An Improved Histogram-Based Features in Low-Frequency DCT Domain for Face Recognition

Qiu Chen, Koji Kotani, Feifei Lee, and Tadahiro Ohmi

Abstract—Previously, we proposed an efficient algorithm for facial image recognition combined with vector quantization (VQ) histogram and energy histogram in low-frequency DCT domains. The former algorithm is essential for utilizing the phase information of DCT coefficients by applying binary vector quantization (BVQ) on DCT coefficient blocks. The latter algorithm, energy histogram can be considered to add magnitude information of DCT coefficients. These two histograms, which contain both phase and magnitude information of a DCT transformed facial image, are utilized as a very effective personal feature. In this paper, we propose a novel quantization optimization method for energy histogram according to the maximum entropy principle (MEP) as a design criterion. Publicly available AT&T 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 optimized energy histogram by maximization of information-theoretic entropy can achieve much higher recognition rate.

Index Terms—Face recognition, binary vector quantization (BVQ), energy histogram, DCT coefficients.

I. INTRODUCTION

Many algorithms have been proposed for solving face recognition problem [1]-[12] due to its potential applications in many fields such as law enforcement applications, security applications and video indexing, etc. These algorithms can be roughly divided to two categories, namely, statistics-based and structure-based approaches. Statistics-based approaches [3]-[5] 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 [3] 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. Structure-based approach [6], [7] uses the relationship between facial features, such as the locations of eye, mouth and nose. It can implement very fast, but recognition rate usually depends on the location accuracy of facial features, so it cannot give a satisfied recognition result. There are many other algorithms have been used for face recognition, such as Local Feature Analysis (LFA) [11], neural network [1], local autocorrelations and multi-scale integration technique [2], and other techniques have been proposed.

Discrete Cosine Transform (DCT) is not only widely used in many image and video compression standards [13], but also for pattern recognition as a means of feature extraction [14]-[23]. The main merit of the DCT is its relationship to the KLT [19]. It has been demonstrated that DCT best approach KLT [24], but DCT can be computationally more efficient than the KLT depending on the size of the KLT basis set.

In our previous work [25], we present a simple, yet highly reliable face recognition algorithm using binary vector quantization (BVQ) method for facial image recognition in compressed DCT domain. Feature vectors of facial image are firstly generated by using DCT coefficients in low frequency domains. Then codevector referred count histogram, which is utilized as a very effective facial feature value, is obtained by BVQ processing.

This algorithm can be considered utilizing the phase information of DCT coefficients by applying binary quantization on the DCT coefficient blocks. If magnitude information of DCT coefficients is added, the composite features of face are expected to be more robust and effective. In our previous work [26], we utilize energy histogram to represent magnitude features of DCT coefficients. These two histograms, which contain both phase and magnitude information of a DCT transformed facial image, are utilized as a very effective personal value. Recognition results with different type of histogram features are first obtained separately and then combined by weighted averaging.

It was found that the distribution of average energy histogram appears non-uniform and concentrated only in very low frequency regions. In this paper, we propose a novel quantization optimization method for energy histogram according to the maximum entropy principle (MEP) as a design criterion.

This paper is organized as follows. A brief introduction to DCT as well as energy histogram is given in Section II. Our proposed quantization optimization method for energy histogram as well as 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 follows:

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

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

where

$$a(\omega) = \begin{cases} \frac{1}{\sqrt{N}} & \text{for } \omega = 0 \\ \frac{2}{\sqrt{N}} & \text{for } \omega = 1, 2, ..., N-1 \end{cases}$$

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

In our previous work [25], we proposed a feature extraction algorithm for face recognition using binary vector quantization (BVQ) 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 the equation (1).

But we found that the distribution of average energy histogram of low-frequency DCT coefficients appears non-uniform and concentrated only in very low frequency regions. According to Maximum Entropy Principle (MEP), the maximum entropy distribution will be achieved when the value of a random variable (counts) equals the average. Such a non-uniform distribution cannot satisfy the MEP, so the recognition performance cannot be expected be best because the average information content will not be maximum.

In this paper, we propose a novel quantization optimization method for energy histogram according to the maximum entropy principle (MEP) as a design criterion.

We investigate the whole quantization structure of low-frequency DCT coefficients to optimize the quantization. At first, the quantization level is set to be only 1 from 0 to maximum value $N$, and then the number of vectors quantized in each quantization region is counted and a histogram of a facial image is generated.

Fig. 2 shows an average histogram of all 400 images in the database (Log scale is used in the Y-coordinate direction).
This histogram shows vectors of low-frequency DCT coefficients are concentrated in very low-frequency regions, and decrease in higher frequency regions. We change the boundary of bins dynamically so as to make the amount of vectors in each bin into equal value, and then we can simply determine the adaptive boundary.

Fig. 2. Average histogram of 400 images. Quantization levels are set at 1 (totally 801), and log scale is used in the Y-coordinate direction.

