Local Variation Joint Representation for Face Recognition with Single Sample per Person

Meng Yang(✉), Tiancheng Song, Shiqi Yu, and Linlin Shen

College of Computer Science and Software Engineering,
Shenzhen University, Shenzhen, China
yangmengpolyu@msn.com

Abstract. Sparse representation based classification (SRC) was originally applied to multiple-training-sample face recognition with promising performance. Recently SRC has been extended to face recognition with single sample per person by using variations extracted from a generic training set as an additional common dictionary. However, the extended SRC ignored to learn a better variation dictionary and to use local region information of face images. To address this issue, we propose a local variation joint representation (LVJR) method, which learns a variation dictionary and does joint and local collaborative representation for a query image. The learned variation dictionary was required to do similar representation for the same-type facial variations, while the joint and local collaborative representation could effectively use local information of face images. Experiments on the large-scale CMU Multi-PIE and AR databases demonstrate that the proposed LVJR method achieves better results compared with the existing solutions to the single sample per person problem.

Keywords: Local variation · Joint representation · Face recognition · Single sample per person

1 Introduction

As one of the most visible applications in computer vision and pattern recognition, face recognition (FR) has been receiving significant attention in the community [17]. In practical FR scenarios such as face identification/verification in uncontrolled or less controlled environment [6, 16], there are many problems which have attracted much attention of researchers. For instance, face recognition with single sample per person is one of the most important FR problems. In the scenarios (e.g., law enforcement, e-passport, driver license, etc), there is only a single training face image per person. This makes the problem of FR particularly hard since very limited information is provided to predict the variations in the query sample. How to achieve high FR performance in the case of single training sample per person (SSPP) is an important and challenging problems in FR.

The performance of FR would be greatly affected by the limited number of training samples per person [26]. First, many discriminant subspace and manifold learning algorithms (e.g., LDA and its variants [15]) cannot be directly applied to FR with SSPP.

H. Zha et al. (Eds.): CCCV 2015, Part II, CCIS 547, pp. 41–50, 2015.
DOI: 10.1007/978-3-662-48570-5_5

Second, sparse representation based classification (SRC) [12], cannot be easily applied to the problem of SSPP, either, since multiple training samples per person are needed to well reconstruct the query face. As reviewed in [26], many specially designed FR methods have been developed. According to the availability of an additional generic training set, the FR methods for SSPP can be divided into two categories: methods without using a generic training set, and methods with generic learning.

The SSPP methods without generic learning often extract robust local features (e.g., gradient orientation [10] and local binary pattern [1]), generate additional virtual training samples (e.g., via singular value decomposition [25], geometric transform and photometric changes [27]), or perform image partitioning (e.g., local patch based LDA [23], self-organizing maps of local patches [22], and multi-manifold learning from local patches [8]). Although these methods have reported improved FR results, they ignored to introduce additional variation information into the single-sample gallery set. Meanwhile, local feature extraction and discriminative learning from local patches can be sensitive to image variations (e.g., extreme illumination and expression), while the new information introduced by virtual training sample generation can be rather limited.

Opposite to the first category of FR with SSPP, methods with generic learning try to borrow new and useful information (e.g., generic intra-class variation) from a generic training set. An intrinsic reason is the fact that face image variations for different subjects share much similarity. Since a generic training set could be easily collected, it has been widely employed in [21, 20, 9] to extract discriminant information for FR with SSPP. For instance, the expression subspace and pose-invariant subspace were learned from a collected generic training set to solve the expression-invariant [21] and pose-invariant [9] FR problems, respectively. Deng *et al.* [3] extended the SRC method to FR with SSPP. The so-called Extended SRC (ESRC) computes the intra-class variation in a generic training set and then uses the generic variation matrix to code the difference between the query and gallery samples. Following ESRC, Zhu *et al.* [16] proposed a block-based method, in which the weight of each block is iteratively updated to reduce occlusion affect on the final coding.

Dictionary learning has been extensively studied in image processing and computer vision [14, 24], etc. However, most of the dictionary learning methods for pattern classification are conducted on the gallery set with multiple samples per class. How to better learn the variation dictionary is still an open question. Recently, Yang *et al.* [13] proposed sparse variation dictionary learning (SVDL) method to learn a variation dictionary and then project the variation dictionary to the space of gallery images.

