Multi-resolution Histograms of Local Variation Patterns (MHLVP) for Robust Face Recognition

Wenchao Zhang[1], Shiguang Shan[2], Hongming Zhang[1], Wen Gao[2], and Xilin Chen[1,2]

[1] School of Computer Science and Technology, Harbin Institute of Technology, 150001 Harbin, P.R. China
{wczhang,hmzhang,xlchen}@jdl.ac.cn
[2] ICT-ISVISION Joint R&D Laboratory for Face Recognition, CAS, 100080 Beijing, P.R. China
{sgshan,wgao}@jdl.ac.cn

Abstract. This paper presents a novel approach to face recognition, named Multi-resolution Histograms of Local Variation Patterns (MHLVP), in which face images are represented as the concatenation of the local spatial histogram of local variation patterns computed from the multi-resolution Gabor features. For a face image with abundant texture and shape information, a Gabor feature map(GFM) is computed by convolving the image with each of the forty multi-scale and multi-orientation Gabor filters. Each GFM is then divided into small non-overlapping regions to enhance its shape information, and then Local Binary Pattern (LBP) histograms are extracted for each region and concatenated into a feature histogram to enhance the texture information in the specific GFM. Further more, all of the feature histograms extracted from the forty GFMs are further concatenated into a single feature histogram as the final representation of the given face image. Eventually, the identification is achieved by histogram intersection operation. Our experimental results on FERET face databases show that the proposed method performs terrifically better than the performance of some classical results including the best results in FERET'97.

1 Introduction

As one of the most successful applications of image analysis and understanding, face recognition has received significant attention in both the wide range of potential applications [1] and scientific challenges [2][3]. Although many commercial face recognition systems have emerged, it is still an active topic in computer vision community. This is partly due to the fact that face recognition is still very challenging in uncontrolled environments with rich variations of pose, illumination etc. Therefore, the goal of the on-going research is to increase the robustness of face recognition systems to these variations. Moreover, to evaluate the effectiveness of different algorithms of face recognition, some evaluation methodologies and face databases have been created, such as the FERET database and protocol [4], which attract more and more researchers to make the further progress on algorithm.

Most of the prevalent approaches for face recognition are based on statistic analysis, such as eigenfaces [5], Fisherfaces [6][7] and Bayesian methods [8]. The methods often need a great deal of data to train, but usually the data available for training are very few. In addition, if the distributions of the test examples are different from that of the training examples, the generality of the method will be weakened.

T. Kanade, A. Jain, and N.K. Ratha (Eds.): AVBPA 2005, LNCS 3546, pp. 937–944, 2005.
© Springer-Verlag Berlin Heidelberg 2005

In this paper, we propose a novel approach for face recognition based on non-statistical model, which named Multi-resolution Histograms of Local Variation Patterns (MHLVP), in which face images are represented as the concatenation of the local feature histogram of multi-resolution Gabor features.

Histograms have been widely used to represent, analyze and characterize images because they could be computed easily and efficiently and their robustness to noise and local image transformations [9]. But they do not capture spatial image information. The multi-resolution histograms [10], however, not only preserves the efficiency, simplicity and robustness of the plain histogram, but also combines intensity with spatial information. Moreover, the Gabor filters combine the spatial and frequency localization [11], which are effective to face image representations [12]. In this work, given a face image, we get forty GFMs computed by convolving the image with each of the multi-scale and multi-orientation Gabor filters. The multi-resolution histograms are extracted from the GFMs represented by the magnitude of the Gabor filters' response, which can provide a measure of the image's local properties [13]. For a face image with abundant texture and shape information, we divided the face image into some certain regions to improve the shape information of the face image represented by the multi-resolution histograms. Also we extracted LBP histograms [14][15] from the GFMs to improve the texture information. Thus, the MHLVP can be used as a texture and shape descriptor for face image.

This paper is organized as follows: In Section 2, the face description with multi-resolution Gabor representations is briefly described firstly, then the local feature histogram is introduced. Experimental design and results in Section 3. Finally the discussion and the conclusion can be found in Section 4.

2 Face Description with Multi-resolution Histograms of Local Variation Patterns

The overall framework of the proposed representation approached MHLVP-based is illustrated in Fig.1. In this method, a face image is modeled by the following procedure: (1) An input face image is normalized and transformed to obtain multiple GFMs in frequency domain by applying multi-scale and multi-orientation Gabor filters; (2) Each GFM is labeled with LBP operator; (3) Each transformation of the GFM is further divided into non-overlapping rectangle regions with specific size, and histogram is computed for each region; (4) All the histograms are concatenated to form the final histogram sequence as the model of the face. The following sub-sections will describe the procedure in detail.

