Face Recognition on Partial and Holistic LBP Features

Xiao-Rong Pu, Yi Zhou, and Rui-Yi Zhou

Abstract—An algorithm for face description and recognition based on multi-resolution with multi-scale local binary pattern (multi-LBP) features is proposed. The facial image pyramid is constructed and each facial image is divided into various regions from which partial and holistic local binary pattern (LBP) histograms are extracted. All LBP features of each image are concatenated to a single LBP eigenvector with different resolutions. The dimensionality of LBP features is then reduced by a local margin alignment (LMA) algorithm based on manifold, which can preserve the between-class variance. Support vector machine (SVM) is applied to classify facial images. Extensive experiments on ORL and CMU face databases clearly show the superiority of the proposed scheme over some existed algorithms, especially on the robustness of the method against different facial expressions and postures of the subjects.

Index Terms—Face recognition, local binary pattern operator, multi-resolution with multi-scale local binary pattern, local margin alignment dimensionality reduction.

1. Introduction

Face recognition is a significant researching field of computer vision and pattern recognition. It impacts important applications in many areas such as entrance control system and video surveillance. Due to its wide range of applications, automatic facial image recognition has attracted much attention in recent years, especially in dimension reducing and features extraction.

One of the crucial aspects in automatic facial recognition is facial representation, which derives a set of features from original facial images to effectively represent faces. The major approaches developed for face recognition are principal component analysis (PCA), linear discriminant analysis (LDA), and independent component analysis (ICA). PCA is commonly referred as the “eigenface” method, which provides an optimal linear transformation from the original image space to an orthogonal eigenspace with reduced dimensionality in the sense of least mean squared reconstruction error (LMSRE). LDA seeks to find a linear transformation by maximizing the between-class variance and minimizing the within-class variance. Both of them represent a face with holistic facial features. The local binary pattern (LBP) has been widely applied to texture classification and face recognition as an effective arithmetic operator in texture description, with rotational and gray scale invariance.

Psychology and brain science studies indicate that human visual system can automatically combine global and local features when recognizing faces. Inspired by these studies, we introduce a new approach, multi-resolution with multi-scale pyramid LBP, for face recognition, which extracts both holistic and partial texture information to represent face images. A nonlinear dimensionality reduction method, local margin alignment (LMA), is adopted to reduce the dimensionality of LBP features. Finally, support vector machine (SVM) is applied to classify facial images.

The paper is organized as follows. The classical LBP and the proposed multi-resolution with multi-scale LBP are described in Section 2. LMA dimensionality reduction method is presented in Section 3. Experimental results are presented in Section 4, and Section 5 concludes the paper.

2. Face Description with Local Binary Patterns

2.1 Local Binary Patterns

The original LBP operator introduced by Ojala et al. has been used extensively for texture discrimination, demonstrating excellent results and good robustness against rotation and global illumination changes. The operator labels the pixels of an image by thresholding a 3×3 neighborhood of each pixel with the center value and considering the results as a binary number. Then the histogram of the labels can be used as a texture descriptor. See Fig. 1 for an illustration of the fundamental LBP operator.

![Fundamental LBP operator](image_url)
The limitation of the fundamental LBP operator is its small $3 \times 3$ neighborhood which can not capture dominant features with large scale structures. Hence, the operator later is extended to use neighborhood of different sizes\(^6\). As shown in Fig. 2, LBP\((P, R)\) means comparing a neighborhood of $P$ on a circle of radius of $R$ to get LBP features. For instance, LBP\((4, 1)\) means comparing a neighborhood of 4 on the circle of radius of 1 to get LBP features.

A LBP operator LBP\((P, R)\) produces $2^P$ distinct values, corresponding to the $2^P$ different binary patterns for the signs of the differences in a neighborhood. The local texture can thus be approximately described with a $2^P$-bin discrete distribution of LBP codes. This yields to the notation of $\text{LBP}_{P,R}^{2^P}$, called as uniform patterns. A uniform LBP pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. For example, $011000000$ and $11100001$ are uniform patterns.

