LBP EDGE-MAPPED DESCRIPTOR USING MGM INTEREST POINTS FOR FACE RECOGNITION

Jou Lin, and Ching-Te Chiu

Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan
Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan
s103064534@m103.nthu.edu.tw, ctchiu@cs.nthu.edu.tw

ABSTRACT

In recent years, face recognition has become a popular topic in academia and industry. Current local methods such as the local binary pattern (LBP), and scale invariant feature transform (SIFT) perform better than holistic methods, but their high complexity levels limit their application. In addition, SIFT-based schemes are sensitive to illumination variation. We propose an LBP edge-mapped descriptor that uses maxima of gradient magnitude (MGM) points. It can completely illustrate facial contours and has low computational complexity. Under variable lighting, experimental results show that our proposed method has a 16.5% higher recognition rate and requires 9.06 times less execution time than SIFT in the FERET database subset fc. In addition, when applied to the Extended Yale Face Database B, our method outperformed SIFT-based approaches as well as saving about 70.9% in execution time. Furthermore, in uncontrolled conditions, our method has a 0.82% higher recognition rate than local derivative pattern histogram sequences (LDPHS) in the Unconstrained Facial Images (UFI) database.

Index Terms— Face Recognition, Maxima of Gradient Magnitude, Local Binary Pattern, Binary Feature

1. INTRODUCTION

For several decades, face recognition has been an active area of research in the field of computer vision. With ongoing improvements in technology, it currently plays an important role in many devices and applications such as surveillance systems, and access control systems. However, there remain a variety of challenges (e.g., illumination, and expression) with respect to real-world conditions. To effectively address these problems, a variety of methods have been proposed.

Current face recognition approaches are typically classified into two categories. The first type includes holistic methods, which utilize subspace learning methods such as principal components analysis (PCA) (i.e., Eigenfaces[1]) and Fisher's Linear Discriminant (FLD) (i.e., Fisherface[2]). The second type includes local methods, which illustrate local patches of an image in detail and then combine statistical information about each patch to form a new vector[3]. The local binary pattern (LBP)[4], [5], and local derivative pattern (LDP)[6] are representative examples of local methods. To realize better performance in difficult circumstances, recently proposed methods combine multiple local features and even apply learning strategies. For example, Vu *et al.* presented patterns of oriented edge magnitudes (POEM) on the basis of the oriented magnitudes in LBP[7]. Lu *et al.* achieved a compact binary face descriptor (CBFD) feature learning method[3]. Moreover, the scale invariant feature transform (SIFT) which extracts distinctive features has achieved great results in matching objects in different views.

Of the current face recognition schemes, few methods are effective against uncontrolled circumstances while maintaining low complexity. Furthermore, few approaches focus on depicting facial profiles in a straightforward manner. Due to its complicated structure, an object can be easily recognized by its shape. For example, Zhang *et al.* proposed a binary local descriptor, called the Edge-SIFT, for mobile searches[10]. However, face images pose difficulty because they have fewer contours with high contrast. Besides, SIFT can also remove interest points along edges[8], which results in the loss of some representative features. Moreover, illumination variation greatly influences SIFT feature detection.

In this paper, we propose an effective, simple, and fast descriptor, the local binary pattern (LBP) edge-mapped descriptor, for face recognition. This descriptor is a string of binary codes that concentrates on describing an individual's profile on the basis of edge-based features, maxima of gradient magnitude (MGM)[11] points which are easily detected despite the presence of illumination variation. Using binary codes, we apply the proposed matching method, obtain acceptable recognition rates, and save much execution time.

The rest of the paper is organized as follows. In section 2, we present the overall system and describe in detail the proposed descriptor and matching method. Then, in section 3 we present the experimental results of our proposed method and compare them with those of other methods. Finally, we draw our conclusions in section 4.

2. LOCAL BINARY PATTERN (LBP) EDGE-MAPPED DESCRIPTOR FOR FACE RECOGNITION

Fig. 1 shows a block diagram of our proposed LBP edge-mapped based system, which comprises four parts: preprocessing, MGM feature detection[11], the proposed LBP edge-mapped descriptor, and the proposed matching method. First, we pre-process a probe image to reduce the noise and enhance contrast. After extracting the interest points via the MGM proposed by Faraji *et al.*[11], we apply our proposed method to describe them.

