Structured Regularized Robust Coding for Face Recognition

Meng Yang(✉), Tiancheng Song, Feng Liu, and Linlin Shen

College of Computer Science and Software Engineering,
Shenzhen University, Shenzhen, China
yang.meng@szu.edu.cn

Abstract. The sparse representation based classifier (SRC) has been successfully applied to robust face recognition (FR) with various variations. To achieve much stronger robustness to facial occlusion, recently regularized robust coding (RRC) was proposed by designing a new robust representation residual term. Although RRC has achieved the leading performance, it ignores the structured information (i.e., spatial consistence) embedded in the occluded pixels. In this paper, we proposed a novel structured regularized robust coding (SRRC) framework, in which the spatial consistence of occluded pixels was exploited by pixel weight learning (PWL) model. Efficient algorithms were also proposed to fastly learn the pixel's weight and accurately recover the occluded area. The experiments on face recognition in several representative datasets clearly show the advantage of the proposed SRRC in accuracy and efficiency.

Keywords: Structure regularized · Robust coding · Face recognition

1 Introduction

Face recognition (FR) has been extensively studied in the past two decades [5], and many representative methods, such as Eigenfaces [6], Fisherfaces [6], LBP [7], have been proposed. In order to deal with facial occlusion, Eigenimages [8-9], probabilistic local approaches [10] and Markov random fields [19] were proposed for FR with occlusion. Although much progress have been made, robust FR to occlusion/disguise is still a challenging issue due to the variations of occlusion such as different categories of disguises, and the unknown intensity of occluded pixels.

Recently, sparse coding [1] and deep learning [17][18] have been widely applied to face recognition. Although deep learning has shown very promising accuracies, it still has some limitations, such as requirements of large amounts of training samples and super computational machines, and lacks of strong theoretical analysis and specific model for face recognition with various occlusions.

A successful work applying sparse coding to robust face recognition is sparse representation based classifier (SRC) [1], which was proposed for robust face recognition, producing very promising performance in FR with occlusion. By coding a query image $\boldsymbol{y}$ as a sparse linear combination of all the training samples via Eq. (1), SRC classifies $\boldsymbol{y}$ by searching for the class that produces the minimal reconstruction error.

H. Zha et al. (Eds.): CCCV 2015, Part II, CCIS 547, pp. 80–89, 2015.
DOI: 10.1007/978-3-662-48570-5_9

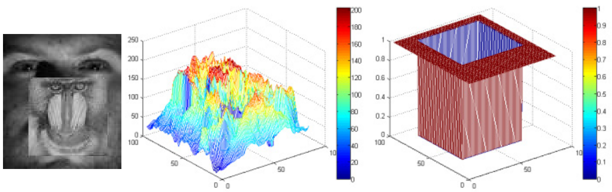
$$\min_{\boldsymbol{\alpha}} \left\| \boldsymbol{y} - \boldsymbol{X\alpha} \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha} \right\|_1 \tag{1}$$

where $\|.\|_1$ is the sparse l_1-norm and each column vector in $\boldsymbol{X}$ is a training sample. In order to make SRC robust to facial occlusion, an identity matrix $\boldsymbol{I}$ was introduced as a dictionary to code the outlier pixels (e.g., occluded pixels):

$$\min_{\boldsymbol{\alpha},\boldsymbol{e}} \left\| \boldsymbol{y} - \boldsymbol{X\alpha} - \boldsymbol{Ie} \right\|_2^2 + \lambda \left\| \boldsymbol{\alpha} \right\|_1 + \lambda \left\| \boldsymbol{e} \right\|_1 \tag{2}$$

By solving Eq. (2), SRC shows good robustness to face occlusions such as block occlusion and disguise. A theoretical support for the success of SRC may be that only a small part of pixels are occluded in most cases (So it is reasonable to require the representation residual $\boldsymbol{e}$ sparse).