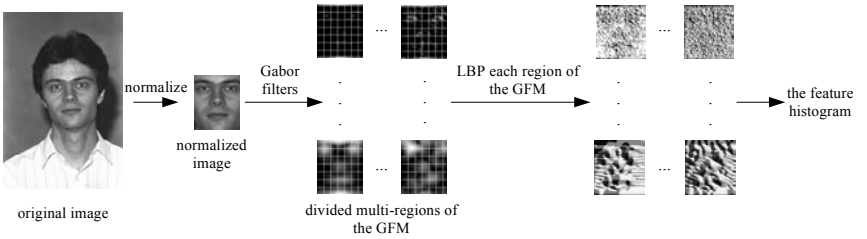


Fig. 1. The framework of Multi-resolution Histograms of Local Variation Patterns

2.1 Face Description with Multi-resolution Gabor Representations

The Gabor feature is effective to face image representation. The multi-resolution description of an image is computed with Gabor filters. The Gabor wavelets (filters, kernels) can be defined as follows, assuming that $\sigma_x = \sigma_y = \sigma$ [11][13]:

$$\Psi(x,y,\varpi_0,\theta) = \frac{1}{2\pi\sigma^2} e^{-\left((x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2\right)/2\sigma^2} \times \left[e^{i(\varpi_0 x\cos\theta+\varpi_0 y\sin\theta)} - e^{-\varpi_0^2\sigma^2/2}\right], \quad (1)$$

where x, y define the pixel position in the spatial domain, ϖ_0 the radial center frequency, θ the orientation of the Gabor wavelet, and σ the standard deviation of the Gaussian function along the x-and y-axes. In addition, the second term of the Gabor wavelet, $e^{-\varpi_0^2\sigma^2/2}$, compensates for the DC value because the cosine component has nonzero mean (DC response) while the sine component has zero mean.

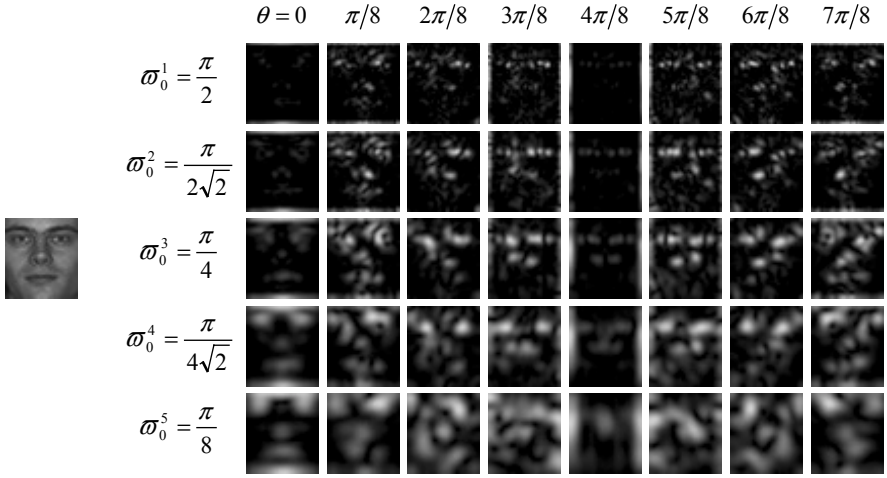
The Gabor representation of an image can be derived by convolving the image and the Gabor wavelets. Let $f(x,y)$ denote the gray level distribution of an image, then the convolution of an image $f(x,y)$ and a Gabor kernel $\Psi(x,y,\varpi_0,\theta)$ is defined as follows:

$$G_{\Psi f}(x,y,\varpi_0,\theta) = f(x,y) * \Psi(x,y,\varpi_0,\theta), \quad (2)$$

where $*$ denotes the convolution operator. With a set of ϖ_0 and θ, a multi-hierarchical Gabor wavelets representation of the face image $f(x,y)$ is composed. The visualizations of the magnitudes of the GFMs in this paper are shown in Figure 2. We select five scales and eight orientations in the Gabor filters.

