

Robust Face Recognition using Voting by Bit-plane Images based on Sparse Representation

Dongmei Wei^{1,2}

¹College of Physics and Electronics, Shandong Normal University

Wenhua donglu 88, Jinan, P.R. China

²School of Information Science and Engineering, Shandong University

Shanda Nanlu 27, Jinan, P.R. China

+8613668805385

E-mail: weidongmei2@163.com

Tianping Li

College of Physics and Electronics, Shandong Normal University

Wenhua donglu 88, Jinan, P.R. China

+8613306408458

E-mail: sdsdlt@163.com

ABSTRACT

Plurality voting is widely employed as combination strategies in pattern recognition. As a technology proposed recently, sparse representation based classification codes the query image as a sparse linear combination of entire training images and classifies the query sample class by class exploiting the class representation error. In this paper, an improvement face recognition approach using sparse representation and plurality voting based on the binary bit-plane images is proposed. After being equalized, gray images are decomposed into eight bit-plane images, sparse representation based classification is exploited respectively on the five bit-plane images that have more discrimination information. Finally, the true identity of query image is voted by these five identities obtained. Experiment results shown that this proposed approach is preferable both in recognition accuracy and in recognition speed.

Keywords

Bit-plane image; voting; sparse representation

1. INTRODUCTION

Face recognition has been researched widely and an enormous amount of approaches have been proposed because of its broad applications in security, law enforcement, human-computer interaction, and entertainment. In which, Principal component analysis (PCA) [1] and Linear discriminant analysis (LDA) [2] have been the most popular subspace analysis techniques to learn the low-dimensional structure of high dimensional face data. Locality Preserving Projections (LPP) [3] preserves the local manifold structures of the face data. Wang [4] constructed the complex face image and recommended Bit-plane generalized PCA (BGPCA).

Sparse representation classification has been paid more attention since Wright [5] regarded face recognition as a globally sparse representation problem and defined Randomfaces. Huang [6] proposed an improved method to produce sparse representation

invariant to image-plane transformation by simultaneously recovering the transformation parameters and sparse representation vector when the test image was not aligned. Sparse representation classification (SRC) [7] introduced an orthogonal matrix, such as identity I , as an occlusion dictionary to code the outlier pixels. Yang [8] innovated Gabor features into SRC and proposed GSRC approach. Robust sparse coding (RSC) [9] casts the sparse coding as a sparsity-constrained robust regression problem. Though these approaches have the higher recognition accuracy, it is impossible to apply to real-time applications because they are time-consuming.

As a combination strategy, voting strategy can achieve better performance and has been widely used in pattern recognition [10, 11, 12]. Lin [13] demonstrated plurality voting is consisted with the Bayesian criterion and it outperforms majority voting in achieving a desirable trade-off between identification accuracy and rejection rate. Paul Watta [14] analyzed the performance of the plurality voting based ensemble classifier.

This paper presents a novel face recognition approach using voting and sparse representation based on binary bit-plane images, having higher recognition rate and rapid speed.

The rest of this paper is organized as follows: Image preprocessing and extraction of bit-plane image are introduced in section 2. The details of the sparse representation are given in Section 3. Proposed approach is presented in section 4. Experiment results and conclusion are given in section 5 and Section 6 respectively.

2. IMAGE PREPROCESS

By extracting each binary digit, the gray image with 256 gray levels can be decomposed into 8 bit-planes, from the least significant bit (bit-plane 1) to the most significant bit (bit-plane 8). The higher-order bit-plane embodies more visually significant data information [15]. Fig. 1 shows an original face image and its 8 bit-planes, the higher-order bit-planes (1-4) carry the majority of the outline features; while the other bit-planes contain the more subtle details in the image.

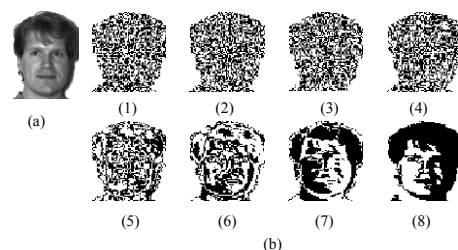


Fig.1 Original face image (a) and its eight bit-planes (b)

In order to enhance the global contrast and to make the subtle details much clearer, all the face images are first equalized using the cumulative distribution function in this proposed approach. The information distribution in the bit-plane images is changed after equalization. Fig.2 provides the equalized image of the one shown in Fig.1 (a) and its 8 bit-planes. There are also outline features in the bit-planes 1 and 2, and the top 3 bit-planes still contain abundance shape information.