It is easy to see in Eq. (2) that the representation residual, i.e., $\boldsymbol{e}$, is regularized by l_1-norm, which may not be optimal when the representation residuals do not follow a Laplacian distribution. Following SRC, He *et al.* [11] proposed a correntropy-based sparse representation (CESR) for robust face recognition, which introduced a Gaussian kernel-based fidelity term to regularize the coding residuals; and Gabor feature was also introduced in the framework of SRC to enhance its discrimination [12]. In order to deal with more general facial occlusion, Yang *et al.* [4] proposed a regularized robust coding (RRC) model by designing a robust representation term, which has shown the state-of-art performance in robust face recognition and attracted much attention in the field.



(a) (b) (c)

Fig. 1. The structured information of occluded image pixels. (a) a face image; (b) pixels' values; (c) pixels' occlusion patterns. It is easy to see occluded pixels' patterns but not the occluded pixels' values are spatial consistent.

Although RRC [4], CESR [11] and Gabor-SRC [12] have achieved leading performance in robust face recognition, all of them measure each pixel's representation residual independently with ignoring the structured information (i.e., spatial consistence) embedded in the 2D image space. In practical face recognition, most of occluded pixels are not independent but spatially consistent (e.g., illumination, expression, block occlusion, facial disguise). Here we should note that the spatial consistence is embedded in occluded pixels but not the occluded pixels' values. Fig. 1 gives an example to show the spatial consistence of image pixels and image pixels' values.

In this paper, we use a weight to indicate whether a pixel is occluded, then the structured information could be easily exploited in the pixel weight learning (PWL) without considering the difference among pixels' values. With the proposed PWL model, the structured information of image pixels could be effectively exploited and a novel framework of structured regularized robust coding (SRRC) was presented for robust face recognition. We also present efficient algorithms to solve PWL model. We evaluate the effectiveness of SRRC on several benchmark datasets, such as CMU Multi-PIE [21] and a joint face database of AR [13] and CAP-Peal [14]. The experiments on these datasets clearly show the advantage of SRRC in accuracy and effectiveness of robust face recognition.

The rest of this paper is organized as follows. Section 2 briefly reviews the related regularized robust coding model. Section 3 presents the proposed structured regularized robust coding framework. Section 4 conducts the experiments, and Section 5 concludes the paper.

2 Brief Review of Related Work

In order increase the robustness of SRC to various outliers, Yang *et al* [4] proposed a regularized robust coding (RRC) model, which was efficiently solved by using an iterative reweighted regularized coding algorithm. In each iteration RRC changes to

$$\min_{\boldsymbol{\alpha}} \left\| \operatorname{diag}(\boldsymbol{w})^{1/2} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\alpha}) \right\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_{l_p} \tag{3}$$

where l_p-norm on $\boldsymbol{\alpha}$ could be l_1-norn or l_2-norm ([2] indicated that the l_2-norm regularized coding could achieve similar accuracy to l_1-norm but with a faster speed), and diag($\boldsymbol{w}$) is a diagonal matrix with the weight vector $\boldsymbol{w}$ as its diagonal vector. Here the element of $\boldsymbol{w}$ is computed as

$$w_j = 1/1 + \exp\left(\mu e_j^2 - \mu\delta\right) \tag{4}$$

where $e_j = y_j - \boldsymbol{r}_j\boldsymbol{\alpha}$, $\boldsymbol{r}_j$ is the j-th row vector of $\boldsymbol{X}$, and μ and δ are two automatically updated scalar parameters in the weight function [4]. Here w_j indicates the importance of the j-th element of $\boldsymbol{y}$ to the coding of $\boldsymbol{y}$. We can observe that the outlier pixels will have small weights to reduce their effects on the coding $\boldsymbol{y}$ on $\boldsymbol{X}$ since they have big residuals.

