IMPROVED COMBINATION OF LBP AND SPARSE REPRESENTATION BASED CLASSIFICATION (SRC) FOR FACE RECOGNITION

Rui Min and Jean-Luc Dugelay

Department of Multimedia Communications, EURECOM, Sophia Antipolis, France
Email: {min, jld}@eurecom.fr

ABSTRACT

Recently, local binary patterns (LBP) based descriptors and sparse representation based classification (SRC) become both eminent techniques in face recognition. Preliminary techniques of combining LBP and SRC have been proposed in the literature. However, the state-of-art method suffers from the "curse of dimensionality" for real world scenarios. In this paper, a novel face recognition algorithm of combining LBP with SRC is proposed; in which the dimensionality problem is resolved by divide-and-conquer and the discriminative power is strengthen via its pyramid architecture. The proposed face recognition method is evaluated on AR Face Database and yields very impressive results.

Index Terms— Face Recognition, Local Binary Patterns, Sparse Representation based Classification, Divide and Conquer, Pyramid Architecture

1. INTRODUCTION

Automatic face recognition (AFR) is one of the classical research topics in computer vision. Thanks to the rapid advancements in signal processing and machine learning, the up-to-date AFR methods can achieve relatively high accuracy in controlled environments (e.g. in biometric systems), and ameliorate the performance degradation caused by different variations (illumination variations, poses changing, occlusions etc.) in uncontrolled environments (e.g. in video surveillance systems). Among numerous existing algorithms, local binary patterns (LBP) and sparse representation based classification (SRC) are two leading techniques in face recognition research. Due to their prominent capabilities in feature extraction and pattern classification respectively, research works based on such techniques for face recognition are very active in recent years.

LBP is known to be a powerful feature extractor for face representation [1]. The success of LBP in face description is due to the discriminative power and computational simplicity of the operator, and its robustness to monotonic gray scale changes caused by, for example, illumination variations. The use of histograms to collect features also makes the LBP approach robust to face misalignment and pose variations. So far many extensions of LBP for face recognition have been proposed, for example (not an exhaustive list): boosting LBP [2], LGBP [3], Multi-Scale LBP [4], MB-LBP [5], CLBP [6] etc. All those extensions target on improving the robustness and accuracy in non-optimal face recognition scenarios.

Recently, Wright et al. [7] introduced a framework for robust face recognition via sparse representation. Here face recognition is casted as penalizing the $\ell 1$-norm of the coefficients in the linear combination of an overcomplete face dictionary. Sparse representation based classification (SRC) has been demonstrated to be superior to nearest neighbor (NN) and nearest subspace (NS) based classifiers in various subspaces (e.g. PCA or LDA). When applied to face recognition, it can also be efficiently customized to handle errors due to occlusion and corruption. Following Wright et al.'s work, in the past year, several extensions of SRC based face recognition were proposed. In [8], Zhou et al. applied a Markov Random Field model to SRC based face recognition for improving performances under severe contiguous occlusion. Yang and Zhang [9] used image Gabor-features for SRC in order to reduce the cost in coding occluded faces meanwhile improving accuracy. In [10], Yang et al. reviewed five representative l1-minimization methods in the context of SRC based face recognition.

Very lately, a preliminary tentative of combining LBP based features with SRC for face recognition is presented by Chan and Kittler [11]. The authors illustrated that histogram descriptors, such as LBP, local phase quantization (LPQ) and Gabor phase pattern (GPP), are more robust to misalignment and illuminations than the holistic features used in SRC. Their approach returns impressive results on the combination of Yale Face Database B and extended Yale Face Database B (38 people under 64 different illumination conditions). Nevertheless, such an approach is infeasible for most real world datasets where only few samples are available for each subject. The reason is due to the "Curse of Dimensionality". In SRC, the ability of discrimination relies on computing the correct sparse solution of an underdetermined system of linear equations. But the solution is non-sparse when the number of non-zero entries in the coefficients vector beyond the equivalence breakdown point (EBP) [12]. In other words, the correct sparse solution can only be recovered when the number of training samples is sufficiently larger than the number of features. Unfortunately, the dimension of LBP histogram for

This work is partially funded by the French national project FR OSEO BIORAFALE.

face representation is generally huge (16384 in an ordinary 8×8 block division). Thus the direct combination of LBP and SRC is non-realistic. (In [11], the authors tested only for 2×2 sub-blocks but with 722 training images.)

