



Common Subspace Based Low-Rank and Joint Sparse Representation for Multi-view Face Recognition

Ziqiang Wang^{1,2}, Yingzhi Ouyang^{1,2}, Weidan Zhu^{1,2}, Bin Sun^{1,2(✉)},
and Qiang Liu^{1,2}

¹ School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China
{201721190135, 20171901011027, 201622190334}@std.uestc.edu.cn,
{sunbinhust, lqiang}@uestc.edu.cn

² Aircraft Swarm Intelligent Sensing and Cooperative Control Key Laboratory of Sichuan Province, Chengdu 611731, China

Abstract. Multi-view face data are very common in real-world application, since different viewpoints and various types of sensors attempt to better represent face data. However, these data have large pose variation, which dramatically degrades the performance of multi-view face recognition. To address this, we propose a common subspace based low-rank and joint sparse representation (CSLRJSR) method, which provides a framework encompassing divergence mitigation and feature fusion. In CSLRJSR method, common subspace is learnt to bridge the view, then low-rank and joint sparse representation are exploited to learn and then fuse the discriminative features. Experiments on multi-view face dataset demonstrate that CSLRJSR outperforms the state-of-the-art methods both in two-view and multi-view situations.

Keywords: Multi-view face recognition · Common subspace · Low-rank representation · Joint sparse

1 Introduction

Multi-view face recognition has been receiving a great deal of attention recently. Multi-view face data provide complementary information of the same object from multiple different views, which have positive effect on multi-view face recognition, compared to single-view face data. Researchers have already realized that fusing multi-view information can greatly improve face recognition performance [1, 2]. However, multi-view face recognition confronts a challenge of the wide variations of pose, illumination, and expression often encountered in realities. Therefore pursuing effective methods is still an urgent problem.

Sparse representation is explored extensively in face recognition [3, 4] and super resolution [5] due to its good performance in revealing underlying structure. Some sparse representation based methods are used for multi-view recognition tasks, which improve recognition performance by indirectly exploiting the inherent relationships

between multi-view data. Wang et al. [6] assumed one view can be represented sparsely by another view through a pair of dictionaries. Huang et al. [7] combined coupled dictionary and feature space learning (CDL) for two-view recognition and synthesis. Mandal et al. [8] proposed a generalized coupled dictionary learning (GCDL) method which learns dictionaries from two different views in a coupled manner, so that in the two views, the sparse coefficients of corresponding classes are correlated maximally. However, these methods are not suitable for scenarios more than two views, due to that they excessively rely on the coupling between views.

Subspace learning based methods have been introduced in pattern classification and data mining. These methods directly apply a specific low-rank projection to different view data. They can reduce differences between views and flexibly deal with scenarios more than two views. Several recognition approaches based on subspace learning were proposed recently, aiming at seeking for a common subspace where the data from multi-view can be compared [9]. The typical subspace learning method-principal component analysis (PCA) [10], is a technology of seeking a subspace in which the variance of the projected samples is maximized. Supervised regularization based robust subspace (SRRS) [11] presented a unified framework for subspace learning and data recovery. Ding et al. [12] proposed low-rank common subspace for multi-view learning (LRCS), which reserves compatible information among each view, through seeking a common discriminative subspace with a low-rank constraint. Robust multi-view data analysis through collective low-rank subspace (CLRS) proposed by Ding et al. [13] in 2018 complemented LRCS with a regularizer to further maximize the correlation within classes with the help of supervised information. Subspace learning methods perform well when reducing the divergence between views, but ignore the full use of extracted features, which actually limit the performance of recognition.

Recently, low-rank representation [14–17] and joint sparse [1, 18] are adopted in sparse representation. Low-rank representation based method has shown promising performance in capturing underlying low-rank structure as well as noise robustness. A joint sparse representation (SMBR) was proposed in [18] that enforce joint sparse constraints to the extracted features. SMBR took fully advantage of the extracted feature information to improve recognition performance.

Aforementioned multi-view face recognition methods still exist some intrinsic problems. To address these issues, we propose a novel multi-view face recognition algorithm, the contributions of this work are as follows:

- This paper focuses on the shared information across different views in order to handle the negative situation brought by multi-view face recognition, and puts forward common subspace based low-rank and joint sparse representation (CSLRJSR) method which unifies domain divergence mitigation and feature fusion.
- The method we propose is a more general algorithm that can be easily extended to handle more than two-view scenarios. Different scenarios of experiments were conducted to demonstrate the effectiveness of our algorithm. In many cases, our method outperforms the state-of-the-art multi-view recognition algorithms.