RRC could be solved by alternatively updating the weight vector $\boldsymbol{w}$ and the coding vector $\boldsymbol{\alpha}$. When the final coding vector $\boldsymbol{\alpha}$ is achieved, RRC conducts the classification via

$$\operatorname{identity}(\boldsymbol{y}) = \arg\min_i \left\| \operatorname{diag}(\boldsymbol{w})^{1/2} (\boldsymbol{y} - \boldsymbol{X}_i\boldsymbol{\alpha}_i) \right\|_2^2 \tag{5}$$

where $\boldsymbol{X}_i$ the training samples of class i, $\boldsymbol{\alpha}=[\boldsymbol{\alpha}_1; \boldsymbol{\alpha}_2;\ldots; \boldsymbol{\alpha}_c]$, and $\boldsymbol{\alpha}_i$ is the coefficient vector associated with i-th disguise pattern.

3 Structured Regularized Robust Coding

Structured information (e.g., spatial consistence) is embedded in the pixels themselves but not the pixels' values. As shown in Fig.1, the nearby occluded pixels may have quite different values but their occluded patterns are same (e.g., their weight values, such as w in RRC, are same). Thus the proposed structured regularized robust coding (SRRC) could be represented as

$$\arg\min_{\boldsymbol{\alpha},\boldsymbol{w}} \left\{ \sum_{i=1}^{n} \rho_{\theta}\left(y_i - \boldsymbol{r}_i\boldsymbol{\alpha}\right) + \lambda \left\|\boldsymbol{\alpha}\right\|_{l_p} \right\} \text{ s.t. } \boldsymbol{w} \text{ is structured regularized} \qquad (6)$$

where $\boldsymbol{w}$ indicates weight values of image pixels, and the representation term is a robust fidelity term like RRC [4]. Here $\boldsymbol{w}$ could be reshaped to a weight map with the same size as face imges,

Similar to RRC, SRRC is solved by iteratively updating the coding vector $\boldsymbol{\alpha}$ and the weight value of each pixel. With known weight matrix $\boldsymbol{w}$, SRRC changes to the coding model of Eq. (3), which could be efficiently solved.

When the coding vector $\boldsymbol{\alpha}$ is known, the robust representation term, i.e., $\rho_{\theta}\left(y_i - \boldsymbol{r}_i\boldsymbol{\alpha}\right)$, could be represented by the approximation of its Taylor expansion (Please refer to the detailed Taylor expansion in [4]). So SRRC model changes to

$$\arg\min_{\boldsymbol{w}} \left\{ \left\|\boldsymbol{w} - \boldsymbol{w}^0\right\|_{l_p} \right\} \text{ s.t. } \boldsymbol{w} \text{ is structured regularized} \qquad (7)$$

where $\boldsymbol{w}^0 = \left[w_1^0 \quad \cdots \quad w_j^0 \quad \cdots \right]$, $w_j^0 = 1 \big/ 1 + \exp\left(\mu e_j^2 - \mu\delta\right)$ is the estimated weight based on the Taylor expansion of $\rho_{\theta}\left(y_i - \boldsymbol{r}_i\boldsymbol{\alpha}\right)$, w_j^0 is the j-th element value of $\boldsymbol{w}^0$.

It can be easily seen that SRRC will degenerate to RRC if there is no structured regularization on $\boldsymbol{w}$. In order to make the robust representation model exploit the structured information, we present a pixel weight learning (PWL) model of Eq.(7) to introduce the structured information.

3.1 Pixel Weight Learning (PWL)

The structured information could be designed in many ways. In this paper, we only use the local consistence of image pixels as the structured information. Then the pixel weight learning (PWL) model could be rewritten as

$$\min_{\boldsymbol{w}} \left\|\boldsymbol{w} - \boldsymbol{w}^0\right\|_{l_p} + \kappa \sum_i \sum_{j \in Ni} \left\|w_i - w_j\right\|_{l_p} \qquad (8)$$

where k is a parameter to control the structured regularization, l_p-norm indicates the l_1-norm and l_2-norm when p=1 and p=2, respectively. For each pixel i, j is a neighboring pixels, and Ni is the set of neighboring pixels of pixel i. With the final term ensures the neighboring pixels have similar weight values. Here the neighboring size could be set by the users. A bigger neighboring region will introduce more global consistence. In this paper we use 4 neighborhoods for a pixel.