One possible solution is to reduce the dimension of extracted features before performing SRC by applying dimension reduction tools (e.g. PCA or LDA as suggested by Yang et al. [13]). It ensures the desired solution is sparse and thus the "curse of dimensionality" is no longer a problem. Such an approach can slightly improve the performance comparing to the baseline algorithm [1].

In this paper, we propose a more powerful approach to combine LBP with SRC for face recognition. In our approach, the dimensionality problem is resolved by the divide-and-conquer scheme [14], where the SRC is performed on LBP histogram extracted from a single sub-block. The obtained sparse coefficients vectors (SCV) from all sub-blocks are then fused together to yield the final output. Besides, our approach has a pyramid architecture which is able to incorporate information from different levels of descriptions. The architecture thus can improve the robustness to various localized variations (inspired by Modular Eigenface [15]). In the experiments, we built a more realistic dataset (comparing to [11]) with different variations (including facial expressions, illumination conditions and time elapses) with fewer training samples using the AR face database [16]. On this dataset, the proposed algorithm is compared with the baseline method and the methods using dimension reduction tools. Our approach yields the best recognition rate (up to 96%).

The rest of this paper is structured as follows. First, tools used in the proposed approach are reviewed in Section 2. Then, the proposed algorithm is detailed in Section 3. Section 4 presents the experimental results. Finally, we draw a conclusion and discuss future directions in Section 5.

2. BACKGROUND&RELATED ALGORITHMS

In this section, necessary knowledge of the tools used in our algorithm is given. First, the LBP based face representation presented in [1] is reviewed in 2.1. Then the algorithm of face recognition using sparse representation [7] is reviewed in 2.2.

2.1. Local Binary Patterns based face representation

The original LBP operator forms labels for the image pixels by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number. The histogram of these $2^8 = 256$ different labels can then be used as a texture descriptor. Each bin (LBP code) can be regarded as a micro-texton. Local primitives which are codified by these bins include different types of curved edges, spots, flat areas etc.

The calculation of the LBP codes can be easily done in a single scan through the image. The value of the LBP code of a pixel (x_c, y_c) is given by:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \tag{1}$$

where g_c corresponds to the gray value of the center pixel (x_c, y_c), g_p refers to gray values of P equally spaced pixels on a circle of radius R, and s defines a thresholding function as follows:

$$s(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0, & otherwise \end{cases} \tag{2}$$

The occurrences of the LBP codes in the image are collected into a histogram. The classification is then performed by computing histogram similarities.

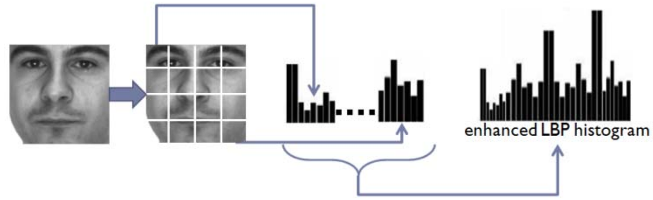

Fig. 1. The procedure of block based face representation.

As suggested in [1], in order to retain the spatial information, a facial image is divided into K non-overlapping regions from which LBP histograms are extracted and concatenated into an enhanced feature histogram. Figure 1 visualizes the procedure of how to compute the block based face representation. When a probe face is inputted into the face recognition system, such an enhanced LBP histogram is computed, and then the histogram similarity between the probe face and all template faces are measured through Chi-square distance (χ^2).