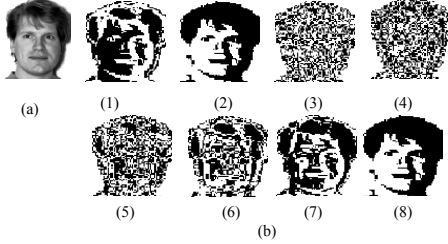


Fig. 2 Equalized face image (a) and its eight bit-planes (b)

3. SPARSE REPRESENTATION CLASSIFICATION

Assume there are K classes and n_k training samples included in the k^{th} class, total number of training samples is $N = \sum_{k=1}^K n_k$. Each face image with size of $w \times h$ can be stacked as an m -dimension vector ($m = w \times h$). Denote the training samples from the k^{th} class as the matrix $\mathbf{X}_k = [\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,l}, \dots, \mathbf{x}_{k,n_k}]$, where $\mathbf{x}_{k,l}$ is an m -dimensional vector corresponding to the l^{th} sample of the k^{th} individual. Over-complete dictionary expressed by $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_K] \in \mathfrak{R}^{m \times N}$ is the concatenation of all training samples. \mathbf{y} expresses the test sample.

Sparse representation based classification (SRC) proves that the test sample \mathbf{y} of subject i can be sparse representation by the over-complete dictionary \mathbf{X} ideally as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} \in \mathfrak{R}^m \quad (1)$$

Where $\boldsymbol{\alpha} = (0, \dots, 0, \boldsymbol{\alpha}_{i,1}, \dots, \boldsymbol{\alpha}_{i,n_i}, 0, \dots, 0)^T$ is sparse representation coefficient vector whose entries are mostly zero except those associated with the i^{th} subject. Eq. (1) means that the sparse representation coefficient vector can be obtained by constrain with l_0 -norm. As the closest convex function to l_0 -norm minimization, l_1 -norm minimization is widely used in spare representation, and both are equivalent if the solution is sufficiently sparse. So the coefficient vector is computed by

$$(\hat{\boldsymbol{\alpha}}) = \arg \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}\|_2 < \varepsilon \quad (2)$$

l_1 -norm minimization is costly, and Zhang [16] demonstrated that the similar classification results can be obtained by replacing the strong l_1 -norm with the much weaker l_2 -norm but with significantly lower complexity. They proposed the collaborative representation based classification with regularized least square (CRC_RLS), in which coefficient vector $\hat{\boldsymbol{\rho}}$ can be computed as

$$(\hat{\boldsymbol{\rho}}) = \arg \min_{\boldsymbol{\rho}} \{ \|\mathbf{y} - \mathbf{X}\boldsymbol{\rho}\|_2^2 + \lambda \|\boldsymbol{\rho}\|_2^2 \} \quad (3)$$

Utilizing regularized least square, we can obtain easily the unique coefficient vector $\hat{\boldsymbol{\rho}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$, where λ is regularization parameter, term of $\lambda \|\boldsymbol{\rho}\|_2^2$ ensures the stability and sparsity embodied in $\hat{\boldsymbol{\rho}}$. Let $\mathbf{P} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T$, \mathbf{P} can be regarded as the projection matrix. At the test stage, \mathbf{y} is only simply projected onto \mathbf{P} and is classified with $\hat{\boldsymbol{\rho}}$ by evaluating which class leads to the minimum reconstruction error. The procedures of the CRC_RLS are summarized as follows.