Although much improvement has been reported, there are several issues remained the generic training based methods for FR with SSPP. First, the variation matrix can be very big and redundant since many subjects in the generic training set are involved. This will increase the computational burden of the final FR algorithm. Second, SVDL required that each subject in the generic training set should include the same number of variations, which may not be available in practical application. Third, ESRC and SVDL both use holistic features which may not effective for facial variation. Although the method proposed by Zhu *et al.* [16] uses local information, the iterative reweighted procedure would need more computation.

To solve the above mentioned problems, we propose to a local variation joint representation (LVJR) method for FR with SSPP. In order to better exploit different types of variation information, we learn a compact variation dictionary with powerful variation representation ability. In addition, a novel joint and local collaborative representation model was also proposed for the representation and classification of a query image. Extensive experiments have been conducted on the large-scale face databases with various variations, including illumination, expression, pose, session, and occlusion, etc. The experimental results show that the proposed LVJR achieves much better performance than state-of-the-art methods for FR with SSPP.

Section 2 presents a brief review of related work. Section 3 gives the proposed LVJR model. Section 4 describes the optimization procedure of LVJR. Section 5 conducts experiments and Section 6 concludes the paper.

2 Brief Review of Related Work

Recently, sparse representation based classification (SRC) [12], has achieved very promising results, which have led to many following works [3, 11]. However, SRC cannot be directly applied to FR with only a single sample per person (SSPP). To address this issue, Deng *et al.* [3] proposed to integrate the intra-class variation matrix extracted from a generic training set to represent the testing sample. Since our work is developed based on ESRC, we give a review of ESRC here.

Denote the intra-class variation matrix of generic training set by $\boldsymbol{V}=[\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_n]$, where $\boldsymbol{V}_i$ is the i^{th}-type variation matrix, and each column of $\boldsymbol{V}_i$ is the difference between a i^{th}-type variation sample and a reference. Let $\boldsymbol{G} = [\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_c]$ and $\boldsymbol{y}$ be the gallery set with a single sample per person and the testing sample, respectively. The procedures of ESRC [3] are described as follows.

1. Sparsely code $\boldsymbol{y}$ on the matrix $[\boldsymbol{G}\ \boldsymbol{V}]$ via l_1-norm minimization :

$$\left[\hat{\boldsymbol{\rho}};\hat{\boldsymbol{\beta}}\right] = \arg\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} \left\|\boldsymbol{y}-[\boldsymbol{G}\,\boldsymbol{V}][\boldsymbol{\rho};\boldsymbol{\beta}]\right\|_2^2 + \lambda\left\|[\boldsymbol{\rho};\boldsymbol{\beta}]\right\|_1 \tag{1}$$

   where λ is a scalar constant, $\boldsymbol{\rho}$ is the coding coefficient associated with $\boldsymbol{G}$, and $\boldsymbol{\beta}$ is the coding coefficient associated with $\boldsymbol{V}$.
2. Classify $\boldsymbol{y}$ via

$$\text{identity}(\boldsymbol{y}) = \arg\min_i \left\|\boldsymbol{y}-\boldsymbol{g}_i\hat{\rho}_i - \boldsymbol{V}\hat{\boldsymbol{\beta}}\right\|_2 \tag{2}$$

   where $\hat{\boldsymbol{\rho}} = [\hat{\rho}_1;\hat{\rho}_2;\cdots;\hat{\rho}_c]$, and $\hat{\rho}_i$ is the coefficient associated with class i.

ESRC has shown interesting results [3], but it doesn't learn a variance dictionary. Recently, Yang *et al.* [13] proposed a variation dictionary learning approach while it needs a strict requirement (e.g., each subject should have the same number of variations) on the generic training set and ignores to use local region information.

3 Local Variation Joint Representation (LVJR)

In this section, we proposed a local variation joint representation (LVJR) method for FR with SSPP. In LVJR, a variation dictionary learning with joint representation method was proposed for constructing a variation dictionary and a joint and local collaborative representation model was proposed for classification.

3.1 Variation Dictionary Learning with Joint Representation

Face images from different subjects have a big inter-class similarity. In FR with SSPP, we also assume that the face images from different subjects would share similar variations. This kind of assumption has been applied to FR [21][9] with improved results.