2.2. Sparse representation based face classification

Supposing a training set A consists of the facial images from k classes, where $A = \{A_1, A_2, \dots, A_k\}$. Ideally, giving sufficient training samples of class i , where $A_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,n_i}\} \in R^{m \times n_i}$, a test facial image $y \in R^m$ belongs to the same class could be well approximated by a linear combination of the training samples from A_i, which can be written as:

$$y = \sum_{j=1}^{n_i} a_{i,j}\, v_{i,j} \tag{3}$$

Since A is the dictionary which includes all the training samples, where $A = \{v_{1,1}, v_{1,2}, \dots, v_{k,n_i}\}$. Then equation (3) can be rewritten in the form as below:

$$y = Ax_0 \in R^m \tag{4}$$

where $x_0 = \{0, \dots, 0, a_{i,1}, a_{i,2}, \dots, a_{i,n_i}, 0 \dots, 0\}^T$ is the coefficient vector in which most coefficients are zero except the ones associated with class i.

Due to the fact that a valid test sample y can be sufficiently represented only using the training samples from the same class, and this representation is the sparsest among all others, to find the identity of y then equals to find the sparsest solution of (4). This is the same as solving the following optimization problem ($\ell 0$-minimization):

$$\hat{x}_0 = \arg\min \|x\|_0 \qquad subject\ to\ \ Ax = y \tag{5}$$

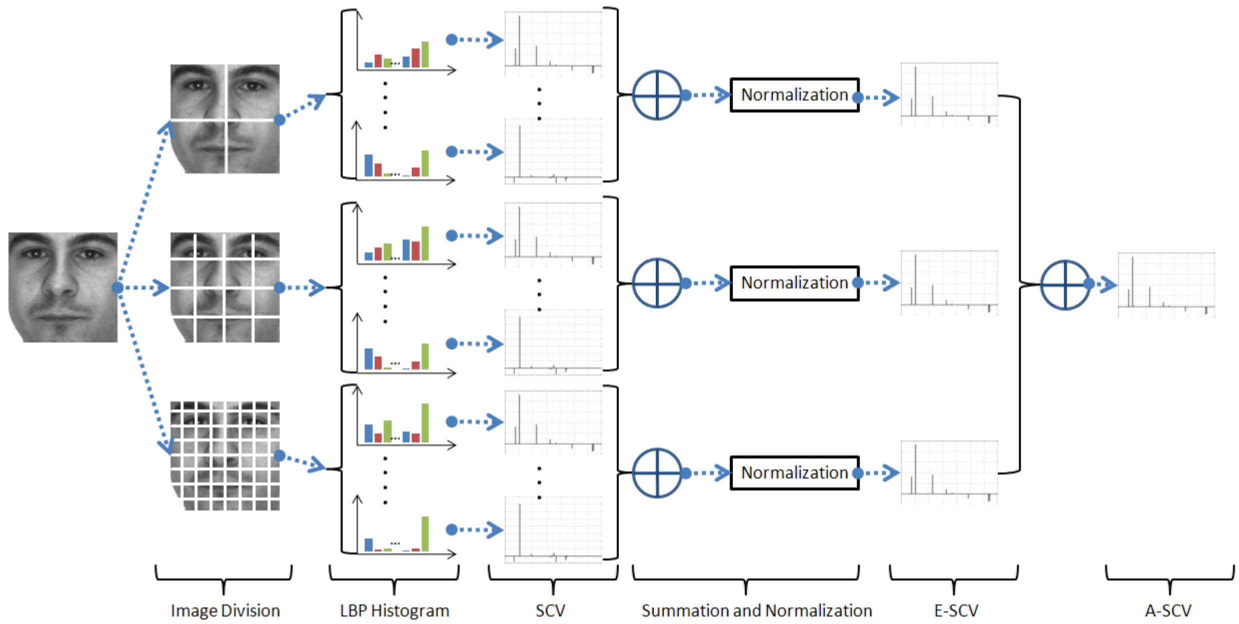

Fig. 2. Architecture of the proposed approach

However, solving the $\ell0$-minimization of an underdetermined system of linear equations is NP-hard. In the case for large number of training samples, it equals to find the minimal $\ell1$-norm solution [12]. Therefore, the SRC procedure presented in [7] is shown as below.

