A NOVEL SRC BASED METHOD FOR FACE RECOGNITION WITH LOW QUALITY IMAGES

*Shicheng Yang, Ying Wen**

Shanghai Key Laboratory of Multidimensional Information Processing
Department of Computer Science and Technology
East China Normal University, Shanghai, China, 200062

ABSTRACT

Sparse representation-based classification (SRC) shows a good performance for face recognition in recent years, but SRC can not be suitable for low quality data with disguise or noise, which are often presented in the practical applications. To solve the problem, in this paper, we propose a novel SRC based method for face recognition with low quality images named sparse low-rank component based representation (SLCR). In SLCR, we utilize the low-rank component from training dataset to construct dictionary. The dictionary composed of low-rank component and non-low-rank component is able to describe the face feature better, especially for low quality training samples. Our recognition rule is based on the minimum class-wise reconstruction residual which leads to a substantial improvement on the proposed SLCR's performance. Extensive experiments on benchmark face databases demonstrate that the proposed method consistently outperforms the other sparse representation based approaches for disguised and corrupted face recognition.

Index Terms— Face recognition, sparse representation, classification, low-rank component, disguised and corrupted training dataset

1. INTRODUCTION

Face recognition is the most popular biometric approach in the past several decades due to its huge application potentials [1], [2], [3], [4], [5], [6]. Feature extraction is important for face recognition. The common techniques are principal component analysis (PCA) [7], linear discriminant analysis (LDA) [8], probabilistic subspace learning [9] and locality preservation [10] and so on. However, these techniques are hard to solve low quality images with disguise and noise [11]. Recently, some works on robust PCA have been proposed to alleviate this problem [12], [13], [14]. Thereinto, low-rank matrix recovery (LR) [12] is a good technique, which can separate corruption better from the training face images than PCA.

In addition, the classifier is also important for face recognition. Sparse representation-based classification (SRC) has been proposed and achieved satisfied results [4]. However, the SRC cannot perform well when the dataset is disguised or corrupted. Thus, some extended SRC methods have been proposed [15],[16],[17],[18],[19]. Deng et al. proposed a superposed sparse representation-based classifier (SSRC) for undersampled face recognition [5], [20]. Chen et al. proposed a low-rank matrix approximation algorithm with structural incoherence (LRSI) combined SRC [21]. Jiang et al. proposed a sparse- and dense-hybrid representation (SDR) framework to alleviate the problems of SRC [6]. Although these extended SRC methods can improve the classifiers' performance, they also present more or less unsatisfied results for disguised and corrupted data.

In this paper, we propose a novel SRC based method for face recognition with low quality images named sparse low-rank component based representation (SLCR) which is robust to disguised and corrupted face recognition. In this method, we apply the low-rank component to the training set to construct the dictionary. The low-rank component obtained by low-rank matrix recovery from the training dataset can separate the effective feature and the corrupted component associated disguises or noise, which can be helpful to accurately recognize. Then we obtain the solution of the proposed SLCR by the Augmented Lagrange Multiplier (ALM) scheme. Finally, we minimize class-wise reconstruction residual to recognize the test image. The experimental results on the extended Yale B, CUM Multi-PIE and AR databases validate that our method performs well for face recognition.

2. RELATED WORKS

Suppose that there are N training images from C object classes. Then, define a training dataset $D = [D_1, D_2, ..., D_C] \in R^{d \times N}$, where D_i consists of the training images of ith class as its columns and d is the dimension of each sample. In sparse representation-based classification (SRC) algorithm, given a

This work was supported by Natural Science Foundation of Shanghai no. 17ZR1408200 and the Science and Technology Commission of Shanghai Municipality under research grant no. 14DZ2260800.
* Corresponding author. E-mail: ywen@cs.ecnu.edu.cn.

test image $y \in R^{d \times 1}$, and the linear representation of y can be rewritten in terms of all training images as:

$$y = D\alpha + z \tag{1}$$

where $\alpha \in R^N$ is a sparse coefficient vector whose entries are zeros except those associated with the ith class and $z \in R^d$ is a noise term with bounded energy $\|z\|_2 < \varepsilon$. The sparse solution α can be approximately recovered by solving the following stable l_1-minimization problem [4]:

$$\min_{\alpha, z} \|\alpha\|_1 + \beta \|z\|_2^2, s.t.\ y = D\alpha + z \tag{2}$$

where β is a constant for a compromise between sparsity and reconstruction.

