V-LGBP: Volume Based Local Gabor Binary Patterns for Face Representation and Recognition

Shufu Xie[1,2], Shiguang Shan[1], Xilin Chen[1], and Wen Gao[1,3]
[1]Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences (CAS), Beijing 100190, China
[2] Graduate University of CAS, Beijing, 100190, China
[3]Institute for Digital Media, Peking University, Beijing, 100190, China
{sfxie,sgshan,xlchen,wgao}@jdl.ac.cn

Abstract

In this paper, we propose volume based local Gabor binary patterns (V-LGBP) for face representation and recognition. In our method, the Gabor feature set of each gray image is regarded as a three dimensional "volume", where the first two dimensions are spatial domain and the third dimension is the Gabor filter index. Then, the neighborhood order relationship in the "volume" is encoded by Local Binary Patterns (LBP), which converts the Gabor transformed images into multiple index maps. Finally, the spatial histograms of all the V-LGBP index maps are concatenated together to represent the facial appearances. In addition, in order to reflect the uniform appearances of V-LGBP, its uniform patterns are redefined via statistical analysis. Extensive experiments on FERET dataset validate the effectiveness of our approach.

1. Introduction

Face recognition has attracted significant attention due to its potential value in security applications and research fields. Much progress has been made in the last two decades [10] . However, due to the fact that the facial appearances are easily affected by the variations of pose, expression, illumination and other factors, it is still an active and challenging research topic. Therefore, many researchers have devoted their efforts to improve the robustness of face recognition systems to these variations. One of the key issues for the successful development of face recognition systems is the development of robust representations.

Numerous face representations have been proposed in the literatures [1][2][3][4][5][7][9]. In the earlier years, many face representations based on global features have been studied widely, such as Eigenface[7], Fisherface[2]. Moreover, a great number of representations based on local features have received more attention due to their robustness to the variations of expression, illumination and other factors. Among these local representations, Gabor feature has been regarded as one of the best representations for face recognition, such as Dynamic Link Architecture (DLA) [3], Elastic Bunch Graph Match (EBGM) [4], Gabor Fisher Classifier (GFC)[5]. More recently, Local Binary Patterns (LBP) has received much attention in face analysis due to its high discriminative power and computational efficiency [1]. In [9], Local Gabor Binary Patterns (LGBP) is proposed by combining the Gabor wavelets and LBP, which has achieved impressive performances for face recognition. In [8], 3D-LBP is proposed for representing dynamic textures by extending LBP to spatio-temporal domain. All these works further show the effectiveness of the local representation methods.

In this paper, motivated by the success of Gabor feature and LBP, we propose volume based local Gabor binary patterns (V-LGBP) for face representation. In our method, the Gabor feature set of each gray image is regarded as a three dimensional "volume", where the first two dimensions are spatial domain and the third dimension is the Gabor filter index. Then, the neighborhood order relationship in the "volume" is encoded by LBP. So, multiple index maps are obtained, each of which is divided into many non-overlapping blocks and the histogram of each block is calculated. Finally, the histograms of all the blocks of all the index maps are concatenated together to represent the facial appearances. Extensive experiments on FERET [6] dataset validate the effectiveness of our approach.

The major contribution of this paper is the V-LGBP representation. Due to encoding the order relationship of both spatial domain and filter responses, it reflects more local variations than LGBP and is expected to achieve better performances. In addition, in order to reflect the uniform appearances of V-LGBP, its uniform patterns are redefined via statistical analysis, which results in low dimensional histogram features without affecting the accuracies.

2. Local Gabor Binary Patterns (LGBP)

LGBP can be regarded as a combination of Gabor feature and LBP, which applies LBP operator to the Gabor wavelet representation [9]. First, the Gabor wavelet representation is calculated by convolving the face image with the Gabor filters:

$$G_{\mu,\nu}(z) = f(z) * \psi_{\mu,\nu}(z) \tag{1}$$

where $f(\cdot)$ denotes the input image, $*$ denotes the convolution operator and $\psi_{\mu,\nu}(\cdot)$ denotes the widely used Gabor kernels with orientation μ and scale ν [3][4][5]. Then, each Gabor transformed image is converted into one LGBP index map by applying LBP operator to its magnitudes:

$$\begin{aligned} &LGBP_{\mu,\nu}(z_c) = [LGBP^0_{\mu,\nu}, LGBP^1_{\mu,\nu}, ..., LGBP^{P-1}_{\mu,\nu}] \\ &LGBP^i_{\mu,\nu} = s(G_{\mu,\nu}(z_i) - G_{\mu,\nu}(z_c)), i = 0,1,..., P-1 \end{aligned} \tag{2}$$