Algorithm 1. The SRC Algorithm

1. Normalize the columns of A to have unit $\ell2$-norm
2. Solve the $\ell1$-minimization problem:

$$\hat{\mathbf{x}}_1 = \arg\min\|\mathbf{x}\|_1 \ \ \textbf{subject to} \ \ \|\mathbf{Ax}-\mathbf{y}\|_2 \leq \boldsymbol{\varepsilon} \tag{6}$$

3. Compute the residuals by:

$$\mathbf{r_i}(\mathbf{y}) = \|\mathbf{y} - \mathbf{A}\boldsymbol{\delta}_\mathbf{i}(\hat{\mathbf{x}}_1)\|_2 \tag{7}$$

for $\mathbf{i} = \mathbf{1}, \ldots, \mathbf{k}$, where $\boldsymbol{\delta}_\mathbf{i}$ is the characteristic function which selects the coefficients associated with the $\mathbf{i}$-th class.

4. output the identity by :

$$\mathbf{identity}(\mathbf{y}) = \arg\min_\mathbf{i}\mathbf{r_i}(\mathbf{y}) \tag{8}$$

3. THE PROPOSED APPROACH

Based upon LBP based face representation and sparse representation based classification we reviewed before, in this section, we propose a new face recognition algorithm that combines LBP features with SRC. In the first part, we outline the architecture of the proposed algorithm. Then we state the motivation of using the divide-and-conquer scheme which encodes LBP features into sparse coefficients vector (SCV) via efficient $\ell1$-minimization. Finally, the classification methodology based on the computed representation (augmented sparse coefficients vector) is discussed.

3.1. Architecture of the proposed algorithm

Figure 2 illustrates the proposed architecture of combining LBP features with SRC for face recognition. A facial image is first divided into several division levels (2×2, 4×4 and 8×8 in our approach). In each division level, the LBP histograms are summarized over all sub-blocks. Instead of constructing an enhanced LBP histogram for each facial image, we treat the LBP histogram of each sub-block individually. The original SRC takes the entire feature space extracted from the whole image as a dictionary component. In our approach the dictionary component is the feature vector extracted from one sub-block. In this way, the SCVs are computed for all sub-blocks of the input image via $\ell1$-minimization (equation (6)). Then the SCVs within the same division level are combined to build an enhanced SCV (E-SCV) using elementary-wise summation. In order to balance the classification impacts from all division levels, the E-SCVs are normalized to have the unit l2-norm. Finally, the normalized E-SCVs from all division levels are summed to construct an augmented SCV (A-SCV) in order to make the classification decision.

The proposed algorithm is highly hierarchical. In the feature extraction phase: the labels for the LBP histogram include information about the patterns on a pixel-level; then the labels are summed over a small region to produce information on a regional level. In the classification phase: the SCV is recovered for each sub-block on a block-level; the E-SCV is computed for each type of division on a division-level; and the A-SCV is produced at the global-level. Such a pyramidal multi-level architecture ensures the accuracy and robustness of the proposed algorithm.