After labeling an image with the LBP operator, the histogram of the labeled image $p(x, y)$ can be defined as

$$H_u = \sum_{x,y} U(p(x, y) = u), \; u = 0, 1, \cdots, n - 1$$

where $n$ is the number of different labels produced by the LBP operator and

$$U(A) = \begin{cases} 1, & A = \text{true} \\ 0, & A = \text{false} \end{cases}$$

A LBP histogram can effectively describe the distribution of the local micro-patterns over a whole face image without any indication about their locations. For efficient face representation, one should also retain spatial information. Thus, a face image can be equally divided into small regions (as shown in Fig. 3). And then, the LBP features extracted from each sub-region are concatenated into a single histogram as\(^{13}\)

$$H_{u,v} = \sum_{x,y} U(p(x, y) = u)U((x, y) \in R_v)$$

where $u = 0, 1, \cdots, n - 1$ and $v = 0, 1, \cdots, m - 1$.

In such a representation, the texture of facial regions is encoded by the LBP (called holistic LBP features) while the shape of the face is recovered by the concatenation of different local histograms (called partial LBP features).

### 2.2 Multi-Resolution with Multi-Scale LBP

Most LBP operators describe the texture distribution of each pixel with its neighborhood only within the original face image. However, the differences between two faces can be demonstrated not only by the texture distribution of each pixel with its neighborhood, but also by the relative connection with further pixels. We propose multi-resolution with multi-scale LBP (Multi-LBP) approach to describe the relative locations of each sub-regions.

A data structure called facial image pyramid is constructed\(^{13}\). It consists of a set of low-pass copies of an original facial image, as shown in Fig. 4.

These reduced resolution levels of the facial image pyramid are obtained through a highly efficient iterative algorithm. The bottom level is the original facial image, which is then transformed to the next level by low-pass filtering and sub-sampling on a factor $s$. This level is then transformed to another level in the same way. A two-tap filter is used to generate the facial image pyramid in Fig. 4. While constructing each pyramid level, its LBP features are extracted. Then all LBP features of each level are concatenated to a single LBP eigenvector with different resolutions.

Multi-resolution with Multi-LBP approach can obtain the relationship among pixels of a facial image in a larger scale, which can contain more face features with the cost of increasing data redundancy. Based on the notion of manifold learning, a LMA algorithm is used to extract the intrinsic characteristics of different types of high dimensional data by performing nonlinear dimensionality reduction\(^9\).
Dimensionality Reduction on Local Margin Alignment

Automatic face recognition is still a great challenge due to the curse of dimensionality and associated large computational demands. Some researches show the variations of face images (such as illumination, identity, pose and facial deformations) can be represented as low dimensional manifolds embedded in the high dimensional image space [14][15]. The LMA algorithm is applied by using nonlinear dimensionality reduction to embed facial deformations in a low-dimensional space, which establishes a regression between two or more data sets by partially aligning their underlying manifolds. It relies upon optimization over a facial representation, where edges of the facial images are computed to preserve local structure in it.

Given c classes of data, there are \( n_i \) \((i = 1, 2, \ldots, c)\) items in class \( i \), where \( n_1 + n_2 + \cdots + n_c = n \). Let \( X = \{x_j \in \mathbb{R}^d, \ j = 1, 2, \ldots, n\} \) be the sample with noise in class \( i \), which is locally linear separable. LMA is applied to reduce the dimensionality of the input data \( X \in \mathbb{R}^{d \times n} \), to get the low-dimension eigenvector \( Z = [z_1, z_2, \cdots, z_n] \in \mathbb{R}^{m \times n} \) \((m \ll d)\) for discriminant analysis. See Fig. 5 for the geometry structure of two classes.