For a type of variation matrix of a local region, $\boldsymbol{V}$, we want to learn a variation dictionary $\boldsymbol{D}$ so that joint representation of these variations could be conducted on $\boldsymbol{D}$. Here joint representation requires that the coding coefficients of the variations in the same category should be similar. The proposed variation dictionary learning model could be written as

$$\min_{\boldsymbol{D},\boldsymbol{A}} \left\|\boldsymbol{V}-\boldsymbol{D}\boldsymbol{A}\right\|_2^2+\gamma\left\|\boldsymbol{A}\right\|_{2,1} \tag{3}$$

where $\|.\|_{2,1}$ is defined as $\|\boldsymbol{A}\|_{2,1}=\sum_k\|\boldsymbol{a}_k\|_2$, $\boldsymbol{a}_k$ is the k-th row vector of the coefficient matrix $\boldsymbol{A}$, γ are a scalar variable. The mixed-norm $\|.\|_{2,1}$ requires the between-row sparisity by using l_1-norm and regularizes the variables in each row vector via l_2-norm, which could make the variation in the same category (e.g., illumination with certain direction, certain type of expression) have similar coding vectors.

3.2 Joint and Local Collaborative Representation

Based on the learned variation dictionary we could develop a joint and local representation model to effectively exploit the local information. Let $\boldsymbol{y}=[\boldsymbol{y}^1,\boldsymbol{y}^2,\ldots,\boldsymbol{y}^K]$, where $\boldsymbol{y}^k$ is the k-th local region of $\boldsymbol{y}$. Similarly, the variation matrix of a generic training set could also be divided into K local regions, and each local region could learn a variation dictionary, $\boldsymbol{D}^k$.

In the joint and local representation phase, we want the coding coefficients of different local regions should be similar because these local regions come from the same query image. In order to efficiently solve the joint representation, we adopt l_2-norm to regularize the coding coefficients inspired by [7]. The proposed joint and local collaborative representation model could be written as

$$\min_{\boldsymbol{\alpha}^k} \sum_{k=1}^{K}\left(\left\|\boldsymbol{y}^k-\left[\boldsymbol{G}^k\ \boldsymbol{D}^k\right]\boldsymbol{\alpha}^k\right\|_F^2+\lambda\left\|\boldsymbol{\alpha}^k\right\|_2^2+\mu\left\|\boldsymbol{\alpha}^k-\bar{\boldsymbol{\alpha}}\right\|_F^2\right) \tag{4}$$

where $\boldsymbol{\alpha}^k=\left[\boldsymbol{\rho}^k;\boldsymbol{\beta}^k\right]$ is the coding coefficient for k-th local region, $\boldsymbol{\rho}^k$ is the coding sub-coefficient vector associated to the gallery set, $\boldsymbol{G}^k$, and $\boldsymbol{\beta}^k$ is the coding sub-coefficient vector associated to the variation dictionary, $\boldsymbol{D}^k$. Here $\bar{\boldsymbol{\alpha}}$ is the mean vector of all $\boldsymbol{\alpha}^k$.

When we solve Eq.(4), the classification could be conducted via

$$\text{identity} = \arg\min_i \left\{ \sum_{k=1}^{K} \omega_k \left\| \boldsymbol{y}^k - \boldsymbol{g}_i \boldsymbol{\rho}_i^k - \boldsymbol{D}^k \boldsymbol{\beta}^k \right\|_2 \right\} \tag{5}$$

where $\omega_k = \exp\left(-\left\| \boldsymbol{y}^k - \boldsymbol{G}\boldsymbol{\rho}^k - \boldsymbol{D}^k \boldsymbol{\beta}^k \right\|_2^2 \Big/ 2\sigma^2 \right)$, $\boldsymbol{\rho}_i^k$ is the coding sub-coefficient associated to the i-th gallery image, and $\sigma^2 = \sum_{k=1}^{K} \left\| \boldsymbol{y}^k - \boldsymbol{G}\boldsymbol{\rho}^k - \boldsymbol{D}^k \boldsymbol{\beta}^k \right\|_2^2 \Big/ K$.

4 Solving Algorithm of JLVR

4.1 Solving Variation Dictionary Learning

The model of variation dictionary learning with joint representation could be efficiently solved by alternatively updating the dictionary $\boldsymbol{D}$ and coding coefficient $\boldsymbol{A}$.

