EXPLORING FEATURE DESCRIPTORS FOR FACE RECOGNITION

[1]*Shuicheng Yan,* [2]*Huan Wang,* [2]*Xiaoou Tang, and* [1]*Thomas Huang*

[1]ECE Department, University of Illinois at Urbana Champaign, USA
[2]Department of Information Engineering, Chinese University of Hong Kong, Hong Kong

ABSTRACT

How to encode a face is a widely studied problem in both pattern recognition and psychology literatures. Many feature descriptors, Gabor feature, Local Binary Pattern (LBP), and Edge Orientation Histogram, have been proposed. In this paper, we give a comprehensive study of these descriptors under the framework of Principal Component Analysis (PCA) followed by Linear Discriminant Analysis (LDA), compared on three different popular similarity measures and two different feature correspondence strategies: holistic and local. Moreover, we present a new feature descriptor named Multi-Radius LBP, and also propose a combination scheme for the LBP and Gabor descriptor. The experiments on the Purdue and CMU PIE databases demonstrate that 1) an obvious recognition boost of LBP is achieved under PCA+LDA framework compared to the direct NN classification; 2) the LBP and Gabor features are comparable as well as mutually complementary, and the combination of these two descriptors brings a significant improvement in classification capability over single ones; and 3) the Multi-Radius LBP shows to outperform all the state-of-the-art feature descriptors.

Index Terms— Feature Descriptor, Similarity Measure.

1. INTRODUCTION

The success of a face recognition algorithm greatly relies on: 1) how to extract effective features to describe an image, and 2) how to infer the similarity of two faces based on the extracted features. Most previous studies focus on the latter part, such as Eigenface [2], Fisherface [2], and view-based recognition approaches [3]. They are all based on the original gray-level values and present different similarity measures based on subspace techniques.

The original gray-level feature often suffers from the illumination and expression variations; and many approaches have been proposed to extract more robust or semantic features from images. Among them, Gabor feature is the most popular one, and the Elastic Bunch Graph Matching (EBGM) [4] method has successfully integrated Gabor features and the local correspondence strategy for the multi-view face recognition problem. Combined with unified subspace analysis [10][12], it has also been used to improve indoor and outdoor face recognition [11]. Recently, Local Binary Pattern (LBP), originally introduced for texture representation, has proved to be a powerful descriptor for face recognition [1]. The EOH descriptor computes the histogram of the edge orientation distribution within the neighborhood of a point; and the work in [5] shows that EOH feature can significantly improve the face detection performance in comparison to the original gray-level feature. In this paper, we apply the EOH descriptor for face recognition.

The work in [1] shows that LBP based on direct Nearest Neighbor Classifier does not always produce satisfying results. In this work, we propose to utilize the framework of Principal Component Analysis (PCA) followed by Linear Discriminant Analysis (LDA) for LBP based face recognition, in which all possible PCA and LDA dimension combinations [10] are explored. A dramatic recognition boost is observed based on this new framework. Also, we propose the Multi-Radius LBP representation for face recognition and give a comprehensive study with the feature descriptors Gabor, LBP, EOH as well as the original gray-level features. These feature descriptors are compared with two different feature correspondence methods, *i.e.* holistic and local. The former one constructs the feature correspondence directly based on the image coordinates while the latter one is based on a set of key feature points, such as the mouth and eye corners, nose, and face contour points. Moreover, three different similarity measures L1, L2, and Cosine are applied to extensively evaluate the effectiveness of these different descriptors.

Gabor feature and LBP characterize the property of local texture distributions in distinct ways. In this work, in addition to the comprehensive comparison, we evaluate their complementary property, and the experimental results on the CMU-PIE and Purdue databases show that the combination of them brings significant performance improvements.

2. REVIEW OF THE FEATURE DESCRIPTORS

In this section, we give an overview of the state-of-the-art feature descriptors for the face recognition problem.

2.1. Gabor Feature

Gabor descriptors are harmonic functions modulated by gaussian distributions. A family of Gabor kernels is the product of a Gaussian envelope and a plane wave [6], defined as

$$\psi_{\vec{k}}(\vec{x}) = \frac{\|\vec{k}\|}{\delta^2} \cdot e^{-\frac{\|\vec{k}\|^2 \cdot \|\vec{x}\|^2}{2\delta^2}} \cdot \left[e^{i\vec{k}\cdot\vec{x}} - e^{-\frac{\delta^2}{2}} \right]. \qquad (1)$$

Here $\vec{x} = (x, y)$ is the position vector in spatial domain; the frequency vector $\vec{k}$ determines the scale as well as the orientation of Gabor kernels and is defined as

$$\vec{k} = k_s e^{i\phi_d}, \qquad (2)$$

where $k_s = k_{max}/f^s$, $\phi_d = \pi d/8$ and f is the spatial factor between kernels in the frequency domain. In all our experiments, five scales and eight orientations are used as reported in [6] with the following parameters: $k_{max} = \pi/2$, $f = \sqrt{2}$ and $\delta = 2\pi$.

