Ensemble of Piecewise FDA Based on Spatial Histograms of Local (Gabor) Binary Patterns for Face Recognition

Shiguang Shan[1], Wenchao Zhang[1,2], Yu Su[1,2], Xilin Chen[1], Wen Gao[1,2]
[1] ICT-ISVISION FRJDL, Institute of Computing Technology, CAS, Beijing, China
[2] School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
{sgshan, wczhang, ysu, xlchen, wgao}@jdl.ac.cn

Abstract

Spatial histogram of Local Binary Pattern (LBP) and Local Gabor Binary Pattern (LGBP) has been successfully applied to face recognition and achieved state-of-the-art performance. Both LBP and LGBP utilize traditional histogram matching method such as histogram intersection for face classification. In this paper, we propose a statistical extension for L(G)BP similarity computation by introducing Fisher Discriminant Analysis (FDA) of the L(G)BP spatial histogram "features". More than a simple application of FDA, we have constructed Ensemble of Piecewise FDA (EPFDA) classifiers, each of which is designed using one segment of the entire spatial histogram features. We show that this extension not only greatly reduces the feature dimension but also brings very impressive performance improvement. Especially, we have made a large step to recognizing all the faces in the standard FERET face database.*

1. Introduction

Face recognition has attracted more and more attention for its scientific values and wide potential applications. Much progress has been made in the last decade as surveyed in [1]. However, the general problem of FR remains unsolved, since most of the systems to date can only successfully recognize faces when images are obtained under constrained conditions. Their performance will degrade abruptly when face images are captured under varying lighting conditions, viewpoints, expressions, and partial occlusion.

To solve these bottlenecks simultaneously, appropriate representation modeling the intrinsic attribute of faces is one of the vital problems. Nowadays, representation approaches based on statistics or learning are undoubtedly absolutely dominant, typical methods including Eigenface [2], Fisherface [3], and Gabor Fisher Classifier (GFC) [4] have attracted sufficient attention. Recently, spatial histogram model of Local Binary Pattern (LBP) [5] have been proposed to represent visual objects, and successfully applied to texture analysis [5] and face recognition [6]. In LBP method, the representation of a face is directly derived from the image without any supervised training set involved, and classification is completed by histogram matching without any learning procedure as well. By using LBP, Ojala et al has reported impressive results on the FERET database [6].

Afterwards, without breaking LBP's non-statistical nature, we extended LBP to LGBP by introducing multi-orientation and multi-scale Gabor filtering before and impressively improved the performance when compared with the pure LBP [7]. By far, LGBP had reported the best results [7] on the FERET database.

As is well know, in LBP (/LGBP) for face recognition, a face image is encoded as multiple spatial histograms, and face classification is completed by combining the similarities between corresponding spatial histograms. In previous works, histogram matching methods have been exploited for computing histogram similarity, such as histogram intersection, Chi square, etc [5,6,7]. However, it is a little surprising for us that few researchers have exploited discriminant analysis on the spatial histograms, which is true at least for LBP-based face recognition.

This paper extends the classification of LBP histogram by exploiting statistical tools, such as Fisher discriminant analysis. More than a simple application of FDA, we have constructed ensemble of piecewise FDA (EPFDA) classifiers, each of which is designed based on one segment of the high-dimensional histogram features. Extensive experiments show that the proposed method impressively outperforms the traditional histogram intersection and Chi square. What is more, a large step has been made towards recognizing all the faces in the standard FERET database, which has been thought very challenging.

2. Previous Work

2.1 LBP and LGBP

Intuitively, LBP operator extracts the local variance features of an image, which are most important for classifying different face images. At each image position, the original 2-D LBP operator encodes the

*Spatial Histograms are defined as the histograms extracted from the image blocks spatially partitioning the whole image uniformly or with some specific rules, that is, one histogram is estimated from each image block.