When the dictionary, $\boldsymbol{D}$, is fixed, Eq.(3) changes to

$$\min_{\boldsymbol{A}} \left\| \boldsymbol{V} - \boldsymbol{D}\boldsymbol{A} \right\|_2^2 + \gamma \left\| \boldsymbol{A} \right\|_{2,1} \tag{6}$$

which could be efficiently solved by the Iterative Projection Method [5]. Denote $\boldsymbol{\Lambda}=\boldsymbol{A}^{(t)}\text{-}(\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{A}^{(t)}\text{-}\boldsymbol{D}^T\boldsymbol{V})/\sigma$, the solution could be written as

$$\boldsymbol{A}^{(t+1)}[k]=\boldsymbol{\Lambda}[k]\cdot\text{Max}(0,1\text{-}\lambda/(2\sigma\|\boldsymbol{\Lambda}[k]\|_2)) \tag{7}$$

where σ is a scalar parameter in [37], Max(.) is a maximal operator, $\boldsymbol{A}^{(t+1)}[k]$ and $\boldsymbol{\Lambda}[k]$ are the k-th row vector of $\boldsymbol{A}^{(t+1)}$ and $\boldsymbol{\Lambda}$ in the t+1 iteration, respectively.

When the coding coefficient, $\boldsymbol{A}$, is fixed, Eq.(3) changes to

$$\min_{\boldsymbol{D}} \left\| \boldsymbol{V} - \boldsymbol{D}\boldsymbol{A} \right\|_2^2 \tag{8}$$

which could be efficient solved atom by atom via the metaface learning [4].

4.2 Solving Joint and Local Collaborative Representation

The proposed joint and local collaborative representation model, Eq.(4), could be efficiently solved. For each local region, the coding coefficient could be derived

$$\boldsymbol{\alpha}^k = \boldsymbol{\alpha}^{k,0} + \tau \boldsymbol{P}^k \bar{\boldsymbol{\alpha}} \tag{9}$$

where $\boldsymbol{\alpha}^{k,0} = \boldsymbol{P}^k \left[\boldsymbol{G}\ \boldsymbol{D}^k \right]^T \boldsymbol{y}^k$, and $\boldsymbol{P}^k = \left(\left[\boldsymbol{G}\ \boldsymbol{D}^k \right]^T \left[\boldsymbol{G}\ \boldsymbol{D}^k \right] + \lambda \boldsymbol{I} \right)^{-1}$.

Based on $\bar{\boldsymbol{\alpha}} = \sum_{k=1}^{K} \boldsymbol{\alpha}^k \Big/ K$ .By summing $\boldsymbol{\alpha}^k$, we could get

$$K\bar{\boldsymbol{\alpha}} = \sum_{k=1}^{K} \boldsymbol{\alpha}^k = \sum_{k=1}^{K} \boldsymbol{\alpha}^{k,0} + \tau \sum_{k=1}^{K} \boldsymbol{P}^k \bar{\boldsymbol{\alpha}} \tag{10}$$

And then we could derive

$$\bar{\boldsymbol{\alpha}} == \left(I - \tau \Big/ K \sum_{k=1}^{K} \boldsymbol{P}^k \right)^{-1} \sum_{k=1}^{K} \boldsymbol{\alpha}^{k,0} \Big/ K \tag{11}$$

Based on Eq.(11) and Eq.(9), we could get an analytical solution of $\boldsymbol{\alpha}^k$.

5 Experiments

In this section, we perform FR with SSPP on benchmark face databases, including large-scale CMU Multiple PIE [19] and AR [2], to demonstrate the performance of SVDL. We first discuss the parameter setting in Section 5.1; in Section 5.2 we test the robustness of LVJR to various variations on CMU Multi-PIE; in Section 5.3, we evaluate LVJR on the AR database.

We compare the proposed LVJR with several state-of-the-art methods on FR with STSPP, including ESRC [3], ESRC-KSVD (the variation dictionary is learned via KSVD[24]), Adaptive Generic Learning (AGL) for Fisherfaces [18], and Discriminative Multi-Manifold Analysis (DMMA) [8], Sparse Variation Dictionary Learning (SVDL) [13], and some baseline classifiers such as SRC [12], Nearest Subspace (NS) and Support Vector machine (SVM). It should be noted that NS is reduced to Nearest Neighbor (NN) in the case of FR with STSPP. Among these methods, NN, SVM, SRC and DMMA do not use a generic training set, while ESRC, AGL, SVDL and JVJR need a generic training set.

