdaBoost can be expressed as:

$$H(x) = sign(\sum_{t=1}^{T} h_t(f_t(x)) - \alpha), \tag{1}$$

where α is a threshold. The confidence value of the classifier H is defined as:

$$conf_A(H;x) = \sum_{t=1}^{T} h_t(f_t(x)) - \alpha. \quad (2)$$

A high confidence value implies that the input data is likely to be a positive sample. It has been shown that Real AdaBoost has better discriminating power than the standard binary-valued AdaBoost. More details can be found in [2].

2.3. Forming the feature vector

We select T rectangle features $g_1, ..., g_T$ from feature pool $\mathcal{F}$ by using Real AdaBoost algorithm. The feature values can be calculated depend on the form of rectangle features. The output of each rectangle feature is 1-D feature value. Based on the feature value, the weak classifier can generate a confident value. In the strong classifier H, it uses the sum of T confident values and a threshold α to classify input data x. In our approach, we use all rectangle features' values in the strong classifier as our feature vector, which can be described as

$$RFV(x) = [g_1(x), g_2(x), ..., g_T(x)]$$

3. EXPERIMENTS

Many researchers use the fa-part of Feret face database [10] to test their methods. We also use this database for our approach. In this section, we show how we prepare the data. Then we compare the accuracy of different data representations by using various kinds of classifiers. To evaluate the efficiency for real-time system, we also calculate the identification time of all of the methods. Finally, we analyze the experimental results to suggest an effective approach for gender recognition.

3.1. Data Preparation

The fa-part of Feret face database consists of regular frontal images. It contains 1,759 images with 1,150 male and 609 female faces from 1,010 individuals. Table 1 shows the number of frontal face images of each individual. For example, there are 252 persons with 2 face images. These faces have many variations, including lighting, beard, and eyeglasses. We align and crop these faces according to the positions of eyes and mouth center and normalize them to size 24×24. After normalizing, we perform histogram equalization for all face images. Some face samples are shown in Figure 5. For each method, the average error rate was estimated with five-fold cross validation. That is we split data into five groups and execute five runs and then average all results. For each run we select one non-overlapped group for testing and other groups for training. In every run, the size of training set is 1,408 (920 males and 488 females) and the test set is 351 (230 males and 121 females).

Table 1. The number of frontal face images of each individual in Feret database.

Number of Images	1	2	3	4	5	6	7	8	9	13
Number of People	626	252	53	28	5	9	16	17	2	1

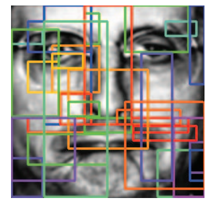

Fig. 6. The color blocks are rectangle features we trained. We just draw the features' position and size instead of types.

3.2. Experimental Results and Analysis

In the experiments, the data representations include thumbnail data (Ori), dimension reduction with PCA (PCA) and our RFV descriptor. We use NN and nonlinear SVMs classifiers for our experiments. We also report the classification results by applying the Real AdaBoost directly.

In Real AdaBoost strong classifier and RFV training, we specify that the detection rate is 99.5% and the false positive rate is 0.1%. Taking the first run as example, when the training process is completed, 33 rectangle features (Figure 6) are selected. We use them in Real AdaBoost strong classifier and also as our RFV descriptor in the test stage. In each method, we select the best result via model selection.

For all data representation and classifiers, combining nonlinear SVMs classifier and the Ori or our RFV representation have better accuracy 92.42% and 92.31% in Table 2. The Ori representation has better accuracy than our descriptor RFV, but its computational load (0.268 second) of 351 test data is heavy. In our experiment, the method which has the best efficiency (0.047 second) of 351 test data is Real AdaBoost strong classifier, but its accuracy (89.80%) is poor. Considering both accuracy and efficiency, the RFV representation with nonlinear SVMs classifier is suggested in our study. The RFV representation only uses 33 dimension data to describe face images for gender recognition. The Figure 7 shows some of the detailed recognition results of the approach (RFV+NN).

4. CONCLUSIONS

We have developed a new approach which combines the rectangle feature vector (RFV) and nonlinear SVMs classifier to identify humans' genders from their face images. Our approach can be integrated with the face detection since they can share the same integral image. An effective and efficient gender identifier can thus be realized.

Acknowledgement

This research was supported by the National Science Council of Taiwan under Grant No. NSC 95-2221-E-001-028-MY3.

Table 2. After executing model selection and five-fold cross validation, the accuracy and computational time (Com. Time) of different methods in test stage are listed in table.

	Accuracy (%)	Com. Time	Dim
Ori+NN	82.91	0.543	576
RFV+NN	**86.50**	**0.0907**	**33**
PCA+NN	82.74	0.2156	55
Real AdaBoost	**89.80**	**0.047**	33
Ori+nSVMs	**92.42**	**0.268**	**576**
RFV+nSVMs	**92.31**	**0.088**	**33**
PCA+nSVMs	92.08	0.6446	55

Fig. 7. This figure show the performance of the method (RFV+NN). The test images are listed at first column and its ten nearest neighbors in training face images are listed at the row. The gender of face images with border are different from the gender of test face images.

5. REFERENCES

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I–511–I–518 vol.1, 2001.

[2] Robert E. Schapire and Yoram Singer, "Improved boosting algorithms using confidence-rated predictions," *Mach. Learn.*, vol. 37, no. 3, pp. 297–336, December 1999.

[3] Yu-Ting Chen and Chu-Song Chen, "A cascade of feed-forward classifiers for fast pedestrian detection," *Asia Conference on Computer Vision*, p. pp. 905V914., 2007.

[4] Yu-Ting Chen and Chu-Song Chen, "Fast human detection using a novel boosted cascading structure with meta stages," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1452–1464, Aug. 2008.

[5] Bo Wu, Haizhou Ai, and Chang Huang, "Facial image retrieval based on demographic classification," *International Conference on Pattern Recognition*, vol. 3, pp. 914–917 Vol.3, Aug. 2004.

[6] G. Shakhnarovich, P.A. Viola, and B. Moghaddam, "A unified learning framework for real time face detection and classification," *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 14–21, May 2002.

[7] Baluja, Shumeet, Rowley, and Henry, "Boosting sex identification performance," *International Journal of Computer Vision*, vol. 71, no. 1, pp. 111–119, January 2007.

[8] B. Moghaddam and Ming-Hsuan Yang, "Learning gender with support faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 707–711, May 2002.

[9] Fok Hing Chi Tivive and A. Bouzerdoum, "A gender recognition system using shunting inhibitory convolutional neural networks," *International Joint Conference on Neural Networks*, pp. 5336–5341, 0-0 2006.

[10] P.J. Phillips, Hyeonjoon Moon, P. Rauss, and S.A. Rizvi, "The feret evaluation methodology for face-recognition algorithms," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 137–143, Jun 1997.