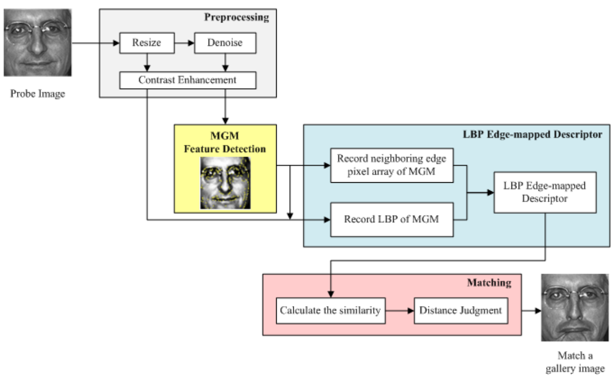

Fig. 1. Block diagram of the proposed LBP edge-mapped based face recognition system.

2.1. LBP Edge-mapped Descriptor

The LBP edge-mapped descriptor is a kind of binary code that illustrates the neighboring illumination and edges information in a proper region of the MGMs[11]. Fig. 2 shows the framework of the proposed descriptor, for which there are three main steps. Given an MGM image I_{MGM} and an edge image I_{edge}, we first record the surrounding edge pixel array of each MGM point through an $n \times n$ edge mask, and obtain the string of codes, $Edge\_pixel\_array$, including n^2 bits.

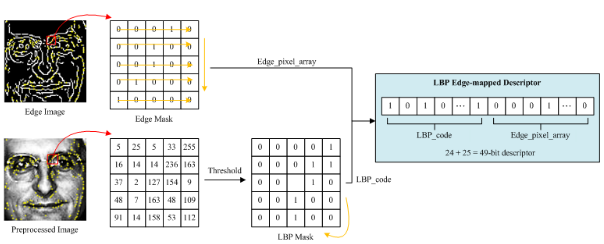

Fig. 2. Framework of LBP edge-mapped descriptor: example using a 5×5 LBP mask and a 5×5 edge mask.

Next, we extract the LBPs from the preprocessed image with MGM points. For each pixel of an image, the primary LBP[4] regards the value of the center pixel in a mask as a threshold, and compares it with its 3×3 neighborhood, generating an 8-bit binary number. For varying the range of the neighborhood, we directly extend the original LBP in our method. To do so, we record all the compared results within an $m \times m$ LBP mask for each feature. Each element of the LBP_code of an MGM point p can be derived:

$$LBP\_code(p, ne) = T(diff_{p,ne}) = T(g_{ne} - g_p), \quad ne = 1, 2, ..., (m^2 - 1) \tag{1}$$

$$T(diff) = \begin{cases} 1, & \text{if } (diff) \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $T(\cdot)$ is the threshold function, and g_p and g_{ne} are the intensity levels of the MGM and its neighboring pixels in the LBP mask, respectively.

Lastly, the LBP edge-mapped descriptor of each feature concludes by directly combining the corresponding LBP code and the edge pixel array, as shown in Eq. (3), in which there are $(m^2 + n^2 - 1)$ bits that depend on the needs of masks in different conditions. The overall process of the proposed LBP edge-mapped descriptor is detailed in Algorithm 1.

$$LBP\_edge\_mapped(p, :) = [LBP\_code(p, :), Edge\_pixel\_array(p, :)]. \tag{3}$$

Algorithm 1 LBP Edge-mapped Descriptor

Input: MGM image I_{MGM}, preprocessed image I_{pre}, edge image I_{edge}, $m \times m$ LBP mask, and $n \times n$ edge mask
Output: LBP edge-mapped descriptor LBP_edge_mapped

```
Find the coordinates (x, y) of all MGM points from I_MGM;
features = [x, y], num = length(features);
for i = 1 to num do
    Use the edge mask to find the edge pixel array of features(i, :) from I_edge;
    Get a string of codes, Edge_pixel_array(i, :);
    Use the LBP mask to derive LBP_code(i, :) using Eqs. (1) and (2) from I_pre;
    Derive LBP_edge_mapped(i, :) using Eq. (3);
end for
```