Different l_p-norm regularization will have different physical meanings. In order to make the learned $\boldsymbol{w}$ similar to $\boldsymbol{w}^0$,we use l_2-norm for the first term. When w_{pq}-$w_{pq'}$ is regularized by l_2-norm, Eq.(8) requires the weight map should be smooth, while l_1-norm regularized version could tolerate some sparse and sharp variance. Here we only consider the case that w_{pq}-$w_{pq'}$ is regularized by l_2-norm since we want the weight values be spatially consistent in general.

Thus the PWL model could be represented as

$$\min_{\boldsymbol{w}} \left\| \boldsymbol{w} - \boldsymbol{w}^0 \right\|_{l_p} + \kappa \sum\nolimits_i \sum\nolimits_{j \in Ni} \left\| w_i - w_j \right\|_2^2 \tag{9}$$

3.2 Solving Algorithm of PWL

In order to efficiently solve Eq.(9), we rewrite the final term of Eq.(9) as

$$\sum\nolimits_i \sum\nolimits_{j \in Ni} \left\| w_i - w_j \right\|_2^2 = \sum\nolimits_{j \in Ni} \boldsymbol{A}_j \boldsymbol{w} \tag{10}$$

where $\boldsymbol{A}_j$ is an indication matrix of the j-th neighboring pixel with all diagonal elements as 1s. For each row of $\boldsymbol{A}_j$ (i.e., each pixel in the image), the value of j-th neighboring pixel is set as -1, with all the elements as 0s. So $\boldsymbol{A}_j \boldsymbol{w}$ is a vector with each element as the difference of a pixel and its j-th neighboring pixel.

Denote $\boldsymbol{v}$=$\boldsymbol{w}$-$\boldsymbol{w}^0$, by replacing $\boldsymbol{w}$ as $\boldsymbol{v}$+$\boldsymbol{w}^0$ we rewrite the PWL model as

$$\min_{\boldsymbol{v}} \left\| \boldsymbol{v} \right\|_F^2 + \kappa \sum\nolimits_j \left\| \boldsymbol{A}_j \boldsymbol{v} + \boldsymbol{A}_j \boldsymbol{w}^0 \right\|_2^2 \tag{11}$$

In this case, we could derive an analytic solution, and the weight matrix solution could be presented as

$$\boldsymbol{v} = -\left(\kappa \sum\nolimits_j \boldsymbol{A}_j^T \boldsymbol{A}_j + I \right)^{-1} \sum\nolimits_j \boldsymbol{A}_j^T \boldsymbol{A}_j \boldsymbol{w}^0 \tag{12}$$

Based on $\boldsymbol{v}$=$\boldsymbol{w}$-$\boldsymbol{w}^0$, we could further derive

$$\boldsymbol{w} = \boldsymbol{P}\boldsymbol{w}^0 \tag{13}$$

where $\boldsymbol{P}=\left(\kappa \sum\nolimits_j \boldsymbol{A}_j^T \boldsymbol{A}_j + I \right)^{-1}$. Since the incidence matrix is predefined, the projection matrix $\boldsymbol{P}$ could be pre-computed and in testing time, only a projection operation with a low computation complexity is needed.

3.3 The Whole Algorithm of SRRC

Based on PWL, the whole algorithm of SRRC is summarized in Table 1.

Table 1. Algorithm of SRRC.