Here, $LGBP_{\mu,\nu}(z_c)$ denotes the LBP code of position z_c on Gabor magnitude map $G_{\mu,\nu}(\cdot)$, P denotes the number of neighbors and $s(\cdot)$ denotes the sign operator, which is defined as follows:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{3}$$

Finally, the binary string is converted into a decimal value which we call "index" to reflect the local pattern of position z_c on $G_{\mu,\nu}(\cdot)$.

For each Gabor kernel, one LGBP index map is obtained. In order to represent the facial appearances well, the LGBP index maps are divided into non-overlapping blocks and the histograms of all the blocks of all the scales and orientations are concatenated together to represent the feature distribution [9].

Compared with LBP, the main merit of LGBP lies in its capacity of modeling local features of varying orientations and scales, which is provided by the Gabor transform. It has achieved higher performance than the original LBP on the FERET database [9].

3. Face Recognition Based on V-LGBP

3.1. Motivation and Basic Ideas

In LGBP, the local patterns of the spatial neighborhood on each Gabor transformed image are encoded to represent the facial micro-structures. In fact, Gabor magnitudes also provide the monotonic measure of the image properties (e.g., "there is an edge present at position *x*")[3]. Thus, for different scale and orientation Gabor filter responses of the same position, they reflect the matching degree of the image patch with different kernels. The order relationship between those filter responses is equal to encode the local saliency (e.g., the response of kernel *a* on position *x* is larger than that of kernel *b*). In our opinion, this relationship is based on the saliency measure of the image patch, which is also useful for face recognition. Thus, by encoding the order relationship of both spatial domain and filter responses on Gabor magnitude maps, V-LGBP is proposed in this paper.

3.2. V-LGBP

According to the analysis in Section 3.1, the V-LGBP descriptor consists of two parts: local patterns of spatial neighborhood (i.e., LGBP) and filter responses. Fig.1 illustrates the computing process of the V-LGBP descriptor with some scale and orientation.

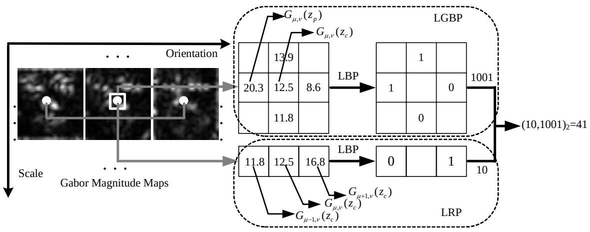

Fig.1 Illustration of the V-LGBP descriptor with some scale and orientation

Formally, V-LGBP is defined as follows:

$$\text{V-LGBP}_{\mu,\nu}(z_c) = [LRP_{\mu,\nu}(z_c), LGBP_{\mu,\nu}(z_c)] \tag{4}$$

Here, $LGBP_{\mu,\nu}(z_c)$ denotes the local pattern of position z_c encoded by LGBP, which is defined in Eq.(2); $LRP_{\mu,\nu}(z_c)$ denotes the local response pattern for position z_c along the dimensionality of Gabor filters, which is defined as follows:

$$\begin{aligned} &LRP_{\mu,\nu}(z_c) = [LRP^0_{\mu,\nu}, LRP^1_{\mu,\nu}, ..., LRP^{K-1}_{\mu,\nu}] \\ &LRP^i_{\mu,\nu} = s(G_{\mu_i,\nu}(z_c) - G_{\mu,\nu}(z_c)), \mu_i \neq \mu \end{aligned} \tag{5}$$

where K denotes the number of Gabor filters used to encode the local response pattern of position z_c and other notations have the same meanings as those defined in Eqs.(2) and (3). Then, the two binary strings are concatenated together and converted into a decimal value to reflect the local patterns of position z_c on $G_{\mu,\nu}(\cdot)$, which is shown in Fig.1.