Given a smooth separable region \( F_{nk} \) between two classes in a high dimensional image space, the geometry structure of two classes is shown in Fig. 5 (b). For the data in class \( i \), we aim to get between-class variance among all the classes and find local direction vector to classify. Since most data in real world can be projected to a low-dimensional space \( F_{nk} \) (as shown in Fig. 5 (c)), the location of the most recently input data can be predicted. It is in favor of increasing the generalization ability. LMA can preserve between-class variance and inner-class structure. The between-class variance can be defined as

\[
\max_{X} \sum_{i=1}^{c} \sum_{j \in N_i(t)} \alpha_j^{(i)} (x_j - x_i^{(t)})^T (x_j - x_i^{(t)}) = 0
\]

\[
\max_{X} \sum_{i=1}^{c} \sum_{j \in N_i(t)} \alpha_j^{(i)} (P_b)\Sigma (x_j - x_i^{(t)}) = 0
\]

\[
\max_{X} \text{Trace} \left[ (P_b)^T X (I - A_{b_i})^T W_{b_i} (I - A_{b_i}) X^T P_b \right] = 0
\]

where \( P_b \in \mathbb{R}^{m \times m} \), \( N_i(t) \) is the collection of \( k_b \) ordered between-class distance of \( x_i \), \( \alpha_j^{(i)} (j \in N_i(t)) \) is the weight of the nearest between-class neighbors, \( I_b \in \mathbb{R}^{m \times m} \) is an identity matrix, and \( A_{b_i} \in \mathbb{R}^{m \times m} \) is a coefficient matrix

\[
A_{b_i}(t,j) = \frac{\alpha_j^{(i)}}{\sum_{j=1}^{n_b} \alpha_j^{(i)}}, (t,j = 1, 2, \cdots, n)
\]

\( W_{b_i} \in \mathbb{R}^{m \times m} \) is a diagonal matrix, and its diagonal element

\[
W_{b_i} = \frac{\sum_{j=1}^{n_b} \alpha_j^{(i)}}{\sum_{j=1}^{n_b} \alpha_j^{(i)}}
\]

is the normalized weight factor of \( x_i \) \((i \in N_i(t))\) in class \( i \). The larger \( W_{b_i} \) is, the more favorable \( x_i \) is in projecting to the optimal global direction.

For a concatenated LBP eigenvector, a linear LMA operator is

\[
Z = (P_{LBP} P_{LMA})^T X
\]

where \( P_{LBP} \) is the LBP eigenvector with different resolutions extracted using Multi-LBP, and \( P_{LMA} \) is the linear projection matrix of LMA. Nonlinear LMA projects onto linear space by using kernel function.

### 4. Experiment Results and Analysis

To evaluate the proposed multi-resolution with the Multi-LBP algorithm, we systematically compare it with the traditional LBP algorithm on ORL and CMU face databases. ORL database contains 400 frontal images with different facial expressions, illumination conditions, hairstyles with or without glasses for 40 subjects, 10 images for each subject. Each sample is a 92×112 grey image, with tolerance for some tilting and rotation of up to 20. CMU database includes 123 subjects, with several facial expressions for each subject. Some of the samples are shown in Fig. 6.

A set of facial image pyramid are designed to support efficient scaled convolution through reduced facial image representation. It consists of a sequence of copies of an original facial image in which both sample density and resolution are decreased in regular steps. All samples of facial image pyramid are divided into blocks, the LBP features of each image are extracted, and then all LBP features of each image are concatenated to a single LBP eigenvector. Finally, LMA is applied to reduce the dimensionality of the concatenated LBP feature vector. Fig. 7 and Fig. 8 illustrate the 2D and 3D projections of facial images whose dimensionality is reduced by LMA.
Fig. 6. Some samples in ORL and CMU face databases: (a) some samples in ORL database and (b) some samples in CMU database.

Fig. 7. 2D projection on LBP-LMA features of 4 subjects in ORL database.