3.2. Motivating divide-and-conquer for combining LBP features with SRC

3.2.1. Uniform LBP

Since the main difficulty of combining LBP features with SRC is due to the dimensionality problem, it encourages us to exploit more parsimonious features instead of the original LBP histogram. According to [17], most of the texture information is contained in a small subset of LBP patterns. Those patterns are called uniform patterns, which contain at most two bitwise transitions (0 to 1 or 1 to 0). By using uniform LBP histogram, the feature dimension thus decreases from 256 to 59. In our approach, the operator we used is $\mathrm{LBP}_{8,2}^{\mathrm{u2}}$ (which means uniform patterns, 8 equally spaced pixels on a circle of radius 2). Since the main purpose of our paper is to introduce a more appropriate way to integrate high dimensional features (like LBP) into SRC, we thus focus on the basic LBP descriptor. Nevertheless, other features such as the extensions of LBP (e.g. LGBP, Multi-Scale LBP) and other histogram descriptors (LPQ, GPP, HOG etc.) could also be adopted into the proposed approach and might achieve higher recognition rates.

3.2.2. Dimensionality Reduction

A possible option to further reduce the feature dimension is to apply the tools like PCA or LDA to project the features from its original space to a reduced space [13]. Denoting the projection function as $R^{d\times m}$ with $d \ll m$, applying R to both sides of equation (4) yields:

$$\tilde{y} \doteq Ry = RAx_0 \in R^d \tag{9}$$

The projection guarantees that the system of equations (9) is under-determined. In addition, by selecting properly the reduced dimension, it can also ensure the system of equations has the unique sparsest solution. According to the experimental results we obtained, methods applying LDA can slightly improve the performance comparing to the baseline algorithm. However, it is still less accurate than the proposed method.

3.2.3. Divide-and-Conquer

The definition of divide-and-conquer motivates us to handle the "curse of dimensionality" using such a strategy:

> "To solve a large instance of a problem, break it into smaller instances of the same problem, and use the solutions of these to the original problem."

As supported by [14], such a strategy is very useful in high-level image processing, and it can be effectively performed using parallel computing. By applying divide-and-conquer to an input image, the computation of $\ell 1$-minimization becomes feasible for LBP based features in the constrained scenarios (e.g. limited number of training samples available for each class).

In addition, the divide-and-conquer algorithm enables us to incorporate information from multi-levels of descriptions by adding a pyramid structure. The composition of diverse representations from different division levels significantly improves the recognition performance.

3.3. Classification based on the augmented sparse coefficients vector

Once the augmented sparse coefficients vector (A-SCV) is computed, classification is conducted based on it. In Wright et al.'s method [7], the probe face is approximated using only the coefficients associated with the i-th class. Then the classification is based on minimizing the residual between the probe image and the approximations. According to equation (8), the identity is assigned by the class which has the least residual.

However, directly adopting Wright et al.'s method to A-SCV is inappropriate. Noticing that A-SCV is derived from SCVs which are computed from sub-blocks of the input image, it cannot be used to approximate the probe face by multiplication with the basis vector which is comprised of the whole facial images. Since a valid probe face can be sufficiently represented only using the training samples from the same class, its SCV is naturally discriminative. A-SCV is actually an enhanced version of SCV. Supposing the values in A-SCV of a valid test sample y can be written as: $p_y(i,j)$,where $i \in [1,k]$, k is the number of classes and $j \in [1, J(i)]$, $J(i)$ is the number of faces in the i-th class. We define a function $\phi_i(y)$ as below:

$$\phi_i(y) = \sum_{j=1}^{J(i)} p_y(i,j) \,/\, J(i) \tag{10}$$

which returns the normalized summation of sparse coefficients within the same class. Then the identity is assigned by:

$$identity(y) = \arg max_i \phi_i(y) \tag{11}$$

The identity is assigned to the class which maximizes the normalized summation of the associated sparse coefficients.

4. EXPERIMENTAL ANALYSIS

To assess the performance of our proposed approach, we performed a set of experiments on AR face database [16]. We will here illustrate the result of the proposed approach (D&C+LBP+SRC with the pyramid architecture), as well as four other methods we discussed in previous sections (LBP+NN [1], LBP+PCA+SRC, LBP+LDA+SRC and D&C+LBP+SRC). Results show that our proposed approach outperforms all the others. In addition, following the same configuration on the same database, we obtained better results than the best one reported in [7].