The 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

ordinal relationship between the center intensity f_c and the intensity of its 8-neighbors $f_p\left(p=0,1,\ldots,7\right)$ as a binary number [5]

$$v_{LBP}(c)=\sum_{p=0}^{7} b\left(f_p-f_c\right)2^p , \qquad (1)$$

where

$$b\left(f_p-f_c\right)=\begin{cases}1, & f_p \geq f_c \\ 0, & f_p < f_c\end{cases} . \qquad (2)$$

From Equ.(1) and (2), one can notice that LBP actually encodes 256 modes of the gradient orientation at each image position in some sense (when not considering its rotation-invariant version). Therefore, LBP is a kind of *symbol vector* rather than feature vector, which means that traditional norm-based (e.g. Euclidean) distance is not appropriate for their classification. So, spatial histogram is generally exploited to further encode the LBP, in which face images are partitioned spatially into multiple rectangle regions and histogram is extracted from each region. Thus, a face image is finally represented as a number of spatial histograms, and classification is completed by computing the histogram similarity.

We had extended LBP to LGBP by introducing Gabor transform before LBP [7]. Simply speaking, in LGBP, LBP operator is conducted on the magnitudes of the multi-orientation and multi-scale Gabor filtered image rather than the original image intensity. Thus, a face is represented as more spatial histograms with varying orientations and scales. Therefore, LGBP can be regarded as a combination of Gabor feature with LBP. Its main merit over LBP lies in its capacity of modeling local features of varying orientations and scales, which is provided by the Gabor transform. LGBP has impressively improved the performance when compared with the original LBP and reported the best results on the FERET database by far [7].

2.2 Similarity Measurement for L(G)BP

As is mentioned above, both LBP and LGBP finally represent a face image as a number of spatial histograms and the similarity between two face images is the summation of the similarities between all the spatial histograms. Formally, the process is formulated as follows:

Assume that a face image I is finally represented as m spatial histograms and denoted by:

$$\Re=[H_1,H_2,\cdots,H_m] , \qquad (3)$$

where H_i is the *i*-th spatial histogram. Then similarity between two face images I and I' is computed by:

$$S(I,I')=S(\Re,\Re')=\tfrac{1}{m}\sum_{i=1}^{m} S(H_i,H_i') \qquad (4)$$

where $S(H_i,H_i')$ denotes the histogram similarity computed using histogram intersection etc. We will then introduce two commonly used histogram matching methods in what follows.

Let a histogram have L bins and the frequency for the *b*-th bin be $h(b)$, then the *histogram intersection* is defined as:

$$S_{HI}\left(H,H'\right)=\sum_{b=1}^{L}\min\left(h(b),h'(b)\right) \qquad (5)$$

and *Chi square* distance is defined as:

$$S_{Chi}\left(H,H'\right)=\sum_{b=1}^{L}\frac{\left(h(b)-h'(b)\right)^2}{h(b)+h'(b)} \qquad (6)$$

3. EPFDA Classifiers Based on L(G)BP

In this section, we first summarize how we come up with the idea of FDA on the spatial histograms of L(G)BP, and then we introduce how piecewise FDA (hereinafter EPFDA) classifier is constructed in details.

3.1 Motivation and Basic Ideas

The first motivation comes from the nature of face recognition, which evidently needs discriminating features rather than pure representation. Therefore, we need further discriminant analysis on the spatial histogram representation in order to improve the recognition performance. FDA is a natural choice.

The other motivation lies in the requirement to reduce the dimensionality, especially for LGBP whose dimension is 40 times that of the LBP. Obviously, such a high dimensional representation not only is terrible for storage requirement but also may cause serious curse of dimensionality. Again, FDA is good choice for this purpose.

However, a direct application of FDA to the holistic high-dimensional histogram seems too straightforward. Moreover, due to small sample size problem, given N training samples of C classes (subjects), the dimension of LGBP generally has to be reduced to be smaller than N-C by PCA before FDA, which will unavoidably results in the loss of much information. Aiming at this problem, we turn to ensemble of piecewise FDA classifiers by partitioning the entire feature vector.

3.2 Ensemble of Piecewise FDA Classifiers Based on Spatial Histograms of L(G)BP

We first rewrite the face representation in Equ.(3) as the following histogram feature vector:

$$X=(X_1,X_2,\cdots,X_m) , \qquad (7)$$

$$X_i=(h_i(b_1),h_i(b_2),\cdots,h_i(b_L)) , \qquad (8)$$

and m and L are respectively the number of spatial histograms and the number of bins for each histogram. We actually try to treat the frequency of each bin for each spatial histogram as one feature and concatenate them to form a feature vector with dimension $m*L$. As mentioned in above section, direct FDA on the high-dimensional feature vector X is not a good choice. In this paper, we partition all the m spatial histograms into k segments:

$$X = (G_1, G_2, \cdots, G_K)\,, \qquad (9)$$

where G_j is the j-th feature segment containing a specific number (n_j) of spatial histograms. Then, for each feature segment G_j, one FDA model is build to transform it to a low-dimensional representation F_j in the discriminant subspace:

$$F_j = (W_j^{lda})^T G_j\,. \qquad (10)$$

Thus, by building K FDA discriminant subspace, a face image is then represented as:

$$F = (F_1, F_2, \cdots, F_K) \qquad (11)$$

Finally, face classification can be achieved by an ensemble classifier combining all the component classifiers, each of which is designed based on one piece of FDA feature. In this paper, a simple sum rule is exploited to compute the similarity between two face images, that is,

$$S_{img}(I, I') = S_{sh-lbp}(X, X') = S_{fda}(F, F') = \sum_{i=1}^{K} S(F_i, F_i') \qquad (12)$$

and Nearest Neighbor (NN) is employed for classification. We have used the cosine similarity (rather than Euclidean distance) for $S(F_i, F_i')$ in Equ.12.

Note that, when training an FDA model, one may have to deal with the singularity of the within-class scatter matrix carefully, because generally the dimension of each G_j may still be much greater than the number of classes and training samples. In this paper, to avoid the singularity problem, PCA is conducted to reduce the dimensionality of the histogram vector to be less than N-C, where N is the number of training examples, and C is the number of classes. The PCA features are then transformed by FDA for final classification.

3.3 Implementation Details for Face Identify

This section describes the details of one of the possible implementations of the proposed method based on LBP and LGBP respectively. Note that the configuration described below is not necessary the optimal and we are on the way to finding the optimal one.

In this paper, a face image is geometrically and photometricly normalized as an image patch of 88x80 pixels. For the ensemble of piecewise FDA based on LBP, the parameters needed in section 3.2 are set to be: K=4, m=220, L=8, and n_j=55 for j=1,2,3,4. Note that the K=4 segments are obtained by partitioning the image patch into 4 image blocks of size 44x40, each of which is further partitioned into 55 sub-blocks (of size 4x8) to compute the spatial histograms. Thus, the dimension of each G_j will be 55*8=440 in our system. Fig.1 (a) has illustrated the procedure.

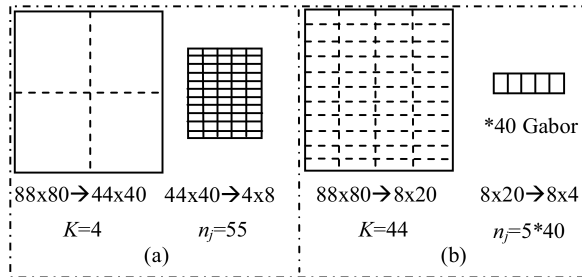

Figure 1. Region partitioning for piecewise FDA and spatial histograms. (a) LBP; (b) for LGBP.

For ensemble of piecewise FDA based on LGBP, as shown in Fig.1(b), face image patch of 88x80 pixels is partitioned into 11*4=44 blocks of size 8x20, that is, K=44 for Equ.(9). Then, for each block, it is further partitioned into 5 sub-blocks of size 8x4. Since 40 Gabor filters (5 scales and 8 orientations) are exploited, we can obtain 5*40=200 spatial histograms, that is n_j =200 for j=1,…,44. Similarly, L=8. Thus, the dimension of each G_j will be 200*8=1,600 in our system.

Note that the number of bins for each spatial histogram in this paper is fixed at 8, which simply quantifies the original 256 bins.

4. Experiments

The FERET database is used to validate the proposed method according to the standard FERET evaluation protocol, which has exactly defined the gallery and probe sets. In our experiments, we strictly evaluate all the methods based on the standard gallery (1196 images of 1196 subjects) and four probe sets, fafb(1195 images), fafc (194 images), dup.I (722 images), and dup.II (234 images). Please refer to [8] for details about the FERET evaluation protocol. Note that 1002 frontal images of 429 subjects are used as the training set for methods that need a training stage.

Table 1 shows the recognition rates of LBP-based methods in our experiments including LBP based on histogram intersection (HI_LBP), LBP based on Chi square (Chi_LBP), Holistic FDA on LBP (HFDA_LBP), and the proposed EPFDA on LBP (EPFDA_LBP), as well as their corresponding LGBP versions (HI_LGBP, Chi_LGBP, HFDA_LGBP, and EPFDA_LGBP).