2.2 Local Variation Patterns

To improve the local feature information, we divided the image into small non-overlapping regions from which LBP histograms are extracted.



(a) Face image (b) The GFMs

Fig. 2. The visualizations of the magnitudes of the GFMs

The original LBP operator, introduced in [14], labels the pixels of an image by thresholding the 3×3 -nieghborhood of each pixel $f_p(p=0,1,\ldots,7)$ with the center value f_c and considering the result as a binary number (3).

$$S(f_p - f_c)=\begin{cases}1, & f_p \geq f_c \\ 0, & f_p < f_c\end{cases}. \tag{3}$$

Then, by assigning a binomial factor 2^p for each $S(f_p - f_c)$, the LBP number is achieved as

$$LBP = \sum_{p=0}^{7} S(f_p - f_c)2^p, \tag{4}$$

which characterizes the spatial structure of the local image texture. Figure 3 shows an example of the basic LBP operator, and the transform result of an image is shown in Figure 4. The LBP images of the GFMs are shown in Figure 5.

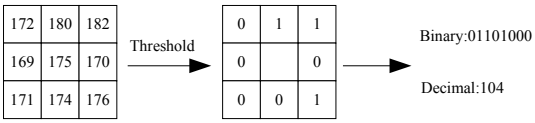


Fig. 3. Example of the basic LBP operator

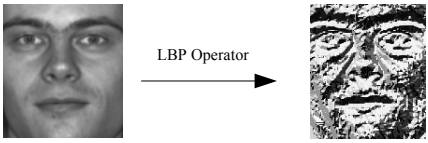


Fig. 4. The face image and the LBP image

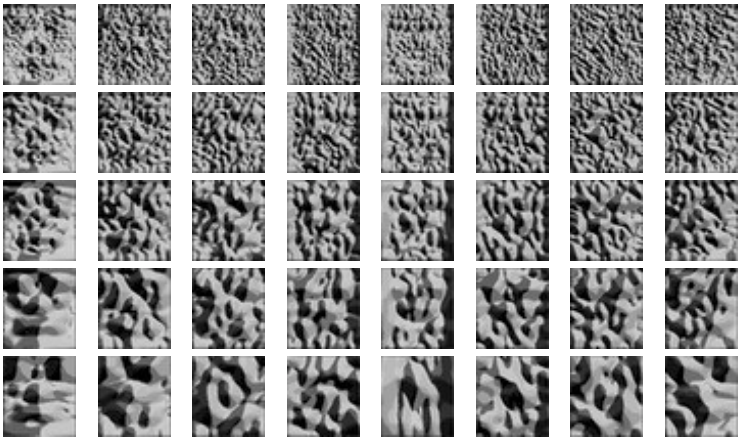


Fig. 5. The LBP images of the GFMs

2.3 Multi-resolution Histograms of Local Variation Patterns Description

The histogram h_f of the image $f(x,y)$ could be defined by (5) which provides the frequency of different values in the image.

$$h_{i,f} = \sum_{x,y} I\{f(x,y) = i\}, i = 0, 1, \ldots, n-1, \tag{5}$$

n is the number of different values of the image and

$$I\{A\} = \begin{cases} 1, A \text{ is true} \\ 0, A \text{ is false}. \end{cases} \tag{6}$$

From the description above, we can get the multi-resolution histograms of local variation patterns description, and each of the Gabor filter image $G_{\psi f}(x,y,\varpi_0,\theta)$ is divided into m regions of $R_0, R_1, \ldots, R_{m-1}$ from which the LBP histograms are extracted (7).

$$h_{i,j,G_{\psi f}} = \sum_{x,y} I\{G_{\psi f}(x,y,\varpi_0,\theta) = i\} I\{(x,y) \in R_j\} i = 0, 1, \ldots, n-1; \; j = 0, 1, \ldots, m-1. \tag{7}$$

Many dissimilarity measures have been proposed for histogram. In this paper, Histogram intersection $D(Sh_1, Sh_2)$ is used as the similarity matching between two histograms [17].

$$D(Sh_1, Sh_2) = \sum_{i=1}^{k} \min(Sh_1^i, Sh_2^i), \tag{8}$$

where Sh_1 and Sh_2 are two histograms, and k is the number of bins in the histogram.

3 Experimental Design and Results

To achieve a fair comparison, we test our face recognition algorithm on the FERET face database and protocol, which have been used widely to evaluate face recognition algorithms and are a de faceto standard in face recognition research field [20].