5.1 Parameter Setting

There are three regularization parameters, γ, λ and μ, in LVJR. γ regularizes the variation dictionary learning, while λ and μ controls the l_2-norm regularization and similarity of coding coefficients in the joint and local collaborative representation. If no specific instruction, we fix $\gamma = \lambda = \mu = 0.005$, and initialize dictionary atom number as 400.

5.2 Evaluation to Various Variation on CMU-PIE Dataset

We test the robustness of all the competing methods by using the large-scale CMU Multi-PIE database [19], whose images were captured in four sessions with simultaneous variations of pose, expression, and illumination. For each subject in each session, there are 20 illuminations with indices from 0 to 19 per pose per expression. Among the 249 subjects in Session 1, the first 100 subjects were used for gallery training, with the remaining subjects for generic training. For the gallery set, we used the single frontal image with illumination 7 and neutral expression. The image is cropped to 100×82. Here LVJR divided a face image into 2×2 local regions.

1) Illumination Variation: as [13], we use all the frontal face images with neutral expression in Sessions 2, 3, and 4 for testing. The generic training set is composed of all the frontal face images with neutral expression in Session 1. Table 1 lists the recognition rates in the three sessions by the competing methods.

From Table 1, we can see that LVJR achieves the best results in all cases, and SVDL performs the second best, followed by ESRC. That shows a learned variation dictionary could generate a better performance. SRC does not get good result since the single training sample of each class has very low representation ability. DMMA is the best method without using generic training set; nonetheless, its recognition rates are not high since the illumination variation cannot be well learned from the gallery set via multi-manifold learning.

Table 1. Face recognition rates on Multi-PIE database with illumination variations.

Session	Session 2	Session 3	Session 4
NN	45.3%	40.2%	43.7%
SVM	45.3%	40.2%	43.7%
SRC [12]	52.4%	46.7%	49.5%
DMMA [8]	63.2%	55.4%	60.4%
AGL [18]	84.9%	79.4%	78.3%
ESRC [3]	92.6%	84.9%	86.7%
ESRC-KSVD	92.7%	84.9%	86.7%
SVDL	94.8%	87.7%	91.0%
LVJR	**96.0%**	**90.9%**	**92.1%**

2) Expression and Illumination Variations: as [13], the testing samples include the frontal face images with smile in Session 1, smile in Session 3, and surprise in Session 2. In each test, the images in the generic training set include all the frontal face images with the corresponding expression and the frontal face image with illumination 7 and neutral expression in Session 1. The recognition rates of all competing methods are listed in Table 2.

Table 2. Face recognition rates on Multi-PIE database with expression and illumination variations.

Expression	Smi-S1	Sim-S3	Sur-S2
NN	46.9%	28.8%	18.0%
SVM	46.9%	28.8%	18.0%
SRC [12]	49.6%	28.1%	20.4%
DMMA[8]	58.2%	31.5%	22.0%
AGL [18]	84.9%	39.3%	31.3%
ESRC [3]	81.6%	50.5%	49.6%
ESRC-KSVD	85.0%	50.4%	51.2%
SVDL	88.8%	58.6%	54.7%
LVJR	**93.7%**	**63.9%**	**67.6%**

We can see that LVJR outperforms all the other methods in all the three tests, with at least 4.9%, 5.3%, and 12.9% improvements over the second best, SVDL. That validates that the local information explored in our proposed LVJR is very helpful for final recognition. In addition, all the methods achieve the best results when Smi-S1 is used for testing because the training set is also from Session 1. Again, the methods using generic training set usually have better performance than the ones without using generic training set.

3) Pose, Illumination and Expression Variations: As [13], the testing samples include face images with pose '05_0' in session 2 and pose '04_1' in session 3, and face images with pose '04_1' and smile expression in Session 3 (please refer to Figs. 1(b)~(d) for examples). In each test, the images in the generic training set include all the face images with the corresponding expression and pose, and the frontal face image with illumination 7 and neutral expression in Session 1. The recognition rates of all competing methods are listed in Table 3.

