SPARSE LOW-RANK COMPONENT CODING FOR FACE RECOGNITION WITH ILLUMINATION AND CORRUPTION

Shicheng Yang[1], *Ying Wen*[1]*, *Lianghua He*[2]*

1. Shanghai Key Laboratory of Multidimensional Information Processing,
Department of Computer Science and Technology,
East China Normal University, Shanghai, China.
2. Department of Computer Science and Technology, Tongji University, Shanghai, China.

ABSTRACT

Sparse representation-based classification shows a good performance for face recognition in recent years, but it can not be suitable for face recognition with illumination and corruption, 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 named sparse low-rank component coding (SLC). In SLC, we utilize the low-rank component from training dataset to construct dictionary. The dictionary composed of low-rank component is able to describe the face feature better, especially for training samples with illumination and corruption. Our recognition rule is based on the minimum class-wise reconstruction residual which leads to a substantial improvement on the performance of SLC. Extensive experiments on benchmark face databases demonstrate that the proposed method consistently outperforms the other sparse representation based approaches for face recognition with illumination and corruption.

Index Terms— Face recognition, sparse representation, classification, low-rank component, illumination and corruption

1. INTRODUCTION

Face recognition is the most popular biometric approach due to its huge application potentials in the past several decades [1], [2]. Feature extraction is important for face recognition. The common techniques are principal component analysis (PCA) [3], linear discriminant analysis (LDA) [4], probabilistic subspace learning [5] and locality preservation [6] and so on. However, these techniques are hard to solve the images with illumination and corruption [7]. Recently, some works based on robust PCA have been proposed to alleviate this problem [8], [9], [10]. Thereinto, low-rank matrix recovery (LR) [8] 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 [11]. However, when the dataset is with illumination and corruption, the SRC cannot perform well. Thus, some extended SRC methods have been proposed [12], [13], [14]. Chen et al. proposed a low-rank matrix approximation algorithm with structural incoherence (LRSI) combined SRC [15]. Yang et.al proposed a robust sparse coding (RSC) by modeling the sparse coding as a sparsity constrained robust regression problem [16]. Wei et al. proposed a classification algorithm for face recognition via a robust auxiliary dictionary learning (RADL) [17]. Although these extended SRC methods can improve the classifiers' performance, they also present more or less unsatisfied results for face recognition with illumination and corruption.

In this paper, we propose a novel SRC based method named sparse low-rank component coding (SLC) which is robust to face recognition with illumination and corruption. 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 component associated illumination or corruption, which can be helpful to accurately recognize. Then we obtain the solution of the proposed SLC by minimizing the influence of residual. Finally, we minimize class-wise reconstruction residual to recognize the test image. The experimental results on the CMU Multi-PIE, Extended Yale B and CAS-PEAL databases validate that our method performs well for face recognition with illumination and corruption.

2. RELATED WORKS

Sparse representation-based classification (SRC) is an effective technique for face recognition. Suppose that there are N training images from C object classes. Then, define a training

This work was supported by National Nature Science Foundation of China (Nos. 61773166, 61772369), Natural Science Foundation of Shanghai no. 17ZR1408200 and the Science and Technology Commission of Shanghai Municipality under research grant no. 14DZ2260800.

*Corresponding author. Email: ywen@cs.ecnu.edu.cn (Ying Wen), helianghua@tongji.edu.cn (Lianghua He).

dataset $D = [D_1, D_2, ..., D_C] \in R^{d \times N}$, where D_i consists of the training images of the ith class as its columns and d is the dimension of each sample. In SRC, given a test image $y \in R^{d \times 1}$, and the linear representation of y can be rewritten in terms of all training images as [11]:

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

where $\alpha \in R^N$ is a sparse coefficient vector whose entries are zeros except those associated with the ith class and $e \in R^d$ is a noise term with bounded energy $\|e\|_2 < \varepsilon$. The sparse solution α can be approximately recovered by solving the following stable l_1-minimization problem [11]:

$$\min_{\alpha} \|\alpha\|_1 + \beta \|y - D\alpha\|_2^2 \tag{2}$$

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

SRC cannot perform well when training images are with illumination and corruption. Yang et.al modelled the sparse coding as a sparsity constrained robust regression problem and proposed the robust sparse coding (RSC). It can be formulated as the following minimization [16]:

$$\min_{\alpha} \|\alpha\|_1 + \beta\rho\,(y - D\alpha) \tag{3}$$

where residual function $\rho\,(\cdot)$ is defined as

$$\begin{aligned} \rho\,(e) &= \sum_{k=1}^{t} (e_k) \\ \rho\,(e_k) &= -\frac{1}{2\mu}(\ln(1 + \exp(-\mu e_k^2 + \mu\delta)) \\ &\quad - \ln(1 + \exp \mu\delta)) \end{aligned} \tag{4}$$

where t is the total number of e_k, e_k is the kth entry of $e = y - D\alpha$, and μ and δ are the parameters in the residual function.

Similar to RSC, Wei et.al proposed a classification algorithm for face recognition via robust auxiliary dictionary learning (RADL). In RADL, the auxiliary dictionary A is learned from external data for properly handling intra-class variants. Instead of solving Eq.(3), RADL considers the following minimization problem [17]:

$$\min_{\alpha} \|\alpha\|_1 + \beta\rho\,(y - D\alpha - Ax) \tag{5}$$

where A is the auxiliary dictionary learned from external data, x is the coefficient of the auxiliary dictionary A, and the residual function $\rho\,(\cdot)$ is the same as RSC used.

3. SPARSE LOW-RANK COMPONENT CODING

3.1. The Model of SLC

Principal component analysis (PCA) is often used for feature extraction in face recognition. By PCA, the training dataset D can be initialized by

$$D = P + Q \tag{6}$$

where P is the principal component, Q is the non-principal component. It finds that the best rank-k estimation of P by minimizing $\|D - P\|_2$ subject to $rank(P) <= k$ and it can be solved by SVD. If an image is corrupted by Gaussian noise, the principal component obtained by PCA can get optimal [8]. 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.

Generally speaking, the dictionary that only contains class-specific information is a low-rank matrix. This is supported by the fact that face images within a class have a low-rank structure [18]. Thus, whatever the noise is Gaussian, we hope to decompose the training sample matrix D into the low-rank component (i.e., principal component) L and the non-low-rank component N. And the low-rank component L is used to describe face features while the non-low-rank component N contains the information associated sparse error.

By low-rank matrix recovery (LR), the training matrix D can be initialized by

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

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

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

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

Once the dictionary L is constructed from training data, we perform recognition of a test image y as

$$y = L\alpha + Nx + z \tag{9}$$

where L is a low-rank dictionary, N is a variation dictionary that associated noises, outlier pixels and occlusions and z is a reconstruction error. Eq.(9) is the model of the proposed sparse low-rank coding. RADL uses the external data to learn the auxiliary dictionary. However SLC, unlike RADL, uses the low-rank and non-low-rank component from training data to construct the dictionary.

3.2. The Solution of SLC

SRC can be applied to classify y, because it assumes that the types of corruption (or occlusion) must be known and present in training dataset D. It is obvious that this assumption might not be practical in real-world's illumination and corruption

Table 1. Experimental results on the CMU Multi-PIE database (%)

Train. Number	3	4	5	6
LR	73.58 ±1.52	79.45 ±1.64	83.72 ±1.63	87.34 ±0.81
LRSI	76.69 ±1.06	81.96 ±1.13	85.63 ±1.52	88.09 ±0.88
SRC	67.58 ±1.53	81.50 ±1.74	88.47 ±0.89	91.67 ±0.60
RSC	87.03 ±0.98	90.22 ±0.93	91.27 ±0.52	91.81 ±0.54
RADL	84.85 ±1.39	88.47 ±0.99	89.01 ±0.86	90.18 ±0.56
SLC	**88.27 ±0.86**	**90.82 ±0.81**	**92.59 ±0.64**	**93.65 ±0.65**