SRC cannot perform well when training images are disguised or corrupted. Furthermore, Deng et.al proposed a prototype plus variation (P+V) representation model. Then, the linear representation of y can be rewritten in terms of all training images as [5]:

$$y = A\alpha + Bx + z \tag{3}$$

where A is the prototype dictionary, B is the variation dictionary, α and x can be recovered simultaneously by l_1-minimization. Note that the residual is related to α and x, and it is computed as follow:

$$r_i(y) = \left\| y - [A, B] \begin{bmatrix} \delta_i(\alpha) \\ x \end{bmatrix} \right\|_2 \tag{4}$$

where $\delta_i(\alpha)$ is a new vector whose only nonzero entries are the entries in α that are associated with class i. They further proposed a superposed SRC (SSRC) based on this model, in which the prototype dictionary A is the geometric centroid per class and the variation matrix B is constructed by the sample based difference to the centroid.

Similar to the P+V model, Jiang et.al proposed a sparse-and dense-hybrid representation (SDR) framework, in which dictionary A (i.e., class-specific component) contains identity information and dictionary B is non-class-specific component. Note that the representation residual in SDR is defined by

$$r_i(y) = \|z - z_i\|_2 = \|A(I - C_i)\alpha\|_2 \tag{5}$$

where I is an identity matrix, C_i is a class-label matrix of the training dataset D for class i, its element $C_i(k, k) = 1$ if the kth training image originates from class i and all other elements of C_i are zero. Subsequently, they proposed a procedure of supervised low-rank (SLR) dictionary decomposition to facilitate the proposed SDR framework (SDR-SLR) [6].

3. SPARSE LOW-RANK COMPONENT BASED REPRESENTATION

In this paper, we propose a new sparse low-rank component-based representation (SLCR) for face recognition. We begin with the motivations of our work.

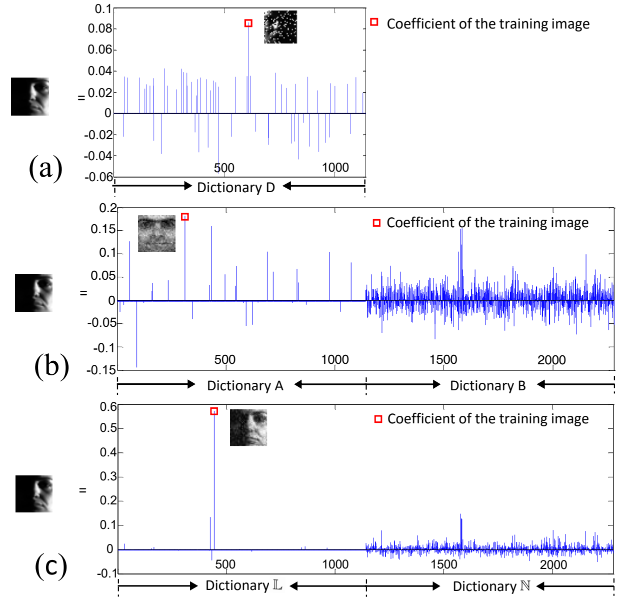

Fig. 1. Recognition with 32×32 training images corrupted by 20% salt-and-pepper noise using (a)SRC, (b)SDR-SLR and (c)SLCR.

SSRC simply used centroid images to capture the class-specific information. SDR-SLR used the reconstructed images by the singular vectors corresponding to the largest singular value to initialize dictionary, in which SVD is used to obtain the dictionary. The dictionary construction using SDR-SLR is better than that of SSRC for face recognition [6]. SVD is same as PCA. By PCA, the training dataset D can be initialized by

$$D = L + N \tag{6}$$

where L is the principal component, N is the non-principal component. It finds that the best rank-k estimation of L by minimizing $\|D - L\|_2$ subject to $rank(L) <= k$ and it can be solved by SVD. If the image is corrupted by Gaussian noise, the principal component obtained by PCA can get optimal [12]. However, PCA is sensitive to small non-Gaussian noise often presented in practical face images. This means that the information captured by PCA remains potential corruption.

In general, the dictionary that only contains class-specific information is the low-rank matrix. This is supported by the fact that face images within a class have a low-rank structure [22]. Thus, whatever the noise is Gaussian, we hope to decompose the training sample matrix D into the low-rank component (i.e., principal component) $\mathbb{L}$ and the non-low-rank component $\mathbb{N}$. And the low-rank component $\mathbb{L}$ is used to describe face feature and the non-low-rank component $\mathbb{N}$ contains the information associated sparse error.