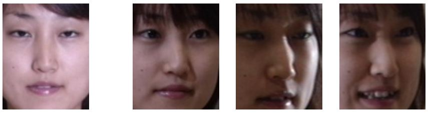

Fig. 1. Face images with pose variations in different sessions. (a) shows the single gallery sample; (b), (c) and (d) show the testing samples with pose, illumination and expression variations in Sessions 2 and 3, respectively.

Table 3. Face recognition rates on Multi-PIE database with pose, expression and illumination variations.

Pose	P05_0-S2	P04_1-S3	Smi-P04_1-S3
NN	26.0%	8.7%	12.0%
SVM	26.0%	8.7%	12.0%
SRC [12]	25.0%	7.3%	10.3%
DMMA[8]	27.1%	5.3%	11.0%
AGL [18]	66.7%	24.9%	23.9%
ESRC [3]	63.9%	31.8%	26.9%
ESRC-KSVD	67.1%	29.9%	25.6%
SVDL	77.8%	38.3%	34.4%
LVJR	**80.4%**	**40.0%**	**35.4%**

From Table 3, we can see that LVJR is still the best methods although the recognition rates of all methods are not high for big pose variation. This experiment also validate that the joint representation of local information on the learned variation dictionary could advance the recognition accuracy. We also run the experiments by learning a variation dictionary without requiring the representation of variation in the same category should be similar, of which the results (e.g., 79.7%, 31.9% and 33.1%) are lower than LVJR.

5.3 Evaluation on Various Variation AR Database

We then conduct FR with SSPP on the AR database [2]. As [16], a subset of AR contains two-session data of 50 male and 50 female subjects (each person has 26 pictures with the normalized size as 165×120) are included in the experiments. For each subjects there are two sessions and for variations (e.g., expression, illumination, disguise, and disguise+illumination). Here for each subject, the neutral face image without disguise and illumination in Session 1 is used as a gallery image. And the first 80 subjects are used to construct the gallery set and query set, with the remaining subjects for a generic training set. Here the face images are divided into 7×7 local regions and γ and λ are set as 0.001 and 0.05, respectively.

The recognition rates of the competing methods for query images from Session 1 are listed in Table 4. LVJR gets much better performance than all the other methods. For instance, compared to SVDL, the improvement for the variation of illumination and disguise is nearly 9%, which shows that local information could be effectively exploited by the proposed LVJR.

Table 4. Recognition accuracy (%) on AR database (Session1)

Pose	Illumination	expression	Disguise	Illumination+disguise
SVM	55.8	90.4	43.1	29.4
SRC [12]	80.8	85.4	55.6	25.3
DMMA[8]	92.1	81.4	46.9	30.9
AGL [18]	93.3	77.9	70.0	53.8
ESRC [3]	99.6	85.0	83.1	68.6
SVDL	98.3	86.3	86.3	79.4
LVJR	**100**	**94.6**	**93.1**	**88.4**

We also compare the proposed LVJR with Local generic representation (LRG) [16] in Session 1 of AR dataset. Following the experimental setting of [16], the face images are divided into 4×4 local regions. The accuracy and average running time on the same machine are listed in Table 5. LVJR is 25 times faster than LGR but with similar accuracy.

Table 5. Recognition accuracy (%) on AR database (Session1) of LGR and LVJR

Variation	Illumination	Expression	disguise
LGR	**100** (0.53second)	97.9(0.52second)	**98.8**(0.53second)
LVJR	**100(0.02 second)**	**98.8(0.02second)**	**98.8(0.02second)**

6 Conclusion

In this paper, we proposed a local variation joint representation method, which learns a variation dictionary with joint representation and does a joint and local collaborative representation. The learned variation dictionary could well exploit the variance information in the generic training set while with a small size. And the joint and local collaborative representation could fully use the local information of face images. The extensive experiments with various face variations demonstrated the superiority of LVJR to state-of-the-art methods for face recognition with SSPP.

