Spatial Histogram Features for Face Detection in Color Images

Hongming Zhang and Debin Zhao

Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin, 150001, China
{hmzhang,dbzhao}@jdl.ac.cn

Abstract. This paper presents a novel face detection approach in color images. We employ spatial histograms as robust features for face detection. The spatial histograms consist of marginal distribution of color image information. Facial texture and shape are preserved by the spatial histogram representation. A hierarchical classifier combining histogram matching and support vector machine is utilized to identify face and non-face. The experiments show that this approach performs an excellent capability for face detection, and it is robust to lighting changes.

1 Introduction

Automatic localization of human faces is significant in many applications, such as human-computer interaction, face animating, face recognition, expression recognition and content-based image retrieval. Face detection is still a challenge because of variability in face orientations, face sizes, face locations, facial expression, partial occlusions, and lighting conditions. A comprehensive survey on face detection in images can be found in [1].

Many approaches have been proposed for face detection in still images. Generally speaking, these approaches can be classified as two categories: model-based technique and feature-based technique. The first one assumes that a face can be represented as a whole unit. Several statistical learning mechanisms are explored to characterize face patterns, such as neural network [2], probabilistic distribution [3], support vector machines [4], principal components analysis [5], naive Bayes classifier [6], and boosting algorithms [7,8]. The second method treats a face as a collection of components. Important facial features (eyes, nose and mouth) are first extracted, and by using their locations and relationships, the faces are detected [9,10].

Skin color is a powerful and fundamental cue that can be used as the first step in the face detection process. Many researchers utilize skin color models to locate potential face areas, then examine the locations of faces by analyzing each face candidate's shape and local geometric information [11,12].

What features are stable and scalable for face detection is still an open problem. Previous works have used many representations for facial feature extraction, such as edges, wavelets, and rectangle features [13,14,7].

K. Aizawa, Y. Nakamura, and S. Satoh (Eds.): PCM 2004, LNCS 3331, pp. 377–384, 2004.

In this paper, we propose a face detection method in color images. Spatial histograms are calculated on color measurements. Based on the spatial histogram representation, discriminating features are extracted for face detection. A hierarchical classifier combining histogram matching algorithm and support vector machine is utilized to identify face and non-face.

The structure of this paper is as follows. In Section 2, we propose the spatial histogram for face representation. A discriminating analysis of spatial histogram is given in Section 3. In Section 4, we describe a face detection system using hierarchical classification. Experimental results are provided in Section 5. At last, we give conclusions in Section 6.

2 Spatial Histogram Features for Face Representation

Human face patterns have much information including skin color cues and facial components configurations. In this section, we describe a new face pattern representation combining color cues and face spatial structures. Basically, we model face regions by their histograms, or marginal distribution, over the face color measurements.

Histogram is a global representation of one image pattern, which is translation invariant. However, for some non-face images and face images, their histogram can be very close or even identical, making the histogram not sufficient for object detection.

In order to overcome this limitation of histogram, we enlarge discrimination information by two ways. First, we adopt color information as measurements of face pattern. Second, we introduce spatial histograms to extract face shape information.

2.1 Color Measurements and Histogram Based Representation

Color information is often effective in image segmentation. For any color, its hue keeps almost constant under different lighting conditions. We consider YUV color space as the interesting color space because it is compatible to human color perception. We first convert the RGB color information to YUV color information:

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ U = -0.147R - 0.289G + 0.436B. \\ V = 0.615R - 0.515G - 0.100B \end{cases} \tag{1}$$

Hue is defined as the angle of vector in YUV color space, given by

$$\theta = \arctan(V/U). \tag{2}$$

We define a pattern as a 32x32 color image with RGB values, and its representation constitutes of five measures: Y, R, G, B and θ. We denote the five measurements by $m(i), i = 1, 2, 3, 4, 5$.

We compute histogram-based pattern representation as follows. First we extract the five measurements from original color images, next we apply histogram normalization on these measurements respectively, and then we use Basic Local Binary Pattern operator(LBP) [15,17] to transform the obtained measures, finally we compute histograms of the processed measurements as representation. We employ LBP opeator since it is invariant against any monotonic transformation of the gray scale or color measure. Fig.1 shows two gray image samples, their LBP images and histograms.