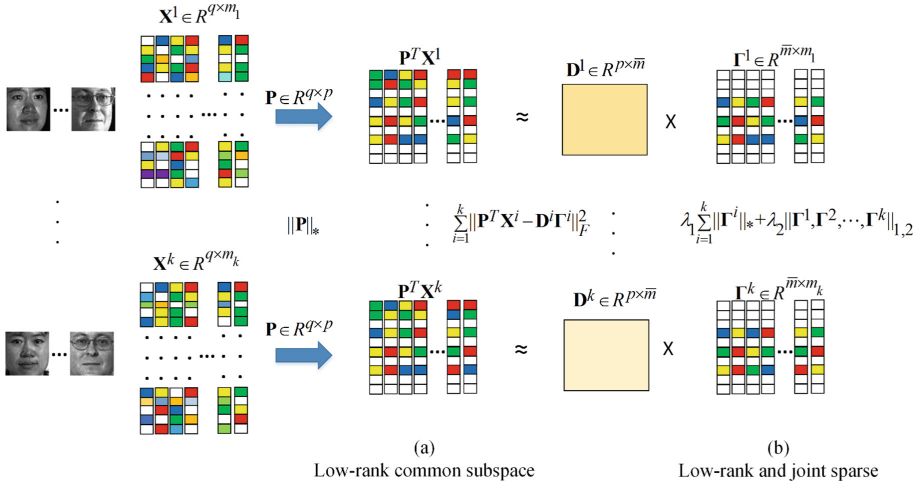


Fig. 1. Framework of our proposed CSLRJSR algorithm. (a) A common projection P for multi-view data to reduce divergences between views. (b) Low-rank and joint sparse constraints are imposed on sparse coefficients to reveal the global structure of multi-view data and fuse extracted features.

2 The Proposed Algorithm

In this section, we briefly present our motivation, and then propose our CSLRJSR method for the task of multi-view face recognition. Finally, we present the design of the optimization solution.

2.1 Motivation

Multi-view face data are ubiquitous in the real world, as the same objects are usually observed at different viewpoints, or even captured with different sensors. Therefore, the data of the same class but different views will show a big difference, which brings great challenges to the analysis of multi-view data. However, some research work [12, 13, 19] shows that the data of different views of the same class have a close relationship in the feature space. Some algorithms based on sparse representations [6–8] handle the task of multi-view recognition by integrating the inherent connection of multi-view data in the algorithm. The main idea of these methods is to learn the dictionary of two views in a coupled way, so that the same classes from the two views are most correlated in some transform space. These algorithms greatly improve the recognition performance of multi-view data. Unfortunately, these algorithms can only handle two views.

Subspace-based algorithms [12, 13, 20] directly process data from multiple views. The core idea of these methods is to directly apply a specific low-rank projection to different view data. In this way, the low-rank common subspace is learned, so that more shared information in different view data of the same class can be found. These algorithms take advantage of the relationship between multi-view data to improve

recognition performance and to generalize the processing of multiple views. However, feature information extracted by these algorithms is not further utilized. Following the core idea of subspace learning, we bridge different view data of the same class based on the common subspace to reduce the divergence between views. And the low-rank and joint sparse constraints are applied to the extracted features to achieve feature fusion, so that the complementary information of the features can be better utilized. Figure 1 shows an overview of the proposed algorithm. Next, we will introduce our algorithm in detail.

2.2 Common Subspace Based Sparse Representation

In multi-view face recognition, each view owns several classes and there are same classes across different views. Between these same classes, there is shared low-rank similar information. Thus, a low-rank common projection \mathbf{P} is utilized in our method to preserve this shared information, so that a same class from different views can be aligned into common subspace. Suppose there is k -view data $\mathbf{X} = [\mathbf{X}^1, \dots, \mathbf{X}^k]$, and each view $\mathbf{X}^i \in \mathbb{R}^{q \times m_i}$ includes m_i training samples and c same classes, where q is the original feature dimension of face sample. For each view $i = 1, \dots, k$, \mathbf{D}^i and $\mathbf{\Gamma}^i$ represent the corresponding dictionary and sparse coefficients, respectively. Therefore, the objective function is defined as:

$$\begin{aligned} \arg \min_{\mathbf{P}, \mathbf{D}^i, \mathbf{\Gamma}^i} \sum_{i=1}^k \|\mathbf{P}^T \mathbf{X}^i - \mathbf{D}^i \mathbf{\Gamma}^i\|_F^2 + \text{rank}(\mathbf{P}) \\ \text{s.t. } \mathbf{P}^T \mathbf{P} = \mathbf{I}, \quad \|\mathbf{d}_j^i\|_2 \leq 1, \quad (i = 1, 2, \dots, k) \end{aligned} \quad (1)$$

where $\text{rank}(\mathbf{P})$ denotes the rank operator of matrix $\mathbf{P} \in \mathbb{R}^{q \times p}$ (p is the reduced dimensionality). The orthogonal constraint $\mathbf{P}^T \mathbf{P} = \mathbf{I}$ (\mathbf{I} is a unit matrix) in Eq. (1) is applied to ensure the obtained \mathbf{P} is a valid solution. \mathbf{d}_j^i is an atom at the j -th column of dictionary \mathbf{D}^i . $\|\mathbf{M}\|_F = \sqrt{\sum_{i,j} \mathbf{M}_{ij}^2}$ denotes the Frobenius norm of matrix \mathbf{M} . Since rank minimization problem is a NP-hard problem in Eq. (1), recent researches adopt nuclear norm as a good surrogate [17]. However, after reducing the divergence between views and obtaining the ability to deal with scenarios more than two views, we notice that the extracted features haven't been fully used yet. Thus, we propose a common subspace based low-rank and joint sparse representation method to further exploit them, which is presented in Sect. 2.3.

2.3 Common Subspace Based Low-Rank and Joint Sparse Representation

In order to fully exploit the extracted features and improve the recognition performance, joint sparse constraints are imposed on the representation coefficients to achieve

feature-level fusion. Hence, we formulate the final objective function by unifying domain divergence mitigation and feature fusion as

$$\begin{aligned} \arg \min_{\mathbf{P}, \mathbf{D}^i, \mathbf{\Gamma}^i} \sum_{i=1}^k \left(\|\mathbf{P}^T \mathbf{X}^i - \mathbf{D}^i \mathbf{\Gamma}^i\|_F^2 + \lambda_1 \|\mathbf{\Gamma}^i\|_* \right) + \lambda_2 \|\mathbf{\Gamma}\|_{1,2} + \|\mathbf{P}\|_* \\ \text{s.t. } \mathbf{P}^T \mathbf{P} = \mathbf{I}, \quad \|\mathbf{d}_j^i\|_2 \leq 1, \quad (i = 1, 2, \dots, k) \end{aligned} \quad (2)$$

where we concatenate k coefficient matrices as $\mathbf{\Gamma} = [\mathbf{\Gamma}^1, \dots, \mathbf{\Gamma}^k]$. And $\|\mathbf{\Gamma}\|_{1,2}$ denotes a joint sparse constraint calculated via $\|\mathbf{\Gamma}\|_{1,2} = \sum_i \|\boldsymbol{\gamma}^i\|_2$ ($\boldsymbol{\gamma}^i$ denotes a vector on the i -th row of matrix $\mathbf{\Gamma}$), which is used to seek sparse nonzero rows so that all views can have similar sparse representations. And features are fused through this way. As for $\|\bullet\|_*$, it denotes a low-rank constraint that is applied to the sparse coefficient $\mathbf{\Gamma}^i$ in order to better expose the global structure of the data, by this way, learnt features become more discriminative. λ_1 and λ_2 are two positive tradeoff parameters. The detailed solution of the proposed algorithm is presented in Sect. 2.4.

2.4 Optimization

In this part, the alternating direction method of multipliers (ADMM) [21, 22] algorithm is used to deal with the optimization problem, as it converges well even several variables are non-smooth. We first introduce three auxiliary variables \mathbf{Z} , \mathbf{L}^i and \mathbf{W} , and then transform Eq. (2) into its equivalent constrained optimization problem as

$$\begin{aligned} \arg \min_{\mathbf{D}, \mathbf{P}, \mathbf{\Gamma}, \mathbf{Z}, \mathbf{L}^i, \mathbf{W}} \sum_{i=1}^k \left(\|\mathbf{P}^T \mathbf{X}^i - \mathbf{D}^i \mathbf{\Gamma}^i\|_F^2 + \lambda_1 \|\mathbf{L}^i\|_* \right) + \lambda_2 \|\mathbf{Z}\|_{1,2} + \|\mathbf{W}\|_* \\ \text{s.t. } \mathbf{\Gamma} = \mathbf{Z}, \mathbf{P} = \mathbf{W}, \mathbf{\Gamma}^i = \mathbf{L}^i, i = 1, \dots, k \end{aligned} \quad (3)$$