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

Reference

1. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J. (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
2. Martinez, A., Benavente, R.: The AR face database. Technical Report 24, CVC (1998)
3. Deng, W.H., Hu, J.N., Guo, J.: Extended src: undersampled face recognition via intra-class variant dictionary. IEEE PAMI **34**(9), 1864–1870 (2012)
4. Yang, M., Zhang, L., Yang, J., Zhang, D.: Metaface learning for sparse representation based face recognition. In: ICIP (2010)

5. Rosasco, L., Verri, A., Santoro, M., Mosci, S., Villa, S.: Iterative Projection Methods for Structured Sparsity Regularization. MIT Technical Reports, MIT-CSAIL-TR-2009-050, CBCL-282
6. Wolf, L., Hassner, T., Taigman, Y.: Effective face recognition by combining multiple descriptors and learned background statistics. IEEE PAMI **33**(10), 1978–1990 (2011)
7. Zhang, L., Yang, M., Feng, X.C.: Sparse representation or collaborative representation: which helps face recognition? In: Proc. ICCV (2011)
8. Lu, J.W., Tan, Y.P., Wang, G.: Discriminative multi-manifold analysis for face recognition from a single training sample per person. IEEE PAMI **35**(1), 39–51 (2013)
9. Li, A.N., Shan, S.G., Gao, W.: Coupled bias–variance tradeoff for cross-pose face recognition. IEEE TIP **21**(1), 305–315 (2012)
10. Tzimiropoulos, G., Zafeiriou, S., Pantic, M.: Subspace learning from image gradient orientations. IEEE PAMI **34**(12), 2454–2466 (2012)
11. Wagner, A., Wright, J., Ganesh, A., Zhou, Z.H., Mobahi, H., Ma, Y.: Towards a practical face recognition system: robust alignment and illumination by sparse representation. IEEE PAMI **34**(2), 372–386 (2012)
12. Wright, J., Yang, A.Y., Sastry, S.S., Ma, Y.: Robust face recogntion via sparse representation. IEEE PAMI **31**(2), 210–227 (2009)
13. Yang, M., Van Gool, L., Zhang, L.: Sparse variation dictionary learning for face recognition with a single training sample per person. In: Proc. ICCV (2013)
14. Rubinstein, R., Bruckstein, A., Elad, M.: Dictionary learning for sparse representation modeling. Proceeding of the IEEE **98**(6), 1045–1057 (2010)
15. Belhumeur, P.N., Hespanha, J., Kriegman, D.J.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE PAMI **19**(7), 711–720 (1997)
16. Zhu, P., Yang, M., Zhang, L., Lee, I.-Y.: Local generic representation for face recognition with single sample per person. In: Cremers, D., Reid, I., Saito, H., Yang, M.-H. (eds.) ACCV 2014. LNCS, vol. 9005, pp. 34–50. Springer, Heidelberg (2015)
17. Zhao, W., Chellppa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Survey **35**(4), 399–458 (2003)
18. Su, Y., Shan, S., Chen, X., Gao, W.: Adaptive generic learning for face recognition from a single sample per person. In: CVPR (2010)
19. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. Image and Vision Computing **28**, 807–813 (2010)
20. Wang, J., Plataniotis, K., Lu, J., Venetsanopoulos, A.: On solving the face recognition problem with one training sample per subject. Pattern Recognition **39**, 1746–1762 (2006)
21. Mohammadzade, H., Hatzinakos, D.: Expression subspace projection for face recognition from single sample per person. IEEE Affective Computing **4**(1), 69–82 (2012)
22. Tan, X., Chen, S., Zhou, Z., Zhang, F.: Recognizing partially occluded expression variant faces from single training image per person with SOM and soft k-NN ensemble. IEEE NN **16**(4), 875–886 (2005)
23. Chen, S., Liu, J., Zhou, Z.: Making FLDA applicable to face recognition with one sample per person. Pattern Recognition **37**(7), 1553–1555 (2004)
24. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE SP **54**(11), 4311–4322
25. Zhang, D., Chen, S., Zhou, Z.: A new face recognition method based on SVD perturbation for single example image per person. Applied Mathematics and Computation **163**(2), 895–907 (2005)
26. Tan, X., Chen, S., Zhou, Z., Zhang, F.: Face recognition from a single image per person: A survey. Pattern Recognition **39**, 1725–1745 (2006)
27. Shan, S., Cao, B., Gao, W., Zhao, D.: Extended fisherface for face recognition from a single example image per person. ISCAS **2**, 81–84 (2002)