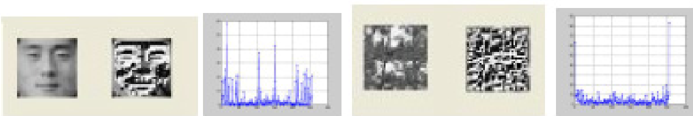

Fig. 1. Image samples, LBP images and LBP histograms

2.2 Spatial Histogram

Human face is a near-regular texture pattern generated by facial components and their configuration. We use 23 different spatial templates to preserve shape of face patterns. These spatial templates are shown in fig.2. Each template is a 32x32 mask. It has a white rectangle, and each white pixel has value 1 and each black pixel has value 0. We denote each template by $rect(x, y, w, h)$, where (x, y) is location and (w, h) is size of the white part in template. The spatial template set is simply denoted by $\{rect(1), rect(2), ..., rect(23)\}$.

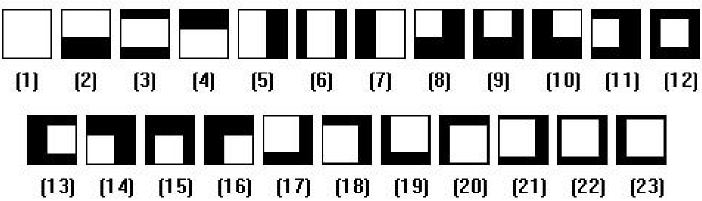

Fig. 2. Spatial templates

For a single spatial template $rect(x, y, w, h)$ and a given measurement $m(i)$, we model image window by histogram. First we convolve the whole preprocessed measurement window with the mask, and then we compute histogram, or marginal distribution, from the resulting measurement. We call this histogram as *spatial histogram*. Given an image P, one spatial histogram is denoted by $SH_{m(i)}^{rect(x,y,w,h)}(P)$.

There are many methods of measuring similarity between two histograms. We adopt histogram intersection measure since it is stable and computational inexpensive. According to [16], the intersection match of two histograms is defined as:

$$D(SH_1, SH_2) = \sum_{i=1}^{k} \min(SH_1^i, SH_2^i), \tag{3}$$

where SH_1 and SH_2 are two histograms, and K is the number of bins in the histograms.

3 Discriminant Analysis of Spatial Histogram

Spatial histogram representation is introduced in last section. In this section, we analysis each spatial histogram's discriminating ability and construct compact feature vectors for face detection.

To train and evaluate our system, we collect 11400 32x32 color face samples as positive training set. 3000 color images contained no face patterns are used as negative training set.

Histogram matching is a direct approach for object recognition. In this approach, we apply histogram intersection as distance measure to object recognition. Given a database with n face samples, we represent each face spatial histogram type by the average spatial histogram of face training samples, defined as

$$SH_{m(i)}^{rect(x,y,w,h)}(face) = \frac{1}{n}\sum_{j=1}^{n} SH_{m(i)}^{rect(x,y,w,h)}(P_j), \tag{4}$$

where P_j is a face training sample, $m(i)$ is the measurement and $rect(x, y, w, h)$ is the spatial template. For any sample P, we define its *histogram-matching feature* as the distance to the average face histogram, given by

$$f_{m(i)}^{rect(x,y,w,h)}(P) = D(SH_{m(i)}^{rect(x,y,w,h)}(P), SH_{m(i)}^{rect(x,y,w,h)}(face)). \tag{5}$$

Each type of spatial histogram has the discriminating ability between face and non-face pattern. To demonstrate this property, we take the gray histogram of the first template as an example. $SH_{m(1)}^{rect(0,0,32,32)}(face)$ is the average model of the positive train samples. Here we adopt histogram-matching feature

$$f_{m(1)}^{rect(0,0,32,32)}(P) = D(SH_{m(1)}^{rect(0,0,32,32)}, SH_{m(1)}^{rect(0,0,32,32)}(face))$$

as one feature. Fig.3 shows its positive and negative distribution over the training set. On this feature, we use threshold to classify face and non-face. By setting the threshold to 0.7, we retain 99.6% face detection rate with false alarm rate 30.2%.

We combine all kinds of histogram-matching features to obtain a compact and discriminating feature vector, given by

$$F = [f_{m(1)}^{rect(1)}, ..., f_{m(1)}^{rect(23)}, ..., f_{m(5)}^{rect(1)}, ..., f_{m(5)}^{rect(23)}]. \tag{6}$$

We call this feature vector as *Joint Spatial Histogram-Matching feature*(JSHM feature). The total set of spatial histogram types has size 23x5=115. Therefore,the JSHM feature is 115 dimensional.