2.2. Local Binary Pattern

The LBP descriptor utilizes binary pattern vectors to express local textures of image patches. The local neighborhood is first thresholded with the center value and then image patterns are converted into binary number vectors, which depict the texture distribution around the center. Among all the possible binary patterns, *uniform patterns*, which contain at most two bitwise circular transitions, are used to represent the concerned center point and a histogram distribution H of these local binary patterns is drawn to give a representation for a local patch S of a pattern labeled image I:

$$H_i = \sum_{(x,y)\in S} \delta(I(x,y), i) / \sum_{(x,y)\in S} 1, \quad i = 1, ..., n \qquad (3)$$

where n is the number of histogram bins and $\delta(z, z')$ is the Kronecker delta function $\delta(z, z') = 1$, if $z = z'$; 0, otherwise.

2.3. Edge Orientation Histogram

EOH descriptor [5] extracts statistical features from the orientation histogram of edges within local image patches. The input images are first convolved with Sobel masks $Sobel_x$ and $Sobel_y$ as

$$G_x(x, y) = Sobel_x * I(x, y), \qquad (4)$$

$$G_y(x, y) = Sobel_y * I(x, y). \qquad (5)$$

Then, the edge orientation angle is computed as

$$\theta(x, y) = tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right). \qquad (6)$$

Finally, an image is divided into small patches, and a histogram of the edge orientation is drawn to represent the statistical characteristic of a local patch. We use eight orientation bins for computing the histogram and the local patch size is fixed as 3-by-3 pixels, from which the best results are obtained in our experiments.

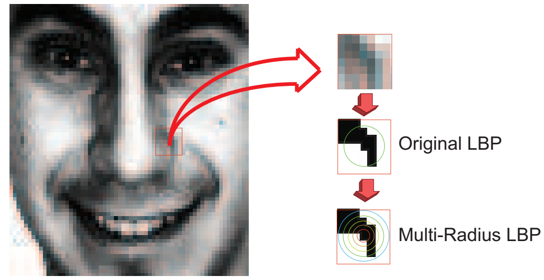

Fig. 1. Illustration of the LBP and Multi-Radius LBP descriptor. A neighborhood is extracted, thresholded with the center value and converted into binary numbers according to the binary distributions.

3. MULTI-RADIUS LOCAL BINARY PATTERN

LBP has been verified to be effective for face recognition, yet for each point, LBP characterizes the property of its neighboring points lying on the circle with fixed radius, hence LBP may be sensitive to the image scale.

A natural way to amend this issue is to explore the properties in multiple scales. In this work, we propose to use the Multi-radius LBP to increase the algorithmic robustness, and argue that the texture distributions over different scales are equally important. We calculate LBP at circles of different radiuses and a 3-D histogram is drawn as a description of the local patch from the pattern labeled images I^r, $r = 1, 2, ..., N_r$, where N_r is the number of radiuses explored.

$$H_i^r = \sum_{(x,y)\in S} \delta(I^r(x,y), i) / \sum_{(x,y)\in S} 1, \quad i = 1, ..., n. \qquad (7)$$

In our experiments, we employ altogether five different radiuses: 1, 1.5, 2, 2.5, and 3 pixels, and the pixel values are bilinearly interpolated whenever the coordinates of sampling points are not integers. We examined various patch sizes and selected the size 8-by-8 and 3-by-3 pixels for the holistic and local region based strategies respectively. Fig. 1 shows the evolvement from LBP to the multi-radius LBP descriptor.

4. HOLISTIC VS LOCAL CORRESPONDENCE

Previous algorithms for 2D face recognition can be roughly divided into two categories according to the applied feature correspondence strategies, holistic or local correspondence. Holistic method regards the entire image as one entity and constructs the feature correspondence directly based on the image coordinates; and holistic method has achieved great success in the cases with controlled pose and expression variations while their performance will be greatly degraded if large pose and expression variations exist [3].