CRC_RLS algorithm:

- a. Normalize the columns of the dictionary \mathbf{X} to have unit l_2 -norm and denote by $\tilde{\mathbf{X}}$, $\mathbf{P} = (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}} + \lambda \mathbf{I})^{-1} \tilde{\mathbf{X}}^T$
 - Normalize the query sample \mathbf{y} and denote it as $\tilde{\mathbf{y}}$, compute the coefficient vector $\hat{\boldsymbol{\rho}} = \mathbf{P}\tilde{\mathbf{y}} = (\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_K)^T$, where $\boldsymbol{\rho}_k \in \mathfrak{R}^{L \times 1}$ is the coefficient vector associated with class k .
 - c. Reconstruct K vectors $\tilde{\mathbf{y}}_k = \tilde{\mathbf{X}}_k \boldsymbol{\rho}_k$ ($k=1 \dots K$).
 - d. Compute the error $r_k = \frac{\|\tilde{\mathbf{y}} - \tilde{\mathbf{X}}_k \boldsymbol{\rho}_k\|_2}{\|\boldsymbol{\rho}_k\|_2}$
 - e. Identity of \mathbf{y} is $k = \arg \min_k \{r_k\}$
-

4. FACE RECOGNITION USING VOTING AND SPARSE REPRESENTATION

The proposed face recognition approach using plurality voting strategy and sparse representation based on bit-plane images (SRV_BP) includes image preprocessing, bit-plane selection and on-line recognition.

4.1 Image Preprocess

Eight training sets are obtained from the original training set after training images are preprocessed according to Section 1. Let \mathbf{x}_{kl}^i denote the i^{th} bit-plane image from the l^{th} image of the class k^{th} and $\mathbf{X}^i = [\mathbf{x}_{11}^i, \dots, \mathbf{x}_{kl}^i, \dots, \mathbf{x}_{KL}^i]$ denote the i^{th} over-complete dictionary corresponding to the i^{th} bit-plane ($i=1, 2, \dots, 8$).

4.2 Bit-plane Selection

Randomly select M ($M < L$) training images from each subject for training, and the rest for testing. $\mathbf{A}^i = [\mathbf{A}_1^i, \mathbf{A}_2^i, \dots, \mathbf{A}_k^i]$ is the training matrix corresponding to the i^{th} bit-plane, where $\mathbf{A}_k^i = [\mathbf{x}_{k1}^i, \dots, \mathbf{x}_{kM}^i]$. All the i^{th} bit-plane test images make up the set $\mathbf{S}^i = \{\mathbf{x}_{kt}^i \mid k=1, \dots, K, t=M+1, \dots, L\}$. Let \mathbf{A}^i instead of \mathbf{X} , the i^{th} project matrix is denoted as $\mathbf{P}^i = [(\mathbf{A}^i)^T \mathbf{A}^i + \lambda \mathbf{I}]^{-1} (\mathbf{A}^i)^T$. The right recognition rate corresponding to the i^{th} bit-plane ξ_i is computed by replacing \mathbf{y} with each element \mathbf{x}_{kt}^i of \mathbf{S}^i and repeating the step (b)~(e) of the algorithm CRC_RLS. Integer i varies from 1 to 8, and the recognition rate vector is

calculated as $\xi = (\xi_1, \dots, \xi_8)$. Fig.3 and Fig.4 show the variation recognition rate of each bit-plane with the number of principal components retained, where BP1 to BP8 are abbreviations for the bit-plane 1 to bit-plane 8, respectively. The recognition rates of bit-plane 2 and of bit-plane 8 are equal no matter how many features are retained, so the two curves, BP2 and BP8, overlap together.

It is clear that the 3rd and the 4th bit-planes contain more texture information which is unhelpful for classification, and experiment validated their recognition rates are always less than 0.25. So both bit-planes are discarded in the procedure of plurality voting. The 2nd bit-plane has the same recognition rate to the 8th bit-plane and they will bring the same identity for the test image, so the 2nd bit-plane is also discarded during the processing of voting.

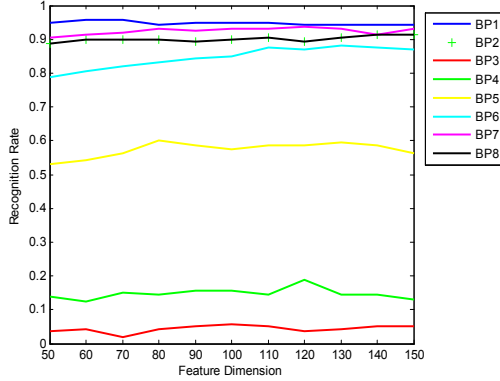


Fig.3 Variation in recognition rate of each bit-plane with the number of principal components retained on ORL database