Solving algorithm of SRRC
1. Initialize $\boldsymbol{\alpha}$
2. Compute residual $\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X\alpha}$.
3. Estimate weights $\boldsymbol{w}$ as via Eq.(4)
4. Weight updating $\boldsymbol{w}$ via PWL of Eq.(9)
5. Solve $\boldsymbol{\alpha}$ via the weighted regularized robust coding , i.e., Eq.(3)
6. Output $\boldsymbol{\alpha}$ until the condition of convergence is met, or the maximal number of iterations is reached.

After several iteration, we could get the final weight vector $\boldsymbol{w}$ and coding vector $\boldsymbol{\alpha}$, and then conduct face recognition via Eq.(5).

4 Experiments

We perform experiments on several benchmark datasets, such as CMU Multi-PIE [21], and a joint database [13][14] to demonstrate the performance of SRRC. In Section 4.1, we test SRRC on face recognition with illumination and expression variations; in Section 4.2 we compare the accuracies and running time on a joint face dataset. Here the joint database was constructed by using AR database (100 persons, 2599 images) [13] and a subset of CAS-Peal (101 persons and 843 images) [14]. For the experiments of face recognition without occlusion, we estimate the weight values of original face image and then use PCA to reduce the feature dimensionality like that in RRC [4].

In all experiments κ of PWL is set as 0.05 in face recognition without occlusion and 0.2 in face recognition with occlusion, respectively. The λ is set as the suggested value in RRC. The competing methods include the latest approaches, such as LLC [20], SRC [1], Gabor-SRC [12], CESR [11], RRC_L1 [4] and RRC_L2 [4]. Similar to RRC, SRRC_L1 and SRRC_L2 represent SRRC using l_1-norm and l_2-norm on $\boldsymbol{\alpha}$, respectively.

4.1 Face Recognition Without Occlusion

AR Database: As in [4], a subset (with only illumination and expression changes) that contains 50 male and 50 female subjects was chosen from the AR database [13] in this experiment. For each subject, the seven images from Session 1 were used for training, with other seven images from Session 2 for testing. The images were cropped to 60×43. The FR rates by the competing methods are listed in Table 2. We can see that SRRC could improve the performance of the second best method, RRC, in most cases. Especially, SRRC_L1 achieves the highest accuracy with visible improvement (e.g., 1.2% with 120-d feature).

Table 2. Face recognition rates on the AR database.

Dimension	54	120	300
NN	68.0%	70.1%	71.3%
SVM	69.4%	74.5%	75.4%
SRC [1]	83.3%	90.1%	93.3%
LLC [20]	80.7%	87.4%	89.0%
RRC_L_2	84.3%	94.3%	95.3%
SRRC_L2	84.4%	94.0%	95.9%
RRC_L_1	87.6%	94.7%	96.3%
SRRC_L1	**88.4%**	**95.9%**	**97.0%**

Multi PIE Database: The CMU Multi-PIE database [21] contains images of 337 subjects captured in four sessions with simultaneous variations in pose, expression, and illumination. Among these 337 subjects, all the 249 subjects in Session 1 were used for training. To make the FR more challenging, two subsets with both illumination and expression variations in Sessions 1 and 3, were used for testing. For the training set, as in [4] and [1], we used the 7 frontal images with extreme illuminations {0, 1, 7, 13, 14, 16, and 18} and neutral expression. For the testing set, 4 typical frontal images with illuminations {0, 2, 7, 13} and smile expressions (smile in Sessions 1 and 3) were used. Here we used the Eigenface with dimensionality 300 as the face feature for sparse coding. Table 3 lists the recognition rates in four testing sets by the competing methods.

From Table 3, we can see that SRRC_L1 achieves the best performance in all tests,. Compared to the baseline method, SRC, SRRC_L1 has 4.4% improvement in Smi-S1 and 16.3% in Smi-S3, respectively. Although the improvement of SRRC over RRC is not big, the introduction of structured information could still bring some benefits.

Table 3. Face recognition rates on Multi-PIE database. ('Smi-S1': set with smile in Session 1; 'Smi-S3': set with smile in Session 3).