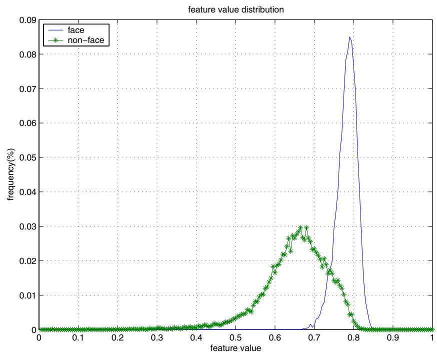

Fig. 3. Feature distribution of face and non-face on gray histogram of the first template

4 Hierarchical Classification for Face Detection System

We apply a hierarchical classification scheme for face detection, shown in Fig 4. First, we use histogram matching as coarse detection stage. In the second stage, a Support Vector Machine is used for fine face detection.

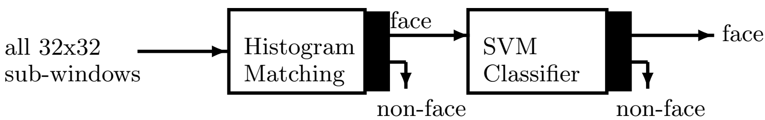

Fig. 4. Hierarchical classification for face detection system

4.1 Histogram Matching

As demonstrated in section 3, each type spatial histogram-matching feature has the ability to identify face and non-face. In practice, we specify three gray spatial histograms, $\{rect(1), rect(2), rect(6)\}$, to detect face pattern jointly. The decision rule is as follows:

$$H(P) = \begin{cases} 1 \text{ face} & \text{if}(f_{m(1)}^{rect(1)}(P) \geq T_1 \text{and} f_{m(1)}^{rect(2)}(P) \geq T_2 \text{and} f_{m(1)}^{rect(6)}(P) \geq T_6) \\ 0 \text{ nonface} & \text{otherwise} \end{cases}, \tag{7}$$

where T_1, T_2, and T_6 are thresholds for matching. In our experiments, the three thresholds are all set to be 0.7.

4.2 Support Vector Machine

Given the JSHM feature and training sets of positive and negative image samples, face detection is regarded as a two-class pattern classification problem. A

SVM classifier[18] performs pattern recognition for a two-class problem by determining the separating hyper plane that maximum distance to the closest points of training set. These closet points are called support vectors.

Using JSHM feature and bootstrap technique, we train a non-linear SVM for face detection on the train data sets. The kernel adopted is a RBF kernel.

5 Experiments

We implement the proposed approach and conduct experiments to evaluate its effectiveness. Many face databases commonly used by researchers, including FERET face recognition dataset and CMU face detection dataset, contain gray-scale images only. Therefore, we have built a large-scale color face image database. This database set is consisted by images from different sources: news photos, movie video pictures, personal digital images and surveillance images. These images contain multiple faces under complex backgrounds and different lighting conditions. In total, 251 color images are collected. There are 356 frontal faces with variations in color, position, size and expression.

Our system can detect frontal faces under complex backgrounds. In Fig.5, some face detection examples are given. The examples demonstrate that our approach can handle multiple faces with different colors and sizes. Also our approach can detect dark skin-tone and bright skin-tone faces under complex lighting conditions.

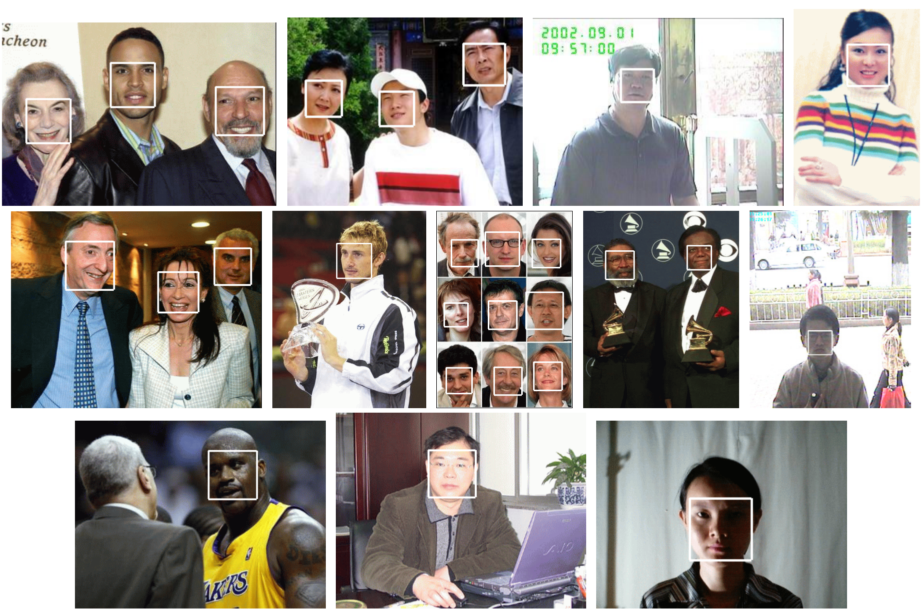

Fig. 5. Face detection results

On the color image dataset, the system correctly detected 334 faces and produced 30 false alarms. The testing result is listed in Table.1. The correct detection rate is 93.8% and the precision is 91.7%, which prove that the proposed approach is effective in detecting faces. Note that precision is defined as the ratio of the number of positive detected faces to the sum of positive detected faces and negative detected faces.