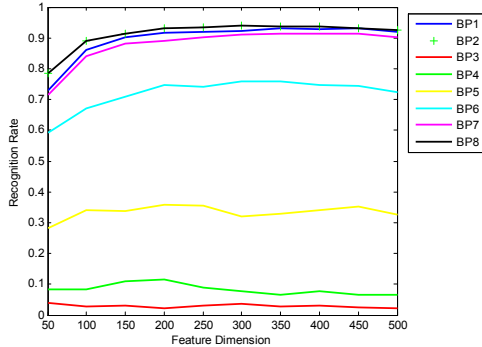


Fig.4 Variation in recognition rate of each bit-plane with the number of principal components retained on AR database

4.3 On-line Recognition

The query image y is equalized and is decomposed into eight binary bit-plane images $y^i (i=1,2,\dots,8)$. In CRC_RLS algorithm, replace X and y with X^i and y^i respectively, and the identity of y^i is computed, k_i , an integer between 1 to K , denoting the identity of $y^i (i=1,5,6,7,8)$. Let $d_1 = k_1, d_2 = k_5, d_3 = k_6, d_4 = k_7, d_5 = k_8$, and assume there are 5 vectors, $v_1, \dots, v_5 \in \mathbb{R}^K$, their values of $v_p (p=1, \dots, 5)$ are assigned by Eq. (4):

$$v_p(m) = \begin{cases} 1 & d_p = m \\ 0 & \text{others} \end{cases} \quad (4)$$

The total number of votes received by class m is presented by $v_m = \sum_{p=1}^5 v_p(m)$, v_m presents a measure of the similarity

between the input and class m . Let vector $v = (v_1, v_2, \dots, v_K)$, and suppose v_M is the only largest element, i.e. $v_M = \max\{v_m, m=1, \dots, K\}$, then M , as the subscript of v_M , denotes the identity of y . If there are more than one subjects having the max number votes, for example, $v_{M_1} = v_{M_2} = \max\{v_m, m=1, \dots, K\}$, the plurality voting fails. In this case, the plurality voting is failure and the testing is rejected to recognition.

5. EXPERIMENTAL RESULTS

In this section, we test the performance of the proposed algorithm (SRV_BP) and compare it with PCA, SRC, CRC_RLS on the AR database and ORL database.

5.1 ORL Face Database

ORL database contains 400 grayscale images from 40 individuals, each individual having 10 images. These images are normalized to 64×64 . For each individual, 6 images are used for training and the rest are used for testing.

For 160 testing images, there are 2 images could not be recognized by plurality voting when the feature dimension equals to 95. Table 1 gives the experiment results.

Table 1 the recognition rates on ORL database (%)

Feature dimension	85	90	95	100	150	400	1000
PCA	86.8	87.5	88.1	87.5	87.5	--	--
SRC	85	79.4	77.5	83.1	88.7	92.5	86.9
CRC_RLS	91.8	92.5	93.1	92.5	90.6	--	--
SRV_BP	94.0	95.2	96.3	95.2	95.1	--	--

5.2 AR Face Database

AR database contains 3288 color face images of 116 individuals, we utilize the light component to be the gray images and crop the face portion of the images whose size are 60×43 . Taking 1400 images from 100 persons to experiment, each person has 14 images. 7 images are used for training and the rest 7 images are used for testing.

When the feature dimension is 300, 11 testing samples of 700 testing images could not be recognized by plurality voting and their identities are decided by the spare second identification. The recognition rates are shown as table 2.

Table 2 the recognition rates on AR database (%)

Feature dimension	54	120	300
PCA	68	70.1	71.3
SRC	83.3	89.5	93.3
CRC_RLS	80.5	90	93.8
SRV_BP	83.3	92.4	96.0

5.3 Comparison of recognition speed

We compare the recognition speed of the proposed method SRV_BP with the competing approaches SRC and CRC_RLS.

The results of ORL and AR are listed in table 3 and table 4 respectively. Optimal feature-dimension is the minimum feature-dimension when the recognition rate reaches highest. The testing time is the average running time to accomplish one testing under the optimal feature-dimension.