	Smi-S1	Smi-S3
NN	88.7%	47.3%
SVM	88.9%	46.3%
SRC [1]	93.7%	60.3%
LLC [20]	95.6%	62.5%
RRC_L_2	95.9%	67.3%
SRRC_L2	96.2%	67.8%
RRC_L_1	97.8%	76.0%
SRRC_L1	**98.1%**	**76.6%**

4.2 Face Recognition on a Joint Face Database

In the test, we conduct FR with more complex disguises (e.g., sunglasses, scarf and hat) with variations of illumination and longer data acquisition interval. 340 images of the first 85 subjects (4 natural and non-occluded images with different illuminations in Session 1) in AR database and 263 images of the first 80 subjects (the non-occluded images) in CAS-Peal are used as the training sets. And 510 face images with sunglass and lighting variations, 510 face images with scarf and lighting variations, and 240 face images with hat and lighting variations are used as the testing dataset. Some samples are shown in Fig. 2.

Fig. 2. The training and testing samples in the joint database.

Table 4. Recognition rates by competing methods on the joint database of AR and CAS-Peal with complex disguise occlusion.

Method	Sunglass	Scarf	Hat
SRC [1]	73.9%	24.9%	26.3%
GSRC [12]	52.4%	66.1%	34.2%
CESR[11]	80.2%	11.0%	26.7%
RRC_L_2	83.5%	75.3%	60.4%
SRRC_L2	87.4%	81.6%	71.3%
RRC_L_1	90.2%	77.3%	67.1%
SRRC_L1	**93.1%**	**83.3%**	**78.3%**

Table 4 lists the results of face recognition on the joint database by competing methods. Clearly, the SRRC methods achieve much better results than SRC, GSRC, CESR and RRC in most cases. RRC achieves the second best performance. SRRC_L2 outperforms RRC_L2 by 3.9%, 6.3% and 10.9% in face recognitions with sunglass, scarf and hat, respectively. SRRC_L1 outperforms RRC_L1 by 2.9%, 6.0%, and 11.2% in face recognitions with sunglass, scarf and hat, respectively;

Apart from recognition rate, computational expense is also an important issue for practical FR systems. In this section, the running time of the baseline method, SRC, and some competing methods which show not bad performance in all cases, including GSRC, RRC_L2, RRC_L1, and SRRC, is evaluated using the FR experiments on the joint face database. The programming environment is Matlab version R2013a. The desktop used is equipped with a 3.5 GHz CPU and 16G RAM. All the methods are implemented using the codes provided by the authors. For SRC, we use a fast l_1-minimization solver, ALM [15], to implement the sparse coding step.

Table 5 lists the average computational expense of different methods. We can observe that both SRRC_L2 and RRC_L2 have the least running time, followed by GSRC and SRC. Although the proposed SRRC has similar computation time to RRC, SRRC could achieve visibly better performance than RRC. Especially, SRRC_L1 has much better performance than RRC_L1 but with less running time.

Table 5. Average runnning time on the joint database with three facial disguises.

Method	Sunglass	Scarf	Hat
SRC (ALM)	0.610	0.579	0.574
GSRC	0.269	0.265	0.277
RRC_L_2	**0.177**	0.153	**0.171**
SRRC_L2	0.200	0.170	0.194
RRC_L_1	1.58	1.34	1.59
SRRC_L1	1.26	1.10	1.26

5 Conclusion

This paper presented a novel structured regularized robust coding (SRRC) framework and an associated pixel weight learning (PWL) model for robust face recognition. We also propose effective algorithms to solve the pixel weight learning model. One important advantage of SRRC is that the structured information (e.g., spatial consistence) could be exploited by the proposed SRRC with PWL. The proposed SRRC methods were extensively evaluated on FR with various variations, such as illumination, expression, random block occlusion, and real facial disgusie. The experimental results clearly demonstrated that SRRC outperforms previous state-of-the-art methods, such as SRC, CESR, GSRC and RRC.