Like LGBP, for each Gabor kernel, one V-LGBP index map is also obtained. Note that Gabor filters of 5 scales and 8 orientations are used in this paper. Then, each V-LGBP index map is divided into multiple non-overlapping blocks and the histogram of each block is calculated. Finally, the histograms of all the blocks of all the scales and orientations are concatenated together to represent the facial appearances. Fig.2 illustrates the V-LGBP representation for facial appearances. Note that histogram intersection [9] is also used to measure the similarity between the two histogram sequences and the Nearest Neighbor rule is adopted for final classification.

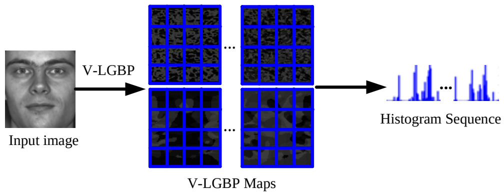

Fig.2 V-LGBP representation for facial appearances

One point to note is the selection of filters to encode the local response pattern. In our implementation, only those responses of the same scale but different orientation Gabor kernels are used. Fig.1 illustrates an example of local response pattern. The reason lies that the comparisons between them provide the effective measure of the local saliency. Although the same orientation but different scale responses also provide a measure of local saliency, their comparisons are less effective than those defined above due to their different modulus and kernel size. Experimental results in Section 4 have also verified this point.

Another concern might be the difference between V-LGBP and 3D-LBP in [8]. The main difference lies that V-LGBP is applied to the Gabor feature set of the same image, and 3D-LBP is applied to the real volume data defined in spatio-temporal domain. Specifically speaking, only different filter responses of the same position is encoded for V-LGBP because they reflect the saliency measure of the same image patch with different filters; But for 3D-LBP, those pixels in the local volume neighborhood are all compared with the central pixel.

3.3. Uniform Patterns of V-LGBP

Previous studies have shown that the uniform patterns are the fundamental properties of local image texture [1][8]. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is considered as circular. In Gabor domain, each position corresponds to one filter response, which reflects the matching degree of the image patch with the corresponding kernel. Further, the magnitudes are robust to the local variations while the gray values are easily affected by noise. So, directly using LBP uniform patterns defined on gray images for V-LGBP is unsuitable. Considering that uniform patterns correspond to the primary patterns, we redefine the uniform patterns of V-LGBP via statistical analysis.

Formally, V-LGBP^{us}, which denotes the uniform patterns of V-LGBP via statistical analysis, is redefined as follows:

$$\text{V-LGBP}^{us} = \{l \mid Rank[l_{percent}] < N\} \qquad (6)$$

where $l_{percent}$ denotes the percentage of pattern l in the whole V-LGBP index map set , $Rank[l_{percent}]$ denotes the index of $l_{percent}$ in term of descending order and N denotes the number of uniform patterns. In our experiments, we select the value of N to ensure that over 90% patterns are covered. By using only uniform patterns and labeling all the remaining patterns with a single label, the index map for V-LGBP^{us} is recalculated as follows:

$$\zeta = \begin{cases} Index(\zeta), & \text{if } \zeta \in \text{V-LGBP}^{us} \\ N, & \text{else} \end{cases} \qquad (7)$$

Here, ζ denotes the original pattern and $Index(\zeta)$ denotes the index of pattern ζ in the redefined uniform pattern set V-LGBP^{us}.

4. Experiments

The FERET [6] database is used to evaluate the proposed approach according to the standard FERET evaluation protocol. In this section, we evaluate our approach based on the standard gallery (1,196 images of 1,196 subjects) and four probe sets: Fb(1,195 images), Fc(194 images), DupI(722 images), and DupII(234 images).

In our experiments, the face images are normalized to the size of 80×88 pixels according to the provided eye positions, and no further preprocessing step has been made. The size of Gabor filter window is fixed to 32×32 pixels.

From the analysis in Section 3, we can see that there are three parameters for V-LGBP: the spatial neighbor size P , the number of Gabor filters K and the block size for histogram. In fact, the values of P and K determine the total number of patterns. If their values are too large, it will lead to the great number of patterns and the high dimensional histogram feature; if they are small, it only encodes few local variations, which has weak representation power. As to the block size, previous work can guide us to select the appropriate values [1][9]. Considering the tradeoff between the accuracies and computation cost, the values of P and K are set to 4 and 2 respectively, which leads to 64 kinds of patterns, and the block size is set to 10×11 pixels.

