An Improved Face Recognition Algorithm Using Histogram-Based Features in Spatial and Frequency Domains

Qiu Chen, Koji Kotani, Feifei Lee, Tadahiro Ohmi

Abstract—In this paper, we propose an improved face recognition algorithm using histogram-based features in spatial and frequency domains. For adding spatial information of the face to improve recognition performance, a region-division (RD) method is utilized. The facial area is firstly divided into several regions, then feature vectors of each facial part are generated by Binary Vector Quantization (BVQ) histogram using DCT coefficients in low frequency domains, as well as Local Binary Pattern (LBP) histogram in spatial domain. Recognition results with different regions are first obtained separately and then fused by weighted averaging. Publicly available ORL database is used for the evaluation of our proposed algorithm, which is consisted of 40 subjects with 10 images per subject containing variations in lighting, posing, and expressions. It is demonstrated that face recognition using RD method can achieve much higher recognition rate.

Keywords—Face recognition, Binary vector quantization (BVQ), Local Binary Patterns (LBP), DCT coefficients.

I. INTRODUCTION

Face recognition seems to be the most natural and effective method compared with other personal biometric features, such as voice, fingerprint, iris pattern, etc. because it is a similar wary human does and there is no need to use special equipments. Many algorithms have been proposed for solving face recognition problem [1]-[11]. These algorithms can be roughly divided into two categories, namely, statistics-based and structure-based approaches. Statistics-based approaches [5]-[7] attempt to capture and define the face as a whole. The face is treated as a two dimensional pattern of intensity variation. Under this approach, the face is matched through finding its underlying statistical regularities. Based on the use of the Karhunen-Loeve transform, PCA [5] is used to represent a face in terms of an optimal coordinate system which contains the most significant eigenfaces and the mean square error is minimal. However, it is highly complicated and computational-power hungry, making it difficult to implement them into real-time face recognition applications.

Although BVQ histogram and LBP histogram have been demonstrated to be very efficient for face recognition, it only uses the counted histogram as the feature information to identify the people and the geometric information of face is unused. So we cannot know which region of facial part the matched codevector belongs to. The region-division (RD) method [30] is adopted in this paper, which combines the position information of facial components and the histogram features for recognition. The combined features of face are expected to be more robust and effective.
This paper is organized as follows. A brief introduction to DCT as well as LBP histogram is given in Section II. Our proposed face recognition method will be described in detail in Section III. Experimental results will be discussed in Section IV. Finally, we make a conclusion in Section V.

II. RELATED WORKS

A. Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) is used in JPEG compression standard. The DCT transforms spatial information to decoupled frequency information in the form of DCT coefficients.

2D DCT with block size of $N \times N$ is defined as:

$$
C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left( \frac{2\pi x u}{N} \right) \cos \left( \frac{2\pi y v}{N} \right)
$$

(1)

$$
f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u, v) \alpha(u) \alpha(v) \cos \left( \frac{2\pi x u}{N} \right) \cos \left( \frac{2\pi y v}{N} \right)
$$

(2)

where, $\alpha(u) = \begin{cases} \frac{1}{\sqrt{N}} & : \text{for } u = 0 \\ \frac{2}{\sqrt{N}} & : \text{for } u = 1, 2, ..., N - 1 \end{cases}$

(3)

B. Face Recognition Using Binary Vector Quantization in Low-Frequency DCT Domains

In [27], we proposed a feature extraction algorithm for face recognition using binary vector quantization (VQ) to generate feature vectors of facial image from DCT (Discrete Cosine transform) coefficients in low frequency domains.

First, low-pass filtering is carried out using 2-D moving filter. Block segmentation step, in which facial image is divided into small image blocks with an overlap, namely, by sliding dividing-partition one pixel by one pixel, is the following. Then the pixels in the image blocks (typical size is 8x8) are transformed using DCT according to (1).

A typical sample of transformed block is shown in Fig. 1. The DCT coefficients of the image block are then used to form a feature vector. From left to right and top to bottom, the frequency of coefficients changes from low to high as shown in Fig. 1. Because low frequency component is more effective for recognition, we only use the coefficients on the left and above to extract features. The equation for calculation is shown as:

$$
a[0] = AC01;
a[1] = AC11;
a[2] = AC10;
a[3] = (AC02 + AC03 + AC12 + AC13) / 4;
a[4] = (AC22 + AC23 + AC32 + AC33) / 4;
a[5] = (AC20 + AC21 + AC30 + AC31) / 4
$$

(4)

where $a[i]$ is the element of extracted feature vector, and $d[i][j]$ is the coefficient value at point $(i, j)$, respectively.

After that, quantization of the feature vectors is implemented. There are only 2 types of value for each $a[i]$, so the number of combination of 6-dimensional vector is 64, which is very easy and fast to be determined. The number of vectors with same index number is counted and feature vector histogram is easily generated, and it is used as histogram feature of the facial image. In the registration procedure, this histogram is saved in a database as personal identification information. In the recognition procedure, the histogram made from an input facial image is compared with registered individual histograms and the best match is output as recognition result.