Local correspondence based algorithm extracts features from the specific key points; and then the features are extracted around these points. The most popular local correspondence based algorithm is Elastic Bunch Graph Matching (EBGM) [4]. Local correspondence based algorithms can well handle the pose and expression variations, yet they often need an accurate key point localization method. In this work, we apply the automatic key point localization algorithm proposed in [7] to locate four key points of a face, *i.e.* the center of two eyes, nose tip and mouth center point. Based on these fiducial points, twelve local patches are designed as in Fig. 2 for the final feature extraction.

Fig. 2. Local patch positions piloted by key points

5. EXPERIMENTS

Recently a series of applications of LBP and above-mentioned descriptors for face recognition have been presented; nevertheless, no detailed comparison of these descriptors has ever been made. In this section, we give a comprehensive study of them along with the proposed Multi-Radius LBP, compared with three different similarity measures and on two distinct feature correspondence strategies, holistic vs local. All the experiments are conducted by utilizing the PCA+LDA framework [10], *i.e.* exploring all possible PCA and LDA dimension combinations, and then Nearest Neighbor method is used for final classification.

Table 1. Holistic-based recognition rates (%) of different feature descriptors on Purdue

Similarities	Raw	EOH	Gabor	LBP	M LBP
L1	68.61	69.44	78.61	79.17	**81.67**
L2	69.72	71.11	79.17	80.28	**81.11**
Cosine	71.67	80.28	86.39	84.44	**86.67**
LBP+Gabor&Cosine		87.50	LBP & NN		78.89

5.1. Experiment Configurations

To extensively examine the performance of different feature descriptors, the popular Purdue databases [8] with 90 persons and CMU-PIE [9] with 68 persons are used in our experiments. The experimental configures on training, gallery and probe sets for two databases are as follow. For Purdue database, two images each person, one neutral expression and one smile expression in the first session, are used for model training. The first image with neutral expression in the first session is used for the gallery set, and other images are used for the probe set. For the CMU-PIE database, we focus on the multi-view face recognition issue. Four images each person are selected in our experiments. Two images each person, one frontal face (pose-11, illumination-12) and one profile face (pose-09, illumination-08), are selected for the model training; and two other images each person at pose-05, illumination-07 and at pose-11, illumination-05 are selected for gallery and probe set respectively. All the face images are normalized by translation, rotation and scaling, such that the centers of two eyes are in fixed positions of an image in size of 64-by-64 pixels; finally, the Histogram Equalization method is applied for photometric normalization.

5.2. Results and Observations

Comparison of different feature descriptors. The performances of Multi-Radius LBP, LBP, Gabor, EOH and raw gray-level features with local or holistic correspondences over different similarities are illustrated in Table 1-3 and Fig 3. From these results, we can see that the proposed Multi-Radius LBP feature descriptor steadily reaches the highest recognition rate, followed by Gabor and LBP. The recognition rate of EOH is lower; yet still higher than that of the raw image data in most cases, since certain amount of noise is removed from the edge extractor of EOH. We can also observe that when there exist pose variations, the performance of LBP without training as in [1], referred to as *LBP & NN* in the result tables, is not acceptable, and the training step gives a significant improvement for LBP in all the cases.

Comparison of similarity measures and two correspondence methods. The experimental results show that the Cosine ($< x, y > /(\|x\|.\|y\|)$) similarity is generally superior to L1 ($\sum_i |x_i - y_i|$) and L2 ($\sqrt{\sum_i |x_i - y_i|^2}$) similarities. Also, the recognition rate of the local region based method on CMU-PIE database is significantly superior to that of the holistic-based one. The holistic method may easily fail when large pose variations exist, due to the lack of explicit semantic correspondence; while the local region based method is more robust to pose variations. For Purdue database where no obvious pose variations exist, when the local patch size is large enough, the local-based method should be similar to holistic-based one; hence, we ignore the local-based method for Purdue database.

Combination of Gabor and LBP. The LBP descriptor gives a condense representation on the texture distribution under a certain scale within the selected patches, while the Gabor descriptor describes the multi-scale and multi-orientation distribution around the neighborhood of each point. Our experimental results reveal the complementary property of these two feature descriptors. Under the framework of PCA+LDA and Cosine similarity, we combine the similarity score matrix of LBP and Gabor and obtain a significant performance improvement in all the cases.