Table 1. Testing results

Number of faces	Positive detected faces	Detection rate	Negative detected faces	Precision
356	334	93.8%	30	334/(334+30)=91.7%

As shown in Fig.6, the proposed method sometimes fails to detect real faces or detects false faces. Some faces are far smaller than 32x32, so they are often missing detected. Some false detected faces are similar to face patterns in color and shape. How to overcome these limitations is one of our future works.

Fig. 6. Face detection examples with missing and false detection

6 Conclusions and Future Work

In this paper, we propose a novel face detection approach in color images. We apply spatial histogram as pattern representation and extract histogram-matching features for face detection. We use a hierarchical scheme combining a histogram matching algorithm and a support vector machine for classification. Experiments show that the proposed approach performs an excellent capability for face detection in color images and it is robust to illumination changes. We will extend this approach to other objects recognition tasks, such as cars, and texts et al.

Acknowledgements. This research is partially supported by National Nature Science Foundation of China (No. 60332010), and ISVISION Technologies Co. Ltd. Thanks to Yuemin Li for her help on SVM code.

References

1. M.H. Yang, D.J. Kriegman, N. Ahuja: Detecting Faces in Images: A Survey. IEEE Translations on Pattern Analysis And Machine Intelligence, 24(1): 34-58, 2002.
2. H.A. Rowley, S. Baluja, T. Kanade: Neural Network-Based Face Detection. IEEE Translations on Pattern Analysis And Machine Intelligence, 20(1): 29-38,1998.
3. K.K. Sung, T. Poggio: Example-Based Learning for View-Based Human Face Detection. IEEE Translations on Pattern Analysis And Machine Intelligence, 20(1): 39-50,1998.
4. E. Osuna, R. Freund, F. Girosi: Training Support Vector Machines: an Application to Face Detection. Proceedings of CVPR, 130-136, 1997.
5. B. Menser, F. Muller: Face detection in color images using principal component analysis. Proceedings of 7th International Congress on Image Processing and its Applications, 13-15, 1999.
6. H. Schneiderman, T. Kanade: A Statistical Method for 3D Object Detection Applied to Faces and Cars. IEEE Conference on Computer Vision and Pattern Recognition, 2000.
7. P. Viola, M. Jones: Robust Real Time Object Detection. IEEE ICCV Workshop on Statistical and Computational Theories of Vison, 2001.
8. S.Z. Li, et al.: Statistical Learning of Multi-View Face Detection. Proc. of the 7th European Conf. on Computer Vision, 2002.
9. K.C. Yow, R. Cipolla: Feature-Based Human Face Detection. CUED/F-INFENG/TR 249, August 1996.
10. J. Miao, B.C. Yin, K.Q. Wang, L.S. Shen, and X.C. Chen: A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template. Pattern Recognition, 32(7): 1237-1248, 1999.
11. K. Sobottka, I. Pitas: A novel method for automatic face segmentation, facial feature extraction and tracking. Signal Processing: Image Communication, 12(3): 263-281, 1998.
12. Q.B. Sun, W.M. Huang, and J.K. Wu: Face detection based on color and local symmetry information. Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, 130-135, 1998.
13. F. Bernhard, K. Christian: Real-Time Face Detection using Edge-Orientation Matching.Audio and Video-based Biometric Person Authentication (AVBPA'2001), 78-83, 2001.
14. C. Garcia, G. Tziritas: Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis. IEEE Transactions on Multimedia, 1(3): 264-277, 1999.
15. M. Pietikäinen, T. Ojala, and Z. Xu: Rotation-invariant texture classification using feature distributions. Pattern Recognition, 33(1):43-52, 2000.
16. M. Swain, D. Ballard: Color indexing. Int. J. Computer Vision, 7(1): 11-32,1991.
17. T. Ojala, M. Pietikäinen, T. Mäenpää: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Macine Intelligence, 24(7): 971-987, 2002.
18. V. Vapnik: Statistical Learning Theory. Wiley, New York, 1998.