C. Local Binary Patterns (LBP) Histogram

The original LBP operator proposed by [29], is used for robust texture description. The operator labels the pixels of an image by thresholding the 3x3-neighbourhood of each pixel with the center value and considering the result as a binary number. Then the histogram of the labels can be used as a texture descriptor. Fig. 2 shows an illustration of the basic LBP operator.

The limitation of the fundamental LBP operator is its small 3x3 neighborhood which cannot capture dominant features with large scale structures. Hence, the operator later is extended to use neighborhood of different sizes. As shown in Fig. 3, LBP($P, R$) means $P$ sampling points on a circle of radius of $R$ to get LBP features. For instance, LBP(8, 2) means comparing a neighborhood of 8 on the circle of radius of 2 to get LBP features.
After labeling an image with the LBP operator, the histogram of the labeled image \( p(x,y) \) can be defined as

\[
H_u = \sum_{x,y} U(p(x,y) = u), u = 0,1,..., n - 1
\]  

(5)

where \( n \) is the number of different labels produced by the LBP operator and

\[
U(A) = \begin{cases} 
1, & A = \text{true} \\
0, & A = \text{false}
\end{cases}
\]  

(6)

An LBP histogram can effectively describe the distribution of the local micro-patterns over a whole face image without any indication about their locations. For efficient face representation, one should also retain spatial information. Thus, a face image can be equally divided into small regions. And then, the LBP features extracted from each sub-region are concatenated into a single histogram as

\[
H_{u,v} = \sum_{x,y} U(p(x,y) = u) U((x,y) \in R_v)
\]  

(7)

where \( u = 0,1,..., n - 1 \) and \( v = 0,1,..., m - 1 \).

III. PROPOSED METHOD

As described in Section II, B, we have proposed a face recognition algorithm by applying binary quantization on the low-frequency DCT coefficient blocks, which was demonstrated to be effective for face recognition by experimental results. Actually, it can be thought that phase information of low-frequency DCT coefficients is extracted by this algorithm. If we could combine spatial information of the facial image, the composite features of face are expected to be more robust and effective.

We utilize LBP to represent facial features in spatial domain. In this paper, we propose an improved face recognition algorithm using combined histogram-based features. Fig. 5 shows proposed face recognition process steps. First, low-pass filtering is carried out using 2-D moving filter. This low-pass filtering is essential for reducing high-frequency noise and extracting most effective low frequency component for recognition. Then the total face area is divided into several regions with respective sizes as shown in Fig. 4, and histogram-based features in spatial and frequency domains of every region are generated separately.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. ORL Database

Face database of AT&T Laboratories Cambridge [25], [26] is used for recognition experiments. In the database, 10 facial...
images for each of 40 persons (totally 400 images) with variations in face angles, face sizes, facial expressions, and lighting conditions are included. Each image has a resolution of 92x112. Five images were selected from each person’s 10 images as probe images and remaining five images are registered as album images. Recognition experiment is carried out for 252 \(_{10C5}\) probe-album combinations by rotation method. The algorithm is programmed by ANSI C and run on PC (Pentium(R)D processor 840 3.2GHz).

B. Results and Discussions

Figs. 6, 7 show the comparison of the recognition results with different features. The average recognition rates obtained by each case with block size of 8x8 are shown here. Recognition success rates are shown as a function of filter size. Recognition results of maximum average rate only using LBP histogram achieved 90.7%, average recognition rate increases combined with BVQ histogram of low-frequency DCT coefficients. By using RD method with the block size of 2x4, the maximum of the average rate 99.06% is achieved, which is almost 5.3% higher than that only using BVQ histogram in our previous work (the maximum of the average rate is 93.7%) [27].

Fig. 6 Comparison of recognition results

Fig. 7 Comparison of recognition results with different features

Fig. 8 shows recognition results using combined features with the same weighting coefficient of two histogram features. Recognition success rates are shown as a function of filter size. “Max,” “Min” and “Ave” stand for the best case, worst case, and average results in 252 \(_{10C5}\) probe-album combinations, respectively. The highest average recognition rate of 99.06% is obtained at the filter size of 5x5. By combining these two different features, namely spatial and frequency domain information of a facial image, the most important information for face recognition can effectively be extracted.

C. Conclusions and Future Work

We have developed a very simple yet highly reliable face recognition method using features extracted from low-frequency DCT domain and spatial domain of a facial image, which is combined with BVQ histogram and LBP histogram. We adopted region-division (RD) method in this paper, which combines the position information of facial components and the histogram features for recognition. Excellent face recognition performance has been verified by using publicly available ORL database. The effect of the image block size will be discussed in our future work, as well as the performance evaluation of the face recognition using larger face database.

ACKNOWLEDGMENT

This work was partially supported by JSPS KAKENHI Grant Number 15K00159.

REFERENCES