Table 2. Local-based recognition rates (%) of different feature descriptors on CMU-PIE

Similarities	Raw	EOH	Gabor	LBP	M-LBP
L1	83.87	82.26	91.94	91.94	**96.77**
L2	87.10	91.94	95.16	96.77	**98.39**
Cosine	90.32	87.10	96.77	96.77	**100**
LBP+Gabor&Cosine	100	LBP & NN			29.03

Table 3. Holistic-based recognition rates (%) of different feature descriptors on CMU-PIE

Similarities	Raw	EOH	Gabor	LBP	M LBP
L1	40.32	40.32	66.13	51.61	**69.35**
L2	41.94	43.55	70.97	56.45	**72.58**
Cosine	48.39	61.29	74.19	72.58	**87.10**
LBP+Gabor&Cosine	90.32	LBP & NN			6.45

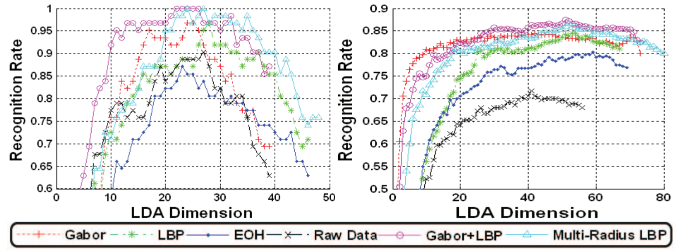

Fig. 3. Recognition rates of different feature descriptors under Cosine similarity over different LDA dimensions for Local-based CMU PIE database (Left) and Holistic-based Purdue database (Right). Note that the PCA dimensions are the ones with the best recognition rates.

6. CONCLUSIONS AND FUTURE WORKS

In this paper, we presented a comprehensive study of popular feature descriptors by considering two factors, *i.e.* feature correspondence methods and similarity measures. Moreover, our proposed Multi-Radius extension of LBP descriptor has shown to outperform all the state-of-the-art feature descriptors. The experiments also demonstrated that the LBP and Gabor features are comparable; meanwhile they are mutually complementary, and the combination of them gives an encouraging improvement of the face recognition performance.

In this work, we focus on the robustness of different feature descriptors to the pose variation, and we plan to give a more comprehensive study of these feature descriptors on larger face databases and by taking more factors, such as illumination and resolution variations, into consideration in our future work.

7. ACKNOWLEDGEMENT

The work of Shuicheng Yan and Thomas Huang was supported in part by the Disruptive Technology Office VACE III Contract NBCHC-060160 issued by DOI-NBC, Ft. Huachuca, AZ; and in part by the National Science Foundation Grant CCF 0426627. The views and conclusions are those of the authors, not of the US Government or its Agencies.

8. REFERENCES

[1] Ahonen T, Hadid A, and Pietikainen M, “Face recognition with local binary patterns,” in *ECCV*, 2004.

[2] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 2002.

[3] David J. Beymer, “Face recognition under varying pose,” Tech. Rep. AIM-1461, 1993.

[4] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg, “Face recognition by elastic bunch graph matching.,” in *ICIP (1)*, 1997, pp. 129–132.

[5] Kobi Levi and Yair Weiss, “Learning object detection from a small number of examples: The importance of good features.,” in *CVPR (2)*, 2004, pp. 53–60.

[6] C. Liu and H. Wechsler, “Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition,” *IEEE Trans. Image Processing*, 2002.

[7] Longbin Chen, Lei Zhang, Long Zhu, Mingjing Li, and Hongjiang Zhang, “A novel facial feature localization method using probabilistic-like output.,” in *ACCV04*, 2004.

[8] A.M. Martinez and R. Benavente, “The ar face database,” *CVC Technical Report #24*.

[9] T. Sim, S. Baker, and M. Bsat, “The cmu pose, illumination, and expression (pie) database of human faces,” *Tech. Report CMU-RI-TR-01-02, Robotics Institute, Carnegie Mellon University*, 2001.

[10] X. Wang and X. Tang, “A unified framework for subspace face recognition,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, pp. 1222–1228, 2004.

[11] X. Wang and X. Tang, “Improving indoor and outdoor face recognition using unified subspace analysis and Gabor features,” *Proc. IEEE Intl Conf. Image Processing (ICIP)*, vol. 3, pp. 1983-1986, Oct. 2004.

[12] X. Wang and X. Tang, Unified subspace analysis for face recognition, *Proc. Ninth IEEE Intl Conf. on Computer Vision (ICCV)*, pp. 679-686, Oct. 2003.