By low-rank matrix recovery, the training matrix D can

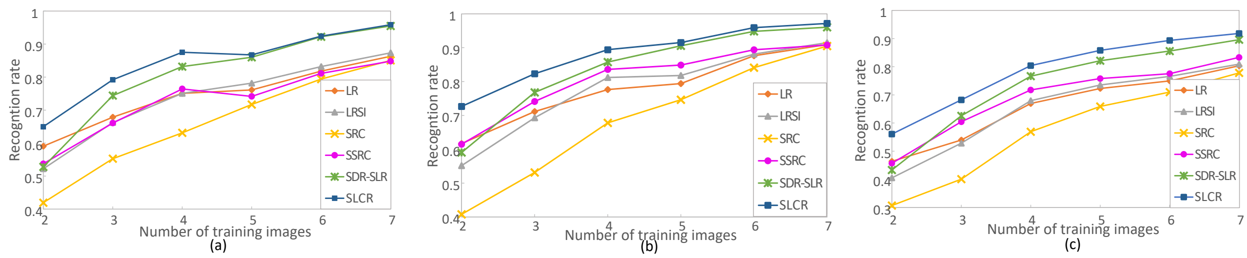

Fig. 3. Experimental results with (a) glasses disguised, (b) scarf disguised and (c) glasses and scarf disguised images on the AR database.

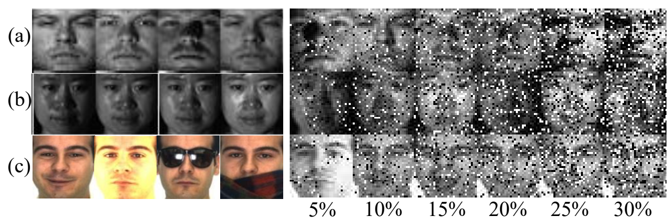

Fig. 2. Some cropped images and some training samples corrupted by salt-and-pepper noise from the (a) extended Yale B, (b) CMU Multi-PIE and (c) AR database.

be initialized by

$$D = \mathbb{L} + \mathbb{N} \tag{7}$$

where $\mathbb{L}$ is the low-rank component from original training matrix D and $\mathbb{N}$ is the non-low-rank component that associated sparse error. This formulation suggests that LR seeks the lowest rank $\mathbb{L}$ that contains all most class-specific information. The lowest rank $\mathbb{L}$ can be approximately recovered by solving the following convex surrogate

$$\min_{\mathbb{L},\mathbb{N}} \|\mathbb{L}\|_* + \gamma \|\mathbb{N}\|_1 , s.t. D = \mathbb{L} + \mathbb{N} \tag{8}$$

where the nuclear norm $\|\mathbb{L}\|_*$, the sum of the singular values, approximates the rank of $\mathbb{L}$ and γ is a constant for a compromise between $\mathbb{L}$ and $\mathbb{N}$. Then, we use the low-rank component $\mathbb{L}$ to construct dictionary.

Once the dictionary $\mathbb{L}$ is constructed from training data, we perform recognition of test image y as

$$y = \mathbb{L}\alpha + \mathbb{N}x + z \tag{9}$$

where $\mathbb{L}$ is low-rank dictionary, $\mathbb{N}$ is variation dictionary that associated noises, outlier pixels and occlusions and z is reconstruction error. Eq.(9) is the proposed sparse low-rank component based representation (SLCR). The sparsity of α is measured by l_0-norm of α. But this problem is NP-hard, we replace l_0-norm of α by l_1-norm of α, i.e. $\|\alpha\|_1$. In order to make the representation error z as small as possible, it is not necessary to put the sparse constraint on x and therefore we use $\|x\|_2$. The solution of the proposed SLCR, α, x and z, is obtained by solving the following optimization problem:

$$\min_{\alpha,x,z} \|\alpha\|_1 + \beta \|x\|_2^2 + \gamma \|z\|_1 , s.t. y = \mathbb{L}\alpha + \mathbb{N}x + z \tag{10}$$

where β and γ are constants for a compromise. We solve the optimization problem by the Augmented Lagrange Multiplier (ALM) scheme [23]. Finally, our recognition rule is also based on the minimum class-wise reconstruction residual. The class-wise reconstruction residual is defined by

$$r_i(y) = \|z - z_i\|_2 = \|\mathbb{L}(I - C_i)\alpha\|_2 \tag{11}$$

where $\mathbb{L}$ is dictionary SLCR used.