For both ORL and CMU databases, half samples are randomly selected as the training set and the remaining samples as the testing set. The results of 4 algorithms (i.e. LBP, EHMM-LBP, Multi-LBP and Multi-LBP-LMA) on ORL and CMU databases are shown in Table 1, which indicates that the Multi-LBP-LMA algorithm is significantly better than the other 3 algorithms. Here, EHMM-LBP\(^{[16]}\) method refers to embedded hidden Markov model (HMM) with LBP, which divides the face into five subregions HMMs: forehead, eyes, nose, mouth, and chin. LBP is applied to extract observation vectors.

Table 1: Face recognition rates on ORL and CMU databases

<table>
<thead>
<tr>
<th>Rec. methods</th>
<th>Rate on ORL</th>
<th>Rate on CMU</th>
<th>Ave. rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>92.3</td>
<td>91.5</td>
<td>91.9</td>
</tr>
<tr>
<td>EHMM-LBP</td>
<td>92.5</td>
<td>92.3</td>
<td>92.4</td>
</tr>
<tr>
<td>Multi-LBP</td>
<td>94.5</td>
<td>94.7</td>
<td>94.6</td>
</tr>
<tr>
<td>Multi-LBP-LMA</td>
<td>96.2</td>
<td>97.4</td>
<td>96.8</td>
</tr>
</tbody>
</table>

Table 2: Recognition rates on different sampling rate

<table>
<thead>
<tr>
<th>Sample rate</th>
<th>Rate on ORL</th>
<th>Rate on CMU</th>
<th>Ave. rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8</td>
<td>95.3</td>
<td>96.1</td>
<td>95.7</td>
</tr>
<tr>
<td>0.5</td>
<td>96.2</td>
<td>97.4</td>
<td>96.8</td>
</tr>
<tr>
<td>0.3</td>
<td>94.9</td>
<td>95.1</td>
<td>95.0</td>
</tr>
</tbody>
</table>

To verify the relationship of various sampling rates with recognition rate, tests have been conducted for three different sampling rates of 0.8, 0.5, and 0.3. Table 2 shows the average recognition rates. It is obvious that the sampling rate should neither be quite high nor extremely low.

Fig. 8. 3D projection on LBP-LMA features of 4 subjects in CMU database.

Fig. 9 shows the recognition rate corresponding to different sampling levels ranging from 1 to 7. It should be noted that a suitable sampling rate and sampling levels are critical to the recognition rate.

Fig. 9. Relationship between sampling levels and recognition rate.
5. Discussion and Conclusions

Face images can be seen as a composition of micro-patterns which can be well described by LBP. We exploit this observation and propose a simple and efficient representation for face recognition. In our approach, a facial image pyramid is constructed. Each face image is then divided into several blocks (facial regions) from which we extract local binary patterns and construct a global feature histogram that represents both the statistics of the facial micro-patterns and their spatial locations. Then LMA is applied to reduce the dimensionality of the concatenated LBP feature vector. Finally, face recognition is performed by using the SVM classifier.

Extensive experiments on ORL and CMU face databases clearly show the superiority of the proposed scheme over traditional LBP algorithms, especially on the robustness of the method against different facial expressions and postures of the subjects.

References


Xiao-Rong Pu was born in Sichuan Province, China, in 1969. She received the M.S. degree in artificial intelligence from Southwest Normal University, Chongqing, China, in 2002 and the Ph.D. degree in computational intelligence from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2007. She was a visiting scholar with the University of Manchester Institute of Science and Technology, Manchester, U.K., in 2004. Currently, she is an associate professor with the School of Computer Science and Engineering, UESTC. Her current research interests include neural networks, biometrics, and affective computing.

Yi Zhou was born in Sichuan Province, China, in 1985. He received the B.S. and M.S. degrees both from School of Computer Science and Engineering in UESTC, in 2008 and 2011, respectively. His research interests include biometrics and affective computing.

Rui-Yi Zhou was born in Sichuan Province, China, in 1990. He is now a senior undergraduate in UESTC. His research interests include biometrics and affective computing.