4.1. Experimental Data and Setup

The AR face database is a standard testing dataset in face recognition research. It contains more than 4000 face images of 126 subjects (70 men and 56 women) with different facial expressions, illumination conditions, and occlusions. For each subject, 26 pictures were taken in two separate sessions (two weeks interval between the two sessions). As Wright et al. did in [7] we configure the same dataset for testing. In the experiment, 100 subjects (half of male and half of female) are selected. For each subject, we chose 14 images with different illumination conditions and facial expressions (see in Figure 3): 7 images from session 1 for training and 7 images from session 2 for testing. The original image resolution is 768x576 pixels. Using the eye coordinates, we cropped, normalized and down-sampled the original images into 128x128 pixels. Some results of cropped and normalized faces are shown in Figure 3.

Fig. 3. Examples of extracted faces (from left to right: Neutral expression, Smile, Anger, Scream, left light on, right light on, all side lights on.)

Our dataset is quite challenging since it involves different variations including facial expressions, illumination conditions and the time elapse for 2 weeks. In addition, comparing to the dataset configured in [11], our dataset is more realistic. Instead of using 38 subjects with 19 training samples per subject, in our dataset, the number of subjects is 100 and there are only 7 training samples for each subject. The increased number of identities as well as the reduced number of training samples increases the difficulty for face recognition. Directly applying SRC to LBP histogram extracted from the whole facial image is infeasible in this scenario.

Other settings of the experiments are listed here: the LBP operator used is $\mathrm{LBP}_{8,2}^{u2}$; the ε in equation (6) is 0.05 (as suggested in [13]); and the $\ell 1$-minimization algorithm is implemented by $\ell 1$-magic [18].

4.2. Experimental Results and Analysis

In this part, several LBP based algorithms are tested in order to demonstrate the superiority of the proposed approach. Conventionally, selecting the image division strategy for LBP based approaches is heuristical or empirical. Results in previous publications are often obtained by division strategies which maximize the recognition rates. Here we tested various algorithms with different division strategies. In our test, an image is divided into 2 × 2, 4 × 4 and 8 × 8 sub-blocks respectively.

Firstly, four previously discussed approaches (LBP+NN, LBP+PCA+SRC, LBP+LDA+SRC and D&C+LBP+SRC) are examined. LBP+NN indicates the original LBP based face recognition by Ahonen et al. [1]. It uses a nearest neighbor classifier and its similarity measure is based on the Chi-square distance (χ^2). LBP+PCA+SRC and LBP+LDA+SRC refer to the methods which apply the dimension reduction tools (Principal Component Analysis and Linear Discriminate Analysis, respectively) to reduce the dimension of extracted LBP features before Sparse Representation based Classification. D&C+LBP+SRC combines LBP features with SRC by using the proposed divide-and-conquer strategy but without the pyramid architecture. For fair comparison, the feature dimension is reduced to 59 in both PCA and LDA based approach, which is same as the length of LBP histogram extracted from one sub-block.

Figure 4 shows the recognition rates of above four methods on AR face dataset. First of all, it shows that by deploying more precise sub-blocks division the recognition rate increases correspondingly. The reason is because face representations obtained from more precise division strategies contain more spatial information. We consider LBP+NN as the baseline algorithm. In the figure, LBP+PCA+SRC outperforms LBP+NN for 2 × 2 sub-blocks; but it returns worse results than LBP+NN on 4 × 4 and 8 × 8 division strategies. LBP+LDA+PCA yields better results than both LBP+NN and LBP+PCA+SRC on 4 × 4 and 8 × 8 division strategies. It corresponds to the fact that by explicitly including the label information of the data constructs a more informative projection. Among all the results, D&C+LBP+SRC on 8 × 8 division strategy yields the best recognition rate. It demonstrates that only using the divide-and-conquer scheme could improve the performance when more spatial information is included.