Next, we implement an experiment to show the difference among SRC, SDR-SLR and SLCR. We randomly select 30 images with 20% salt-and-pepper noise from the extended Yale B database as training set and randomly select a test image to get the results of SRC, SDR-SLR and SLCR. To illustrate the difference among them, Fig.1 (a), (b) and (c) show the sparse coefficients of a testing image using SRC, SDR-SLR and SLCR, respectively. It is worth noting that sparse coefficients obtained by SLCR are sparser than those of SRC and SDR-SLR. In addition, due to the noise in training samples of the correct subject, SRC and SDR-SLR tend to select the samples of many other subjects to represent the test image. The most significant coefficient of SRC anf SDR-SLR is associated with the wrong subject that leads to misclassification in this example. Contrary to SRC and SDR-SLR, the top significant coefficient of an example of SLCR is for the training image of the same identity as the test image. This means SLCR obtains a correct result in such a situation. From this example, we can see that our proposed SLCR method obtains more sparse coefficients and the value of the top significant coefficient is more accurate.

4. EXPERIMENTS

We choose the extended Yale B, CUM Multi-PIE and AR face databases to compare the performance of our method with LR, LRSI, SRC, SSRC and SDR-SLR in different circumstances. All images we used in each experiment are cropped with the size of 32×32 and all experiments are repeated 10

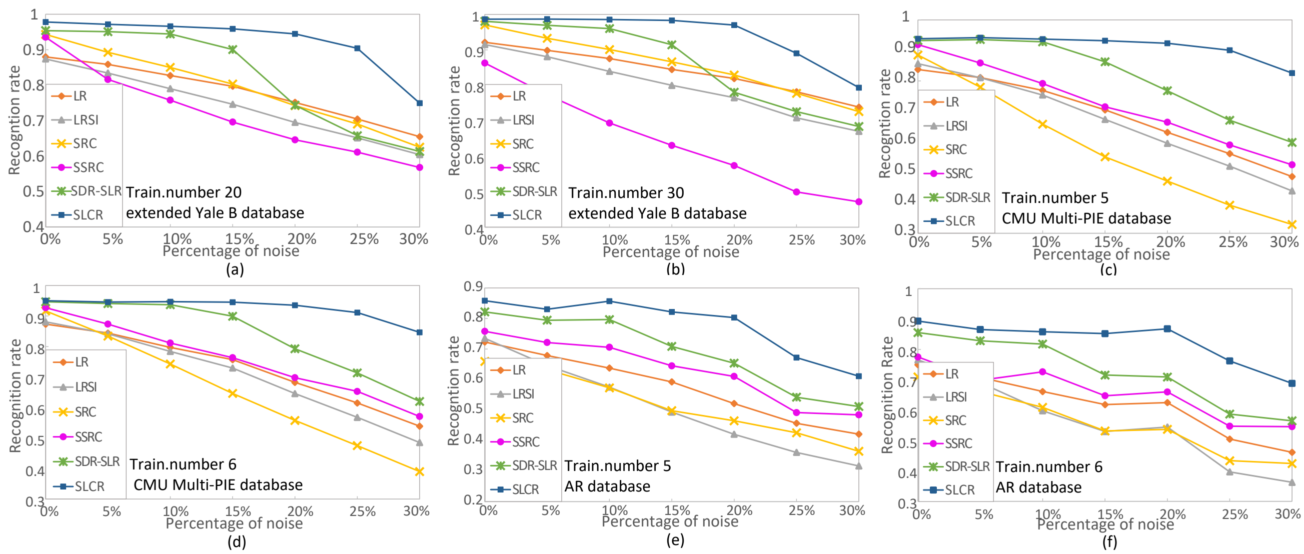

Fig. 4. Experimental results on corrupted training images.

times. Some cropped images from three databases are shown in the left part of Fig.2.

4.1. Experiments on Disguised Training Dataset

To verify the performance of SLCR, we choose AR dataset to testify our method. AR database includes facial variations and disguises. We randomly choose 2-7 images per individual as the training set and the remaining images are for the test set. The average results on disguised database are plotted in Fig.3, respectively. Fig.3(a) shows the result when training set consists of sunglass and neutral images. Fig.3(b) indicates the result when training set comprises scarf and neutral images. Fig.3(c) denotes the result when training set contains sunglass, scarf and neutral images. As shown in the Fig.3, SLCR consistently outperforms SDR-SLR, SSRC, LR and LRSI. It can be seen that SLCR is robust for disguised face recognition.