FERET provided a training set containing 1002 images, and a gallery consisted of images of 1,196 people with one image per person. Four probe sets, fb, fc, Dup.I, and Dup.II are provided for testing. In the fb probe set, there are 1195 images with different facial expressions from the gallery set. In the fc probe set, there are 194 images taken under different lighting condition from that of fa. In the probe categories Dup.I and Dup.II, there are 722 and 234 images taken respectively a month and a year later than the gallery.

To test the robustness of the method against different facial expression, lighting and aging, we do the experiments on all the four probe sets. In addition, to demonstrate the validity of our method, we compared the performance of our method with Correlation [18], LDA [19] and LBP. The similarity measure used in LDA and Correlation algorithms in our experiments to compare feature vectors is the normalized correlation, which is defined as (9).

$$\delta(\mathbf{x}_1, \mathbf{x}_2) = \frac{-\mathbf{x}_1^T \mathbf{x}_2}{\|\mathbf{x}_1\| \|\mathbf{x}_2\|}. \tag{9}$$

In the LBP method, we use the basic LBP operator and the same divided region form as that of the MHLVP in Figure 5. The nearest neighbor rule is then used to classify the face images. The recognition results of our proposed method are shown in Table 1 and the Cumulative Match Score (CMS) curves are plotted in Figures 6 (a)-(d). The MHLVP-based method outperforms clearly the other methods on all of the test sets. In the three test sets of fc, duplicate I and duplicate II, the results are even better than the best results reported in the FERET'97 evaluation [20] and that of [21]. It should be noted that the results of the basic LBP method might be different from the results mentioned in [21] due to the different forms of image division and LBP operators.

Table 1. The rank-1 recognition rates of different algorithms for the FERET probe sets

Method	fb	fc	duplicate I	duplicate II
MHLVP	0.942	0.959	0.676	0.594
LBP	0.947	0.294	0.536	0.269
LDA	0.934	0.727	0.551	0.312
Correlation	0.700	0.268	0.276	0.094
Best Results of FERET'97[20]	0.96	0.82	0.59	0.52

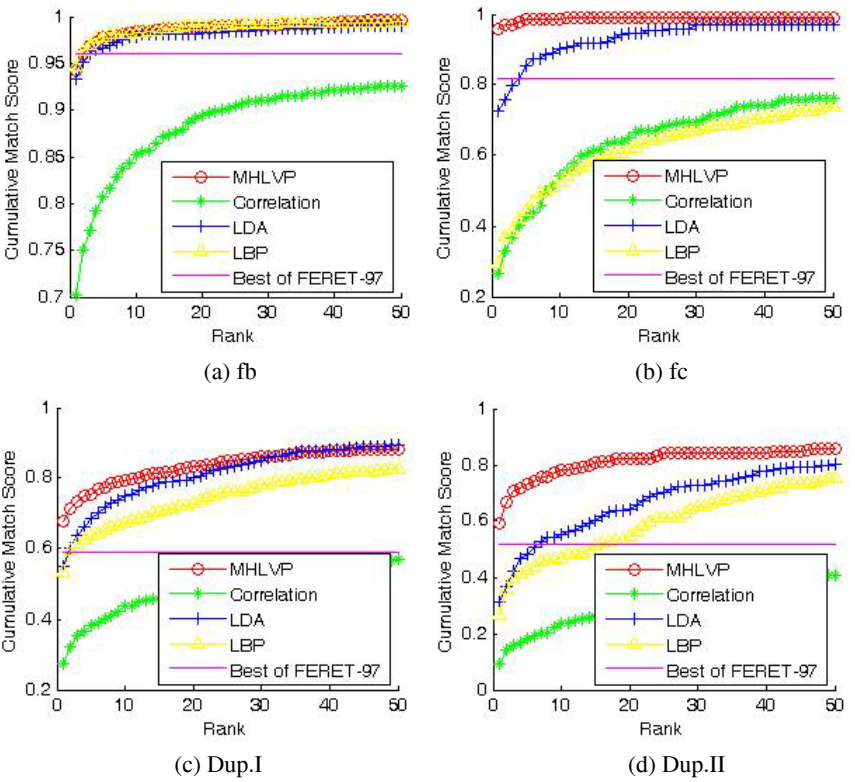


Fig. 6. Comparison of several methods on the FERET fb, fc, Dup.I, and Dup.II probe sets. Note that the FERET line just illustrated the best rank-1 recognition rate in the FERET'97 results