Acknowledgement. This work is partially supported by the National Natural Science Foundation of China under Grants no. 61402289, 61272050 and 61403257, National Science Foundation of Guangdong Province (2014A030313558 and 2014A030313556), and Shenzhen Scientific Research and Development Funding Program under Grant JCYJ20140509172609171.

Reference

1. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. IEEE Trans. Pattern Analysis and Machine Intelligence **31**(2), 210–227 (2009)
2. Zhang, L., Yang, M., Feng, X.C.: Sparse representation or collaborative representation: which helps face recognition? In: Proc. ICCV (2011)
3. Fidler, S., Skocaj, D., Leonardis, A.: Combining Reconstructive and Discriminative Subspace Methods for Robust Classification and Regression by Subsampling. IEEE Trans. Pattern Analysis and Machine Intelligence **28**(3), 337–350 (2006)
4. Yang, M., Zhang, L., Yang, J., Zhang, D.: Regularized robust coding for face recognition. IEEE Trans. Image Processing **22**(5), 1753–1766 (2013)
5. Zhao, W., Chellppa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Survey **35**(4), 399–458 (2003)
6. Belhumeur, P.N., Hespanha, J.P., Kriengman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Analysis and Machine Intelligence **19**(7), 711–720 (1997)
7. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Analysis and Machine Intelligence **28**(12), 2037–2041 (2006)
8. Leonardis, A., Bischof, H.: Robust recognition using eigenimages. Computer Vision and Image Understanding **78**(1), 99–118 (2000)
9. Chen, S., Shan, T., Lovell, B.C.: Robust face recognition in rotated eigenspaces. In: Proc. Int'l Conf. Image and Vision Computing, New Zealand (2007)
10. Martinez, A.M.: Recognizing Imprecisely localized, partially occluded, and expression variant faces from a single sample per class. IEEE Trans. Pattern Analysis and Machine Intelligence **24**(6), 748–763 (2002)
11. He, R., Zheng, W.S., Hu, B.G.: Maximum correntropy criterion for robust face recognition. IEEE Trans Pattern Analysis and Machine Intelligence **33**(8), 1561–1576 (2011)

12. Yang, M., Zhang, L.: Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part VI. LNCS, vol. 6316, pp. 448–461. Springer, Heidelberg (2010)
13. Martinez, A., Benavente, R.: The AR face database. CVC Tech. Report No. 24 (1998)
14. Gao, W., Cao, B., Shan, S.G., Chen, X.L., Zhou, D.L., Zhang, X.H., Zhao, D.B.: The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. IEEE Trans. on System Man, and Cybernetics (Part A) **38**(1), 149–161 (2008)
15. Yang, A.Y., Ganesh, A., Zhou, Z.H., Sastry, S.S., Ma, Y.: A review of fast l1-minimization algorithms for robust face recognition (2010). arXiv:1007.3753v2
16. Jia, H., Martinez, A.: Support vector machines in face recognition with occlusions. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2009)
17. Sun, Y., Chen, Y., Wang, X., Tang, X.: Deep learning face representation by joint identification-verification. In: Proc. Neural Information Processing Systems (2014)
18. Zhu, Z., Luo, P., Wang, X., Tang, X.: Multi-view perceptron: a deep model for learning face identity and view representations. In: Proc. Neural Information Processing Systems (2014)
19. Zhou, Z., Wagner, A., Mobahi, H., Wight, J., Ma, Y.: Face recognition with contiguous occlusion using markov random fields. In: Proc. IEEE Conf. Computer Vision (2009)
20. Wang, J.J., Yang, J.C., Yu, K., Lvx, F.J., Huangz, T., Gong, Y.H.: Locality-constrained linear coding for image classification. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2010)
21. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. Image and Vision Computing **28**, 807–813 (2010)