Table 3 Speed on ORL database

Method	Recognition rate (%)	Optimal feature-dimension	Testing time (s)
SRC	92.5	400	0.3452
CRC_RLS	93.1	95	0.0087
SRV_BP	96.1	94.5	0.0300

Table 4 Speed on AR database

Method	Recognition rate (%)	Optimal feature-dimension	Testing time (s)
SRC	93.3	300	1.8258
CRC_RLS	93.8	300	0.0051
SRV_BP	96.0	300	0.0289

The proposed method SRV_BP includes image histogram equalization, bit-plane extraction, running CRC_RLS several times and plurality voting. Hence SRV_BP spends more time than CRC_RLS in recognizing, but its speed is faster than the SRC prominently.

6. CONCLUSION

A new face recognition approach with sparse representation and plurality voting based on the bit-plane images has been discussed in this article. According to the recognition rate and order of each bit-plane, five binary bit-plane images with more discrimination information are selected to vote the identity of testing image. In each bit-plane images, the testing image is represented sparsely by the training images under the l_2 -norm constraint, which solves the "lack of samples" problem in face recognition and avoids the expensive l_1 -minimization. Experimental results indicate that the proposed approach is superior both in the recognition rate and the speed.

7. ACKNOWLEDGMENT

This study was sponsored by Project of Shandong Province Higher Educational Science and Technology Program

8. REFERENCE

[1] Turk, M, Pentland, A. 1991. Eigenfaces for recognition. *J. Cognitive Neuroscience*, 3(1): 71–86.

[2] Belhumeur, P.N, Hespanha, J.P, Kriegman, D.J. 1997. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE T Pattern Anal Mach Intell*, 19(7): 711–720.

[3] He, X. Yan, S. Hu, Y. Zhang, H. 2003. Learning a locality preserving subspace for visual recognition. *In Proceedings of IEEE International Conference on Computer Vision (Nice, France)*, pp 385-392.

[4] Wang, H.Y. Leng, Y. Wang, Z. Wu, X. 2007. Application of image correction and bit-plane fusion in generalized PCA based face recognition. *Pattern Recognition Lett*, 28(16):2352-2358.

[5] Yang, A. Wright, J. MaM Y. Shankar, S. 2007. Feature selection in face recognition: A sparse representation perspective. Technical Report UCB/ECS-2007-99, UC Berkeley.

[6] Huang, J.Z. Huang, X.L. Metaxas, D. 2008. Simultaneous image transformation and sparse representation recovery. *In*

Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, (Anchorage, AK), pp 1-8.

[7] Wright, J. Yang, A. Ganesh, A. Sastry, S. Ma, Y. 2009. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal Mach Intell*, 31 (2) 210 – 227.

[8] Yang, M. Zhang, L. 2010. Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. *In Proceedings of ECCV'10 Proceedings of the 11th European conference on Computer vision: Part VI. Heraklion (Crete Greece: Springer-Verlag Berlin, Heidelberg)*, pp 448-462.

[9] Yang, M. Zhang, L. Yang, J. Zhang, D. 2010. Robust sparse coding for face recognition. *In Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*. (Providence, RI). pp 625-632.

[10] Erp, M. Vuurpijl, L. Schomaker, L. 2002. An overview and comparison of voting methods for pattern recognition. *In Proceedings of the eighth international workshop on frontiers in handwriting recognition*. Canada, pp 195–200.

[11] Ho, T.K. Hull, J.J. Srihari, S.N. 1994. Decision combination in multiple classifier systems. *IEEE T Pattern Anal Mach Intell*, 16(1):66–75.

[12] Kittler, J. Hatef, M. Duin, R. Matas, J. 1998. On combining classifiers. *IEEE Trans. Pattern Anal Mach Intell*, 20(3):226–239.

[13] Lin, X. Yacoub, S. Burns, J. Simske, S. 2003. Performance analysis of pattern classifier combination by plurality voting. *Pattern Recognition Lett*, 24(12):1959–1969.

[14] Mu X.Y. Watta, P. Hassoun, M.H. 2008. Analysis of a plurality voting-based combination of classifiers. *IEEE International Joint Conference on Neural Networks*. pp: 304 – 309.

[15] Rafael, C. G. Richard, E.W. 2003. Digital Image Processing. Prentice Hall, NJ, USA.

[16] Zhang, L. Yang, M. Feng, X. 2011. Sparse representation or collaborative representation: Which helps face recognition? *In Proceedings of IEEE International Conference on Computer Vision*, Barcelona, pp 471-478.