2.2. Matching Method

After obtaining the LBP_edge_mapped codes of all MGM points[11], next we use our proposed matching method to determine the best match for the probe image. First, we calculate the similarity scores, sim_score, of features by applying an XNOR gate and then summing entire binary codes:

$$sim\_score = sum(XNOR(LBP\_edge\_mapped(fe1, :), LBP\_edge\_mapped(fe2, :))) \tag{4}$$

$$fe1 = 1, 2, ..., num1,\ fe2 = 1, 2, ..., num2 \quad (5)$$

where $fe1$ and $num1$ are an MGM point and the number of MGMs from the probe image, while $fe2$ and $num2$ are an MGM point and the number of MGMs from the gallery image. After computing the similarity scores for all the features, we use Eqs. (6) and (7) to determine the possible matching pair, $match\_pair$, of the features:

$$match\_pair = \begin{cases} 1, & \text{if } max(sim\_score(fe1, fe2)) \geq th \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where the threshold, th, is a proper constant, which we define based on the total bits of the descriptor, $code\_size$:

$$th = 0.8 \times code\_size. \quad (7)$$

To eliminate the incorrectly matching pairs (Fig. 3(a)), we consider the location of each feature. We apply Eqs. (8) and (9) to discard those that are incorrect and obtain the final pair $final\_pair$. This whole process is known as distance judgment, and it can be expressed as follows:

$$final\_pair = \begin{cases} 1, & \text{if } (dis) < dis\_th \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$dis = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (9)$$

where (x_1, y_1) and (x_2, y_2) are the coordinates of the matching MGM points in the probe and gallery images respectively. Fig. 3 shows the effect of distance judgment for the unconstrained condition. It is apparent that no matter what situation we encounter, matching pairs that differ greatly will be removed following the execution of distance judgment.

Finally, we decide the best matching subject for the probe image by identifying the maximum number of $final\_pair$ among all the gallery images.

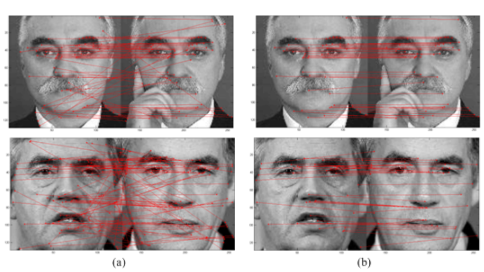

Fig. 3. Effect of distance judgment in the UFI Database (unconstrained conditions)[16]. (a) Before. (b) After.

3. EXPERIMENTAL RESULTS

In our experiments, we implemented the proposed method in C++ with OPENCV 3.0 and Matlab software on a computer with an Intel(R) Core(TM) i5-4200 CPU @ 1.60GHz 2.30GHz, 8GB RAM, 64-bit. The details of results and analysis are presented in the following sections.

3.1. FERET Database

The FERET Database[12], [13] contains many facial images in a variety of different conditions In this work, we regarded subset fa, which contains 1196 frontal images of 1196 individuals, as the gallery set, and subset fb, which includes 1195 individual images with diverse facial expression, and subset fc, comprising 194 people with one image per person under different illumination conditions, as probe sets.

Table 1 shows the recognition rates and execution times of the different methods under variable lighting. From the table, we can see that our method whose descriptor was composed of a 9×9 LBP mask and a 7×7 edge mask has the highest recognition rate. In addition, the execution time of our method is 9.06 times less than that of SIFT even though our descriptor size is the largest of these methods.

Table 2 shows the recognition rates and execution times using different algorithms for variation of facial expression. In this task, our descriptor consisting of a 9×9 LBP mask and a 9×9 edge mask outperformed other combinations.Of these five methods, the SIFT-based schemes have better performance (e.g., SIFT is 1.5% higher) than our method because different expression results in various contours. However, our approach is 7.50 times faster than SIFT.

Table 1. Recognition rates and execution times for different methods in the FERET database subsets *fc* vs. *fa*[13].

Enhancement	CLAHE[15]				
Features	131.66 MGM points[11]			266.41 SIFT points[8]	
Descriptor	LBP1 [4]	LBP2 [1]	LBP Edge -mapped	SIFT [8]	Module-based LBP with SIFT[9]
Descriptor size	8 bits	80 bits	129 bits	128 bits	72 bits
Recognition rate	19.1%	45.9%	54.1%	37.6%	28.9%
Execution time	0.1483	0.1641	0.1804	1.6346	9.5486
Time ratio	0.82	0.91	1	9.06	52.93

* Features: Averages based on the database subset fa.
1. Descriptor: 9×9 LBP mask.

Table 2. Recognition rates and execution times for different methods in the FERET database subsets *fb* vs. *fa*[13].