Table 2. Experimental results on the Extended Yale B database (%)

Train. Number	3	4	5	6
LR	51.21 ±1.77	57.42 ±1.92	62.25 ±1.08	66.94 ±0.66
LRSI	52.63 ±1.82	59.15 ±1.99	63.66 ±1.45	68.13 ±0.53
SRC	41.37 ±1.27	56.88 ±1.67	66.41 ±1.13	73.37 ±0.97
RSC	65.63 ±1.67	74.35 ±1.74	79.40 ±0.79	83.44 ±1.55
RADL	65.23 ±1.67	73.06 ±2.29	76.66 ±1.36	77.96 ±2.48
SLC	**65.18 ±3.15**	**74.30 ±2.32**	**80.17 ±1.18**	**85.89 ±1.20**

Table 3. Experimental results on the CAS-PEAL database (%)

Train. Number	3	4	5	6
LR	73.08 ±8.19	78.85 ±6.13	87.35 ±5.41	87.40 ±6.27
LRSI	74.30 ±8.29	80.91 ±6.80	88.79 ±5.26	89.61 ±4.70
SRC	62.00 ±7.19	75.97 ±6.01	87.18 ±5.66	89.94 ±4.94
RSC	70.70 ±4.02	78.01 ±2.79	90.52 ±1.17	92.73 ±1.50
RADL	68.19 ±3.59	76.91 ±3.72	90.76 ±1.27	92.92 ±1.58
SLC	**78.37 ±4.17**	**85.17 ±4.75**	**92.37 ±1.82**	**93.86 ±1.37**

scenarios. To address this issue, we use the residual function to minimize the influence of z. In the theory of robust M-estimators [19], standard residual functions include Huber, Cauchy, and the Welsch functions. The residual function we applied is defined as

$$\begin{aligned}\rho(z) &= \sum_{k=1}^{n} (z_k) \\ \rho(z_k) &= -\frac{1}{2\eta}(\ln(1+\exp(-\eta z_k^2+\eta\sigma)) \\ &\quad -\ln(1+\exp\eta\sigma))\end{aligned} \tag{10}$$

where n is the total number of z_k, z_k is the kth entry of $z = y - L\alpha - Nx$, and η and σ are the parameters in the residual function. We use this residual function because it has shown promising results in recent literatures of robust face recognition [16], [20]. In order to make the representation error z as small as possible, it is not necessary to put any constraint on x. Thus, we consider the following minimization problem instead of Eq.(5):

$$\min_{\alpha} \|\alpha\|_1 + \lambda\rho(y - L\alpha - Nx) \tag{11}$$

where λ is a constant for a compromise.

Then we solve the optimization problem by the technique of variable substitution and the chain rule for calculating the derivatives. From the derivation [17], Eq.(11) can be calculated by repeatedly solving

$$\min_{\alpha} \|\alpha\|_1 + \lambda\|W(y - L\alpha - Nx)\|_2 \tag{12}$$

where W is a weighting matrix obtained by derivation [17] and expressed as

$$\begin{aligned}W &= diag(w(z_1), w(z_2), ..., w(z_d))^{\frac{1}{2}} \\ w(z_k) &= \frac{\exp(-\eta z_k^2+\eta\sigma)}{1+\exp(-\eta z_k^2+\eta\sigma)}\end{aligned} \tag{13}$$

where η and σ are the parameters in the residual function. With W is fixed, we can apply existing techniques such as Homotopy, Iterative Shrinkage-Thresholding, or Alternating Direction Method for obtaining the optimal solution. Finally, our recognition rule is based on the minimum class-wise reconstruction residual.

Let Ci be a class-label matrix of the training data D for class i, its element $C_i(k,k) = 1$ if the kth column of D originates from class i and all other elements of Ci are zero. The class-wise reconstruction residual is defined by

$$r_i(y) = \|W(z - z_i)\|_2 = \|W(L(I - C_i)\alpha)\|_2 \tag{14}$$

where W is a weighting matrix, L is a low-rank dictionary and I is an identity matrix. The testing images y is classified into the class that produces the minimum residual $r_i(y)$.