Note that the LBP operator used in this paper is a simple version based on 3x3 neighborhoods. But it is still completely comparable, since the same LBP is used in all these methods. And, for better comparison we also list the results of LBP published in ECCV2004 [6] and best FERET97 results [8], which has been the best results on these datasets.

Table 1. Comparisons on the standard FERET probe sets

Method	Standard FERET probe sets			
	fafb	fafc	dup.I	dup.II
Chi_LBP	0.919	0.572	0.560	0.333
HI_LBP	0.936	0.577	0.580	0.389
HFDA_LBP	0.969	0.680	0.688	0.376
EPFDA_LBP	0.979	0.701	0.695	0.444
Chi_LGBP	0.970	0.969	0.755	0.697
HI_LGBP	0.970	0.954	0.749	0.688
HFDA_LGBP	0.977	0.825	0.802	0.739
EPFDA_LGBP	**0.996**	**0.990**	**0.920**	**0.889**
ECCV04 LBP[6]*	*0.970*	*0.790*	*0.660*	*0.640*
Best FERET97[8]*	*0.962*	*0.820*	*0.591*	*0.521*

Note*: these results are only suitable for limited reference, because their preprocessing methods might be different from ours.

From Table.1, we notice that FDA extension of L(G)BP can generally improve classification accuracy compared with HI_L(G)BP and Chi_L(G)BP, except HFDA_LGBP on fafc probe set. And the most exciting observation is that the proposed EPFDA_LGBP has achieved the most terrifically encouraging recognition rates on all the four probes. For fafb, the proposed EPFDA_LGBP method has been able to recognize almost all the faces (with only 5 errors). For fafc, we have also nearly reached the limitation (with only 2 errors). For aging probe sets, dup.I and dup.II, a great step has been made forward to about 90%, considering the best results ever reported before our work. From Table 1, the result of HFDA_LGBP on fafc is as low as 0.825, which seems one “exception”. We think this may be caused by the poor generalizability of holistic FDA, because no similar image as *fafc* probes are included in the FERET training set.

In addition, EPFDA_L(G)BP has effectively reduced the feature dimension. In our experiments, the results in Table 1 for EPFDA_LBP is achieved with each FDA’s dimension being about 100D, thus totally 4*100=400D for each face image. And for EPFDA_LGBP, the final total EPFDA dimension is about 11,000(44*250), which has been much lower than the original 70,400(44*1600).

5. Conclusions

We have proposed a novel Fisher discriminant extension of the LBP and LGBP, which is an ensemble classifier combining multiple piecewise FDA models. Experiments on standard FERET database show that the proposed method has made a great step forward towards recognizing all the ever challenging probes in fafc, dup.I, and dup.II datasets.

The results of the paper suggest us several observations: (1) LGBP is indeed impressively better than pure LBP; (2) The proposed FDA extension of L(G)BP histograms generally outperforms histogram matching method; (3) The proposed Ensemble of piecewise FDA classifiers of LGBP is proved to be a very promising method for face recognition.

Acknowledgement

This research is partially sponsored by Natural Science Foundation of China under contract No.60332010, "100 Talents Program" of CAS, the Program for New Century Excellent Talents in University (NCET-04-0320).

Reference

[1] W.Y. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, Face recognition: A literature survey. ACM Computing Survey, pp. 399-458, 2003.

[2] M. Turk and A. Pentland, Eigenfaces for Recognition, J. Cognitive Neuroscience,Vol.3, pp. 71-86, 1991.

[3] P. Belhumer, P. Hespanha, and D. Kriegman, Eigenfaecs vs. fisherfaces: Recognition using class specific linear projection. TPAMI, 19(7), pp.711-720, 1997.

[4] C. J. Liu, Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, IEEE Trans. on IP, 11(4), pp.467-476, 2002.

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää, Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns. TPAMI, 24(7), pp.971-987, 2002.

[6] T. Ahonen, A. Hadid, and M. Pietikäinen, Face recognition with Local Binary Patterns. In Proceeding of ECCV2004, pp.469-481, 2004.

[7] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition, ICCV2005, Vol. I, pp.786-791, 2005.

[8] P.J. Phillips, H.M. Syed, A. Rizvi, and P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms. TPAMI, 22(10), pp.1090-1104, 2000.