Enhancement	Histogram Equalization[14]				
Features	126.35 MGM points[11]			234.40 SIFT points[8]	
Descriptor	LBP1 [4]	LBP2 [1]	LBP Edge -mapped	SIFT [8]	Module-based LBP with SIFT[9]
Descriptor size	8 bits	80 bits	161 bits	128 bits	72 bits
Recognition rate	49%	78.4%	84.9%	86.4%	86.11%
Execution time	0.1524	0.1732	0.1971	1.4785	7.1935
Time ratio	0.77	0.88	1	7.50	36.50

*

Features: Averages based on the database subset fa.

3.2. The Extended Yale Face Database B

The Extended Yale Face Database B [17], [18] contains 38 frontal images of 38 individuals in 64 different illumination conditions. In this experiment, we regarded the condition, *P00A+00E+00*, as the gallery set and the other 63 conditions as probe sets.

Table 4 shows the average recognition rates and execution times for the five algorithms for the various illumination conditions. Our method, whose descriptor consisted of a 9×9 LBP mask and a 9×9 edge mask, yields the highest average recognition rate and is 3.44 times faster than SIFT. Fig. 4 shows the recognition rates of different methods in each individual lighting condition. We can see that most approaches attain nearly a 100% recognition rate in ordinary circumstances (e.g., from conditions 6 to 12 in Fig. 4), while LBP-based schemes including our method perform better than SIFT-based approaches in extreme conditions (e.g., from conditions 22 to 32 in Fig. 4).

Table 3. Recognition rates and execution times for different methods in the Extended Yale Face Database B[18].

Features	MGM[11]			SIFT[8]	
Descriptor	LBP1 [4]	LBP2 [1]	LBP Edge -mapped	SIFT [8]	Module-based LBP with SIFT[9]
Descriptor size	8 bits	80 bits	161 bits	128 bits	72 bits
Average recognition rate	49.21%	65.04%	67.42%	37.93%	33.42%
Average execution time	0.2046	0.2411	0.2931	1.0073	8.5074
Time ratio	0.70	0.82	1	3.44	29.03

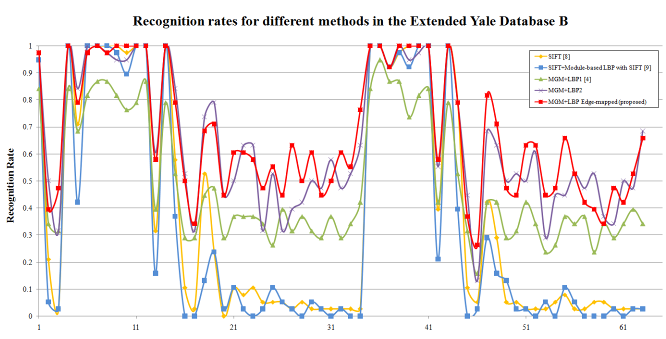

Fig. 4. Recognition rates for different methods in each condition in the Extended Yale Face Database B[18].

3.3. UFI Database

The Unconstrained Facial Images (UFI) Database[16] consists of real-world images extracted from a large set of photos owned by the Czech News Agency. There are two partitions in the UFI database: *Cropped images* and *Large images*, and we choose the cropped group to evaluate our algorithm in uncontrolled circumstances. The *Cropped images* dataset contains 4316 images of 605 subjects in the training set where each person has about 7.1 images on average, and in the probe set there are 605 images of 605 people.

Fig. 5 shows the recognition rates of different methods in uncontrolled situations, for which our proposed descriptor was composed of an 11×11 LBP mask and a 9×9 edge mask. We compared the performance of our method with LBP histogram sequences (LBPHS)[5], LDP histogram sequences (LDPHS)[6], POEM histogram sequences (POEMHS)[7], face specific LBP (FS-LBP)[19], and SIFT[8]. From Fig. 16, we can see that POEMHS[7] performs the best of all these algorithms, and SIFT also has a better recognition rate due to the invariant characteristic of the features. We also see that the recognition rate of our method, MGM+LBP Edge-mapped, is 0.82% higher than LDPHS[6].