Table 1 tabulates the comparison with some previous approaches on the four standard probe sets. Note that “V-LGBP-s” denotes different scale but the same orientation filters used to encode the local response pattern, “V-LGBP-so” denotes different scale and different orientation filters used, and "V-LGBP-o" denotes different orientation but the same scale filters used; "V-LGBP-o^{u2}" denotes the original LBP uniform pattern used for "V-LGBP-o" and "V-LGBP-o^{us}" denotes the uniform patterns redefined in this paper.

Table 1 Comparison with some previous approaches on the standard FERET probe sets

Methods	Fb	Fc	DupI	DupII
GFC in [5]	0.96	0.84	0.70	0.57
Best Results of LBP in [1]	0.97	0.79	0.66	0.64
LGBP	0.97	0.97	0.76	0.71
V-LGBP-s	0.97	**0.99**	0.77	0.71
V-LGBP-so	**0.98**	**0.99**	0.79	0.75
V-LGBP-o	0.97	**0.99**	**0.80**	**0.77**
V-LGBP-o^{u2}	0.97	0.98	0.77	0.74
V-LGBP-o^{us}	0.97	0.98	**0.80**	**0.77**

One can see that the results of "V-LGBP-o" are basically higher than LGBP due to encoding the local patterns along the filter response dimensionality. In addition, the results of "V-LGBP-o^{us}" are comparable with those of "V-LGBP-o" while its histogram feature dimension is lower than that of "V-LGBP-o". The results of "V-LGBP-o^{u2}" on DupI and DupII sets are a little lower than those of "V-LGBP-o^{us}", which shows the effectiveness of our redefined uniform patterns. One can also see that, the results of "V-LGBP-s" and "V-LGBP-so" on DupI and DupII sets are basically lower than "V-LGBP-o", which shows that the order relationship between the filter responses of different orientations but the same scale are more useful for face recognition.

5. Conclusion

In this paper, we propose volume based local Gabor binary patterns (V-LGBP) for face representation and recognition. Because V-LGBP encodes the local patterns of both spatial domain and filter responses, it represents more local variations than LGBP. In addition, the uniform patterns of V-LGBP are redefined via statistical analysis, which reduces the dimension of histogram feature without affecting accuracies. Extensive experiments on FERET dataset validate the effectiveness of our approach.

Because different facial parts have different discriminative powers for face recognition, one future effort can be done to make use of the discriminative powers of different facial parts. Another effort can be made to extract the discriminative features from the histogram sequences by utilizing discriminant analysis (e.g., LDA).

Acknowledgement

This paper is partially supported by National Natural Science Foundation of China under contract No.60332010 and No.60772071; Hi-Tech Research and Development Program of China under contract No.2007AA01Z163; 100 Talents Program of CAS; and ISVISION Technology Co. Ltd.

References

[1] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE TPAMI 28(12), 2006.

[2] Peter N. Belhumeur, João P.Hespanha, and David J.Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE TPAMI 19(7), 1997.

[3] Martin Lades, Jan C. Vorbrüggen, Joachim Buhmann, Jörg Lange, Christoph v.d. Malsburg, Rolf P. Würtz, and Wolfgang Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture", IEEE Trans. On Computers, 42(3), 1993.

[4] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching", IEEE TPAMI 19(7), 1997.

[5] Chengjun Liu and Harry Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition," IEEE TIP 11(4), 2002.

[6] P.Jonathon Phillips, Hyeonjoon Moon, Syed A.Rizvi, and Patrick J.Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," IEEE TPAMI 22(10), 2000.

[7] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, 3(1), 1991.

[8] Guoying Zhao and Matti Pietikäinen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", IEEE TPAMI 29(6), 2007.

[9] Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen, and Hongming Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition," ICCV, 2005.

[10] W.Zhao, R.Chellappa, P.J.Phillips, and A.Rosenfeld, "Face recognition: A literature survey," ACM Computing Surveys, vol.35, 2003.