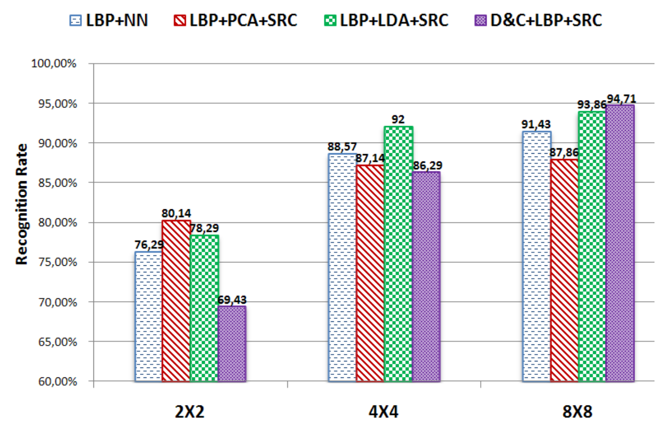

Fig. 4. Recognition rates of four LBP based methods (LBP+NN, LBP+PCA+SRC, LBP+LDA+SRC and D&C+LBP+SRC) on three different division strategies (2 × 2, 4 × 4 and 8 × 8 sub-blocks).

Figure 5 shows the performance of our proposed approach (D&C+LBP+SRC with the pyramid architecture). In the figure, the improvement of recognition rate is significant. Using only the base level (8 × 8 sub-blocks) returns the same result as D&C+LBP+SRC in Figure 4. But the recognition rate increases progressively when adding more levels in the pyramid. The recognition rate tends to converge as the increasing of pyramid levels. Hence it is inefficient to add more levels in the pyramid since the complexity increases accordingly.

It should be noticed that the computational burden of the proposed algorithm is higher than the previously discussed algorithms. Suppose that the proposed algorithm has n levels in the pyramid architecture, it requires to compute $(4^n - 1)/3$ times of $\ell 1$-minimization. Therefore we only used a limited number of pyramid levels (i.e. 2 × 2, 4 × 4 and 8 × 8) in our approach. The experimental results reveal that our algorithm with chosen levels yields significant recognition rates. In addition, when the parallel computing is considered, the complexity of the proposed approach can be reduced to $\Theta(n^{1/2})$ using the algorithm demonstrated in [14].

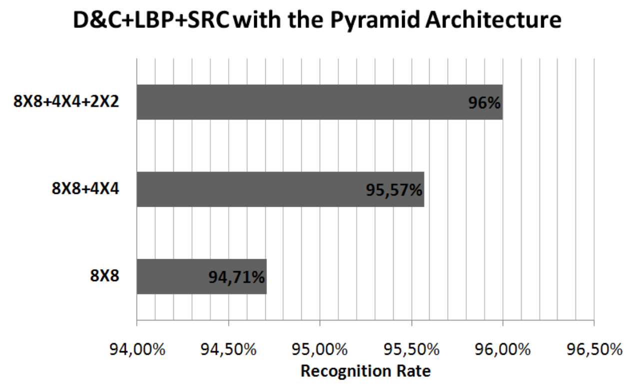

Fig.5. Recognition rates of the proposed approach (D&C+LBP+SRC with the pyramid architecture) in different pyramid levels (8×8, $8 \times 8 + 4 \times 4$ and $8 \times 8 + 4 \times 4 + 2 \times 2$).

On the same dataset, the proposed algorithm reaches the recognition rate at 96%, which is higher than the best result reported by Wright et al. [7] (Fisherface+SRC, up to 94.7%).