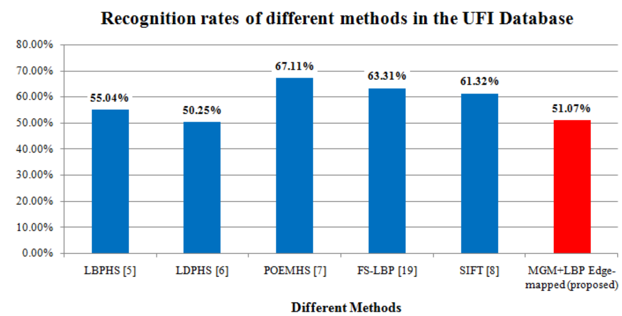

Fig. 5. Recognition rates of different methods in the UFI Database (*Cropped image* dataset)[16].

4. CONCLUSION

In this paper, we proposed a novel face recognition descriptor called the LBP edge-mapped descriptor that uses MGM[11] and a simple matching method for face recognition. The LBP edge-mapped descriptor can completely depict human contours while also maintaining low computational complexity. We experimentally verified the performance of our method in three respects: variable illumination, different expression, and real-world situations. Under different lighting conditions, our method performed better than SIFT-based approaches. Our method also proved to be 9.06 times and 3.44 times faster than SIFT[8] in the FERET database subset fc[13] and the Extended Yale Face Database[18], respectively. In addition, our method yielded acceptable recognition rates while requiring about 86.7% less execution time in comparison with SIFT for expression variation in the FERET database subset fb[13]. In uncontrolled conditions, our method yielded a 0.82% higher recognition rate than LDPHS[6] in the UFI Database[16].

5. REFERENCES

[1] M. Turk and A. Pentland, “Eigenfaces for Recognition,” *Journal of Congnitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman,“Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, Jul. 1997.

[3] J. Lu, V. E. Liong, X. Zhou, and J. Zhou, “Learning Compact Binary Face Descriptor for Face Recognition,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 37, no. 10, pp. 2041-2056, Oct. 2015.

[4] T. Ojala, M. Pietikäinen, and D. Harwood, “A Comparative Study of Texture Measures with Classification Based on Feature Distributions,” *Pattern Recognition*, vol. 29, no. 1, pp. 51-59, 1996.

[5] T. Ahonen, A. Hadid, and M. Pietikäinen, “Face Description with Local Binary Patterns: Application to Face Recognition,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[6] B. Zhang, Y. Gao, S. Zhao, and J. Liu, “Local Derivative Pattern Versus Local Binary Pattern: Face Recognition with High-Order Local Pattern Descriptor,” *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 533-544, Feb. 2010.

[7] N. S. Vu, H. M. Dee, and A. Caplier, “Face Recognition Using The POEM Descriptor,” *Pattern Recognition*, pp. 2478-2488, 2012.

[8] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, Jan. 2004.

[9] Y. K. Shen and C. T. Chiu, “Local Binary Pattern Orientation Based Face Recognition,” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1091-1095, Apr. 2015.

[10] S. Zhang, Q. Tian, K. Lu, Q. Hung, and W. Gao, “Edge-SIFT: Discriminative Binary Descriptor for Scalable Partial-Duplicate Mobile Search,” *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2899-2902, Jul. 2013.

[11] M. Faraji, J. Shanbehzadeh, K. Narollahi, and T. B. Moeslund, “Extremal Regions Detection Guided by Maxima of Gradient Magnitude,” *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5401-5415, Dec. 2015.

[12] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, “The FERET Database and Evaluation Procedure for Face-recognition Algorithms,” *Image and Vision Computing J*, vol. 16, no. 5, pp. 295-306, 1998.

[13] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, “The FERET Evaluation Methodology for Face-Recognition Algorithms,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090-1104, Oct. 2000.

[14] Rafael C. Gonzalez and Richard E. Woods. (2008). *Digital Image Processing: An Adapted Version* (3^{rd} ed.). Taiwan: Pearson Education Taiwan.

[15] K. Zuiderveld, “Contrast Limited Adaptive Histogram Equalization,” *Graphic Gems IV*, Academic Press, pp. 474-485, 1994.

[16] L. Lenc and P. Král, “Unconstrained Facial Images: Database for Face Recognition under Real-word Conditions,” *Mexican International Conference on Artificial Intelligence (MICAI)*, pp. 349-361, Oct. 2015.

[17] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, “From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643-660, Jun. 2001.

[18] K. C. Lee, J. Ho, and D. J. Kriegman, “Acquiring Linear Subspace for Face Recognition under Variable Lighting,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 684-698, May 2005.

[19] L. Lenc and P. Král, “Automatically Detected Feature Positions for LBP Based Face Recognition,” *Artificial Intelligence Applications and Innovations*, pp. 246-255, 2014.