Fig. 3. Low-frequency DCT coefficients for energy histogram.

Fig. 4. Face recognition process steps using combined histogram-based features. 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.

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 the equation (1). After generations of low-frequency DCT coefficients, binary quantization of the feature vectors is implemented as described in Section II (B), and then BVQ histogram of low-frequency DCT coefficients is created.

On the other hand, energy histogram of low-frequency DCT coefficients is also generated after 2D-DCT. Because low frequency component is more effective for recognition, we only use the coefficients on the left and above to extract features. This can also reduce computation time compared with using the whole DCT coefficients. In this paper, we use 4x4 coefficient blocks as shown in Fig. 3, the same domain as BVQ histogram used at the upper left corner of the DCT coefficient block which retain the higher energy level of the image. The DC coefficient is not included in the features to reduce the influence of the lighting conditions of the images. Once the adaptive boundaries of bins are determined, the energy histogram is created by using formula (4) as shown in Section II (C).

These two histograms, which contain both phase and magnitude information of a DCT transformed facial image, are utilized as a very effective personal value. Recognition results with different type of histogram features are first obtained separately and then combined by weighted averaging.

Fig. 5. Comparison of recognition results.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. ORL Database

Face database of AT&T Laboratories Cambridge [29], [30] 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 \((10\binom{5}{2})\) 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

In our experiments, the bin size of energy histogram is set
to 100, and the adaptive boundaries of bins are determined as described in Section III. Fig. 5 shows the comparison of the recognition results with different features. The average recognition rates obtained by each case with block size of 8×8 are shown here. Recognition success rates are shown as a function of filter size. Although recognition results only using energy histogram of low-frequency DCT coefficients (“N8_energy_hist”) are not satisfied, average recognition rate increases combined with BVQ histogram of low-frequency DCT coefficients (“N8_combined (1:1)”), which the maximum of the average rate 96.1% is achieved. In this paper, by using the novel quantization optimization method for energy histogram according to MEP as the design criterion, the recognition results only using energy histogram of low-frequency DCT coefficients (“N8_energy_hist MEP”) rise up and the maximum of the average rate of 88.3% is achieved. Furthermore, the maximum of average recognition rate increases to 97.25% which combined with BVQ histogram of low-frequency DCT coefficients (“N8_combined_MEP (1:1)”), which is 1.2% higher than non-optimization method.

It can be said that by combining these two different features, namely phase information and magnitude information of low-frequency DCT coefficient blocks, the most important information for face recognition can effectively be extracted.

V. CONCLUSIONS AND FUTURE WORK

We have developed a very simple yet highly reliable face recognition method using features extracted from low-frequency DCT domain, which is combined with BVQ histogram and energy histogram. Based on the maximum entropy principle (MEP), we optimize the quantization for energy histogram by changing the boundary of bins dynamically so as to make the amount of vectors in each bin into equal value, and then we can simply determine the adaptive boundary. Excellent face recognition performance has been verified by using publicly available ORL database.

REFERENCES


Q. Chen received the Ph.D. degree in electronic engineering from Tohoku University, Japan, in 2004. In 2011, he became an associate professor at New Industry Creation Hatchery Center (NICHe), Tohoku University, and he is currently an associate professor at the Department of Information and Communication Engineering, Kogakuin University. His research interests include pattern recognition, computer vision, information retrieval and their applications.
K. Kotani received the B.S., M.S. and Ph.D. degrees all in electronic engineering from Tohoku University, Japan, in 1988, 1990 and 1993, respectively. He joined the Department of Electronic Engineering, Tohoku University as a research associate in 1993. From 1997 to 1998, he belonged to the VLSI Design and Education Center (VDEC) of the University of Tokyo. He is currently an associate professor at the Department of Electronic Engineering, Tohoku University. He is engaged in the research and development of high performance devices/circuits as well as intelligent electronic systems.

F. F. Lee received the Ph.D. degree in electronic engineering from Tohoku University, Japan, in 2007. Since then, she has been an assistant professor at New Industry Creation Hatchery Center (NICHe), Tohoku University, and now she is a professor at the University of Shanghai for Science and Technology. Her research interests include pattern recognition, video retrieval and multimedia processing.

T. Ohmi served as a research associate in the Department of Electronics, Tokyo Institute of Technology, from 1966 to 1972. Then, he moved to Research Institute of Electrical Communication, Tohoku University and became an associate professor in 1976. In 1985, he became a professor at the Department of Electronics, Faculty of Engineering, Tohoku University. Since 1998, he has been a professor at New Industry Creation Hatchery Center (NICHe), Tohoku University.

His research field covers whole Si-based semiconductor and flat panel display technologies in terms of material, process, device, circuit, and system technologies. He is known as an originator of "Ultra1clean Technology," which introduced ultraclean and scientific way of thinking into semiconductor manufacturing industry and became indispensable technology today.