5. CONCLUSION

In this paper, we proposed a novel face recognition algorithm which combines LBP based feature extraction with sparse representation based classification (SRC). Comparing to the state-of-art method, the proposed approach is more appropriated for realistic scenarios (usually there are only few training samples available for each class). The dimensionality problem is resolved by applying the divide-and-conquer strategy and the discriminative power is strengthen by a pyramid architecture which incorporates information from different levels of descriptions. The proposed algorithm is compared with the baseline algorithm [1] as well as the methods combining LBP with SRC based on dimension reduction tools (PCA and LDA); it also yields better recognition results than the best result reported by Wright et al. [7] on the same database. Furthermore, the proposed approach is not only restricted to the basic LBP features but also compatible with other high dimensional histogram descriptors which might achieve improved recognition results.

The future works include: 1. exploring other histogram descriptors (such as LGBP, HOG, LPQ etc.) for the potential performance increase; 2. selecting the optimal number of levels for the pyramid architecture; 3. using more sophisticated fusion scheme instead of simple summation when fusing the obtained sparse coefficients vectors.

6. REFERENCES

[1] T. Ahonen, A. Hadid, and M. Pietikäinen, “Face recognition with local binary patterns,” *Computer Vision, ECCV 2004 Proceedings, Lecture Notes in Computer Science* 3021, Springer, 469-481.

[2] X. Huang, S. Z. Li, Y. Wang, "Jensen-Shannon Boosting Learning for Object Recognition," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05),* vol. 2, pp.144-149, 2005

[3] 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,” in *Proc. IEEE Conf. Computer Vision (ICCV),* vol. 1, pp. 786–791, 2005.

[4] C.-H. Chan, J. Kittler and K. Messer, “Multi-scale Local Binary Pattern Histograms for Face Recognition,” *LNCS*, Vol. 4642, pp. 809-818, 2007.

[5] S. Liao and S. Z. Li, “Learning Multi-Scale Block Local Binary Patterns for Face Recognition,” *The 2nd International Conference on Biometrics* (ICB 2007), *Springer LNCS*, vol. 4642, pp.828-837, Aug. 2007.

[6] Z.H. Guo; L. Zhang and D. Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification," *IEEE Transactions on Image Processing*, vol.19, no.6, pp.1657-1663, June 2010

[7] J. Wright, A. Ganesh, A. Yang, and Y. Ma, “Robust face recognition via sparse representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.31, no.2, pp. 210-227, 2009.

[8] Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma, “Face Recognition with Contiguous Occlusion using Markov Random Fields,” *IEEE International Conference on Computer Vision (ICCV)*, 2009

[9] M. Yang and L. Zhang, “Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary,” *11th European Conference on Computer Vision (ECCV 2010),* Sept. 2010.

[10] A. Yang, A. Ganesh, S. Sastry and Y. Ma, “Fast L1-Minimization Algorithms and an Application in Robust Face Recognition: A Review,” In *Proc. of IEEE 17th International Conference on Image Processing (ICIP 2010)*, Sept. 2010.

[11] C.-H. Chan and J. Kittler, “Sparse Representation of (Multiscale) Histograms for Face Recognition robust to Registration and Illumination Problems”, In *Proc. of IEEE 17th International Conference on Image Processing (ICIP 2010)*, Sept. 2010.

[12] D.L. Donoho, “For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution,” *Comm. Pure and Applied Math.*, vol. 59, no. 6, pp. 797–829, 2006.

[13] A. Yang, J. Wright, Y. Ma, and S. Sastry, “Feature selection in face recognition: A sparse representation perspective,” UC Berkeley Tech Report UCB/EECS-2007-99, 2007.

[14] Q. F. Stout, “Supporting Divide-and-Conquer Algorithms for Image Processing,” *Journal of Parallel and Distributed Computing 4* (1987), pp. 95-115.

[15] A. Pentland, B. Moghaddam, and T. Starner. “View based and modular eigenspaces for face recognition. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 1994.

[16] A.M. Martinez and R. Benavente, " The AR face database," Technical report, CVC Technical report, no. 24, 1998.

[17] T. Ojala, M. Pietikainen, T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol.24, no.7, pp.971-987, Jul 2002.

[18] http://www.acm.caltech.edu/l1magic/