4 Discussion and Conclusion

This paper proposes a novel face representation, MHLVP, which is impressively insentative to appearance variations due to lighting, expression and aging. Moreover, the modeling procedure of MHLVP does not involve in any learning process, that is, it is non-statistical learning based. The effectiveness of the MHLVP comes from several aspects including the multi-resolution and multi-orientation Gabor decomposition, the LBP and the local spatial histogram modeling. The experimental results on FERET face database have evidently shown that the MHLVP method performs terrifically better than other approaches for all four standard probe sets. In the three test sets of fc, duplicate I and duplicate II, the results are even better than that of the best reported on the FERET evaluation.

Some improvements may further be made by optimizing the bins of the histogram to reduce the dimension the feature vector, and/or selecting the different scales and orientations Gabor filters to represent the shape and texture information of the face image. In addition, we could divide images into different amount of regions according to the variations of different scales.

Acknowledgements

This research is partially sponsored by Natural Science Foundation of China under contract No.60332010, "100 Talents Program" of CAS, ShangHai Municipal Sciences and Technology Committee (No.03DZ15013), and ISVISION Technologies Co., Ltd.

References

1. Phillips, P.J., Grother, P., Micheals, R.J., Blackburn, D.M., Tabassi, E., Bone, J.M.: Face recognition vendor test 2002 results. Technical report (2003)
2. Chellappa, R., Wilson, C.L., Sirohey, S.: Human and Machine Recognition of faces: A survey. Proc. of the IEEE 83 (1995) 705-740
3. Zhao, W.Y., Chellappa, R., Phillips, P.J. and Rosenfeld, A.: Face Recognition: A Literature Survey. ACM Computing Survey (2003) 399-458
4. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face recognition algorithms. Image and Vision Computing 16 (1998) 295-306
5. Turk, M. and Pentland, A.: Face recognition using eigenfaces. In Proc. IEEE Conference on Computer Vision and Pattern Recognition (1991) 586-591
6. Belhumer, P., Hespanha, P., and Kriegman, D.: Eigenfaecs vs fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (1997) 711-720
7. Swets, D.L. and Weng, J.: Using Discriminant eigenfeatures for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (1996) 831-836
8. Moghaddam, B., Nastar, C., Pentland, A.: A Bayesian similarity measure for direct image matching. In: 13th International Conference on Pattern Recognition (1996) II: 350-358
9. Hadjidemetriou, E., Grossberg, M.D. and Nayar, S. K.: Multiresolution Histograms and Their Use for Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (2004) 831-847
10. Hadjidemetriou, E., Grossberg, M. D., and Nayar, S. K.: Spatial Information in Multiresolution Histograms. Proc. Computer Vision and Pattern Recognition Conf. (2001) I: 702-709

11. Porat, M. and Zeevi, Y.: The Generalized Gabor Scheme of Image Representation in Biological and Machine Vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 10 (1988) 452-468
12. Liu, C.J., Wechsler, H.: Gabor feature based classification using the enhanced fisher linear discriminant modal for face recognition image processing. IEEE Transactions on Image Process 11 (2002) 467-476
13. Lades, M., Vorbrüggen, J.C., Buhmann, J., Lange, J., Malsburg, C., Würtz, R.P., Konen, W.: Distortion invariant object recognition in the dynamic link architecture. IEEE Transactions on Computers 42 (1993) 300-311
14. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 29 (1996) 51-59
15. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 971-987
16. Daugman, J.G.: Two-dimensional spectral analysis of cortical receptive field profile. Vision Research 20 (1980) 847-856
17. Swain, M., Ballard, D.: Color indexing. Int.J.Computer Vision 7 (1991) 11-32
18. Brunelli, R., and Poggio, T.: Face Recognition: Features vs. Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (1993) 1042-1053
19. Duda, R., Hart, P., Stork, D., Pattern Classification. Wiley Interscience, USA, Second Edition (2001)
20. Phillips, P.J. Syed, H.M., Rizvi, A. and Rauss, P.J.: The FERET Evaluation Methodology for Face-Recognition Algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (2000) 1090-1104
21. Timo, A., Abdenour, H. and Matti, P.: Face Recognition with Local Binary Patterns. ECCV 2004 Proceeding, Lecture Notes in Computer Science 3021, Springer (2004) 469-481