4. EXPERIMENT RESULTS

In this section, we choose the CMU Multi-PIE [21], Extended Yale B [22], and CAS-PEAL [23] face databases to compare the performance of our method with LR [8], LRSI [15], SRC [11], RSC [16], RADL [17] and SLC in different experimental circumstances. Thereinto, RADL applies variation data instead of external data to obtains the auxiliary dictionary with the same setting as [17]. All images are cropped with size 32 × 32 and all experiments are repeated 10 times. Some cropped images from three databases are shown in the left part of Fig.1.

4.1. Experiments on Appearance

We validate the performance of SLC for face recognition with variations such as posture and expression changes but without illumination and corruption changes. Thus, we carry out this experiment on the CMU Multi-PIE face database and randomly choose 3-6 images per individual as the training set and the remaining images are for the testing set. The average recognition rates and the standard deviations are shown in Table 1. Obviously, SLC consistently outperforms LR, LRSI, SRC, RSC and RADL. Because the low-rank component from the training dataset constructs the dictionary in SLC and it contains important facial feature. Accordingly, this experiment shows that SLC's performance is better than that of the others.

4.2. Experiments on Illumination Change

Illumination has important influence on face recognition. In the first test, we carry out this experiment on the Extended

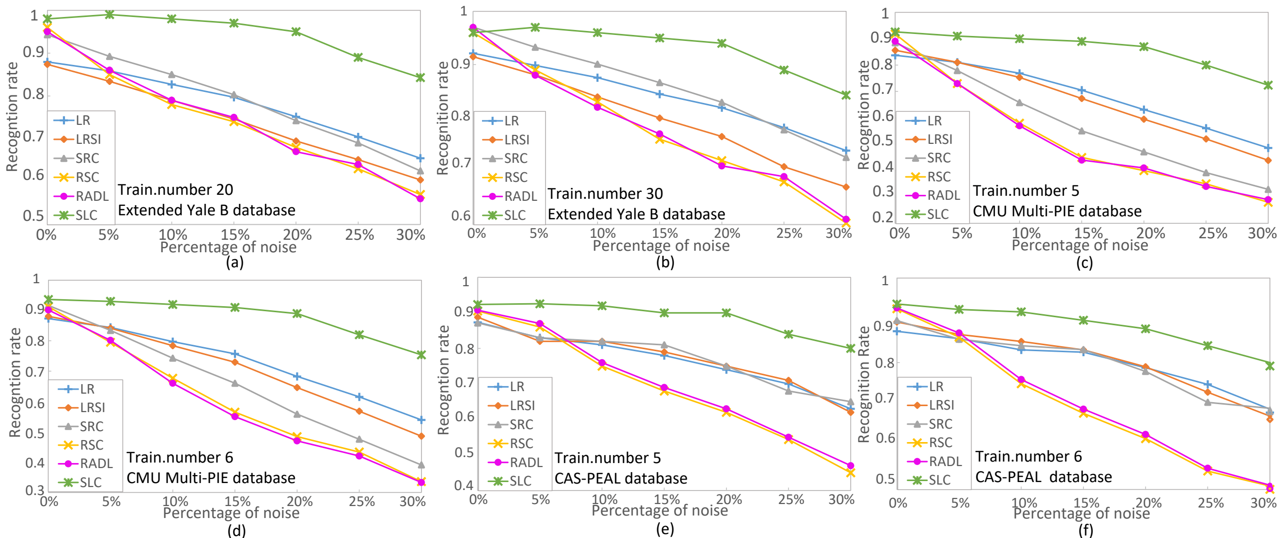

Fig. 2. Experimental results on corrupted training images.