4.2. Experiments on Images Corrupted by Noise

In the experiments, to test the noise robustness of the proposed method, we use the extended Yale B, CUM Multi-PIE and AR databases and all training samples are corrupted by different level noise. In the right part of Fig.2, from top to bottom, training images are from the extended Yale B, CUM Multi-PIE and AR databases respectively, and from left to right, training images are corrupted by salt-and-pepper noise from 5% to 30%, respectively. Considering different databases having different number samples, we randomly choose 20 and 30 images per individual from extended Yale B database, 5 and 6 images per individual from PIE and AR databases as the training set and the rest as the testing set, respectively. Fig.4 plots the average results of different training set in different databases. From Fig.4, we can see that the noise in training set indeed impacts the performance of face recognition and the performances of aforementioned methods decline with increasing level of salt-and-pepper noise. But from Fig.4, it is noteworthy that SLCR consistently has better performance than the others, particularly when the level of salt-and-pepper noise is more than 10%. Furthermore, the superiority of SLCR is obvious with the gradually increase of salt-and-pepper noise. Nevertheless, the proposed SLCR outperforms all other algorithms consistently for all levels of salt-and-pepper noise. The above experiments demonstrate that SLCR is robust to the noise.

From the above experiments on the disguised and noise datasets, we can see that the proposed method has a good performance for face recognition with low-quality images.

5. CONCLUSION

We proposed a novel SRC based method for face recognition with low quality images named sparse low-rank component based representation (SLCR). In this method, we employed LR on the training dataset to obtain low-rank component and non-low-rank component represented the effective feature and the information with sparse error respectively to construct dictionary. We obtained the solution of the proposed SLCR by the ALM scheme and minimized class-wise reconstruction residual to recognize the test image. In this way, SLCR achieved better classification performance and overcame the problem of insufficient representative training dataset. The experiments demonstrated that the proposed SLCR method outperforms the other SRC based methods for disguised and corrupted face recognition.

6. REFERENCES

[1] Y. Sun, X. Wang, and X. Tang, "Deeply learned face representations are sparse, selective, and robust," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 2892–2900.

[2] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in *European Conference on Computer Vision*. Springer, 2016, pp. 499–515.

[3] O. Arandjelovic, "Gradient edge map features for frontal face recognition under extreme illumination changes," in *Proceedings of the British machine vision association conference*. BMVA Press, 2012, pp. 1–11.

[4] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, 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.

[5] W. Deng, J. Hu, and J. Guo, "Extended src: Undersampled face recognition via intraclass variant dictionary," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1864–1870, 2012.

[6] X. Jiang and J. Lai, "Sparse and dense hybrid representation via dictionary decomposition for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 5, pp. 1067–1079, 2015.

[7] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[8] P.N. Belhumeur, J.P. Hespanha, and David J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[9] B. Moghaddam, "Principal manifolds and probabilistic subspaces for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 780–788, 2002.

[10] X. He, S. Yan, Y. Hu, P. Niyogi, and H.J. Zhang, "Face recognition using laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328–340, 2005.

[11] F. De la Torre and M.J. Black, "Robust principal component analysis for computer vision," in *IEEE International Conference on Computer Vision*. IEEE, 2001, vol. 1, pp. 362–369.

[12] E.J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM (JACM)*, vol. 58, no. 3, pp. 11, 2011.

[13] F. De la Torre and M.J. Black, "A framework for robust subspace learning," *International Journal of Computer Vision*, vol. 54, no. 1-3, pp. 117–142, 2003.

[14] Q. Ke and T. Kanade, "Robust l_1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2005, vol. 1, pp. 739–746.

[15] Z. Zhou, A. Wagner, H. Mobahi, J. Wright, and Y. Ma, "Face recognition with contiguous occlusion using markov random fields.," in *IEEE International Conference on Computer Vision*, 2009, pp. 1050–1057.

[16] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 372–386, 2012.

[17] H. Zhang, J. Yang, Y. Zhang, N.M. Nasrabadi, and T.S. Huang, "Close the loop: Joint blind image restoration and recognition with sparse representation prior," in *IEEE International Conference on Computer Vision*. IEEE, 2011, pp. 770–777.

[18] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in *IEEE International Conference on Computer Vision*. IEEE, 2011, pp. 543–550.

[19] J. Lai and X. Jiang, "Class-wise sparse and collaborative patch representation for face recognition," *IEEE Transactions on Image Processing*, vol. 25, no. 7, pp. 3261–3272, 2016.

[20] W. Deng, J. Hu, and J. Guo, "In defense of sparsity based face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 399–406.

[21] Y.CF. Wang, C.P. Wei, and C.F. Chen, "Low-rank matrix recovery with structural incoherence for robust face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, pp. 2618–2625.

[22] Y. Zhang, Z. Jiang, and L.S Davis, "Learning structured low-rank representations for image classification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 676–683.

[23] D. Bertsekas, *Constrained optimization and Lagrange multiplier methods*, Academic press, 2014.