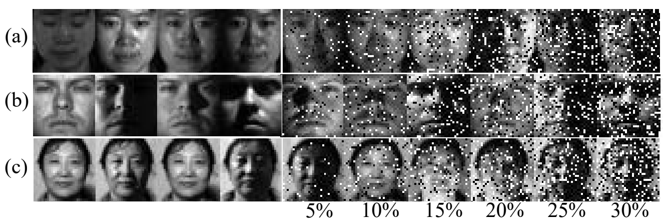

Fig. 1. Some cropped images and some training samples corrupted by salt-and-pepper noise from the (a) CMU Multi-PIE, (b) Extended Yale B and (c) CAS-PEAL database.

Yale B face database. We randomly select 3 - 6 images per individual as the training set and the rest images as the testing set. The average recognition rates and the standard deviations are shown in Table 2. Thanks to its low-rank dictionry, the result of SLC is more approximate to the practical result than that of the other methods. It shows that SLC is better than the other methods for face recognition with illumination, when the number of training images is more than 4.

In the second test, we utilize 178 subjects from the Lighting category of CAS-PEAL database. Each subject has 9 image and is captured under 9 kinds of variations in illumination. We randomly select 3 - 6 images per individual on the CAS-PEAL database as the training set and the rest images as the testing images. The averaged recognition rates and standard deviations are showed in Table 3. We can see the maximum accuracy of the proposed SLC (78.37%) over RADL (68.19%) reaches about 10.18% at 3 training samples. Generally speaking, SLC achieves better than other methods for face recognition with illumination.

4.3. Experiments on Corruption Change

Face recognition for images corrupted by tense noise is a challenge task. In the experiments, to test the corruption robustness of the proposed method, we use the CMU Multi-PIE, Extended Yale B and CAS-PEAL databases and all training samples are corrupted by different level noise. In the right part of Fig.1, from top to bottom, training images are from the CMU Multi-PIE, Extended Yale B and CAS-PEAL 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 the Extended Yale B database, 5 and 6 images per individual from the CMU Multi-PIE and CAS-PEAL databases as the training set and the rest as the testing set, respectively. Fig.2 plots the average results of different training set. From Fig.2, it is noteworthy that SLC consistently has better performance than the others. Because of the low-rank component from the training images, the dictionary contains more feature information instead of the other information associated noise. Hence, the superiority of SLC is obvious with the gradually increase of salt-and-pepper noise. It shows that SLC has a good performance for face recognition with corruption.

5. CONCLUSION

We proposed a sparse low-rank component coding (SLC) for face recognition with illumination and corruption. In this method, we employed LR on the training dataset to construct dictionary. And they represented the effective feature and the information with sparse error respectively. We obtained the solution of the proposed SLC and minimized class-wise reconstruction residual to recognize the testing image. In this way, SLC achieved better classification performance and overcame the problem of training dataset with illumination and corruption. The experiments demonstrated that the proposed SLC method outperforms the other SRC based methods for face recognition with illumination and corruption.

6. REFERENCES

[1] S. Tan, X. Sun, W. Chan, L. Qu, and L. Shao, "Robust face recognition with kernelized locality-sensitive group sparsity representation," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4661–4668, 2017.

[2] A. Majumdar, R. Singh, and M. Vatsa, "Face verification via class sparsity based supervised encoding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1273–1280, 2017.

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

[4] P. N. Belhumeur, J. P. Hespanha, and D. 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.

[5] 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.

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

[7] 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.

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

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

[10] Q. Ke and T. Kanade, "Robust l/sub 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.

[11] 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.

[12] 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.

[13] 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.

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

[15] C. Chen, C. Wei, and Y. F. Wang, "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.

[16] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Robust sparse coding for face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2011, pp. 625–632.

[17] C. Wei and Y. F. Wang, "Undersampled face recognition via robust auxiliary dictionary learning," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1722–1734, 2015.

[18] 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*. IEEE, 2013, pp. 676–683.

[19] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 57–91, 1996.

[20] M. Yang, L. Zhang, J. Yang, and D. Zhang, "Regularized robust coding for face recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 5, pp. 1753–1766, 2013.

[21] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.

[22] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[23] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The cas-peal large-scale chinese face database and baseline evaluations," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 38, no. 1, pp. 149–161, 2008.