Equation (3) can be addressed using the Augmented Lagrangian Method (ALM) [21]. The Augmented Lagrangian function $f_{\alpha_Z, \alpha_L, \alpha_W}(\mathbf{P}, \mathbf{D}^i, \mathbf{\Gamma}, \mathbf{Z}, \mathbf{L}^i, \mathbf{W}; \mathbf{A}_Z, \mathbf{A}_L^i, \mathbf{A}_W)$ is defined as:

$$\begin{aligned} \arg \min_{\mathbf{D}^i, \mathbf{P}, \mathbf{\Gamma}, \mathbf{Z}, \mathbf{L}^i, \mathbf{W}} \sum_{i=1}^k \left(\|\mathbf{P}^T \mathbf{X}^i - \mathbf{D}^i \mathbf{\Gamma}^i\|_F + \lambda_1 \|\mathbf{L}^i\|_* + \langle \mathbf{A}_L^i, \mathbf{\Gamma}^i - \mathbf{L}^i \rangle + \frac{\alpha_L}{2} \|\mathbf{\Gamma}^i - \mathbf{L}^i\|_F^2 \right) \\ + \lambda_2 \|\mathbf{Z}\|_{1,2} + \langle \mathbf{A}_Z, \mathbf{\Gamma} - \mathbf{Z} \rangle + \frac{\alpha_Z}{2} \|\mathbf{\Gamma} - \mathbf{Z}\|_F^2 \\ + \|\mathbf{W}\|_* + \langle \mathbf{A}_W, \mathbf{P} - \mathbf{W} \rangle + \frac{\alpha_W}{2} \|\mathbf{P} - \mathbf{W}\|_F^2 \end{aligned} \quad (4)$$

where \mathbf{A}_Z , \mathbf{A}_L^i , \mathbf{A}_W are three Lagrange multipliers and α_Z , α_L , α_W are the positive penalty parameters, $\langle \mathbf{A}, \mathbf{B} \rangle$ denotes $\text{tr}(\mathbf{A}^T \mathbf{B})$, and $\mathbf{A}_Z = [\mathbf{A}_Z^1, \mathbf{A}_Z^2, \dots, \mathbf{A}_Z^k]$.

The truth is that it is difficult to jointly optimize the variables in Eq. (4). Fortunately, we can get optimization results in an iterative way. That is, we solve each variable by keeping the others fixed. Moreover, we denote \mathbf{P}_t , \mathbf{D}_t^i , $\mathbf{\Gamma}_t$, \mathbf{Z}_t , \mathbf{L}_t^i , \mathbf{W}_t , $\mathbf{A}_{Z,t}$,

$\mathbf{A}_{\mathbf{L},t}^i$, $\mathbf{A}_{\mathbf{W},t}$, $\alpha_{\mathbf{Z},t}$, $\alpha_{\mathbf{L},t}$ and $\alpha_{\mathbf{W},t}$ as the solutions optimized in the t -th iteration ($t > 0$). Then in the $(t+1)$ -th iteration, those solutions are updated as follows:

Updating \mathbf{L}^i :

$$\mathbf{L}_{t+1}^i = \arg \min_{\mathbf{L}^i} \frac{\lambda_2}{\alpha_{\mathbf{L},t}} \|\mathbf{L}^i\|_* + \frac{1}{2} \|\mathbf{L}^i - (\mathbf{\Gamma}_t^i + \frac{\mathbf{A}_{\mathbf{L},t}^i}{\alpha_{\mathbf{L},t}})\|_F^2 \quad (5)$$

Updating $\mathbf{\Gamma}^i$:

$$\mathbf{\Gamma}_{t+1}^i = (\mathbf{D}_t^{i^T} \mathbf{D}_t^i + \alpha_{\mathbf{Z},t} \mathbf{I} + \alpha_{\mathbf{L},t} \mathbf{I})^{-1} (\alpha_{\mathbf{Z},t} \mathbf{Z}_t^i + \alpha_{\mathbf{L},t} \mathbf{L}_{t+1}^i - \mathbf{A}_{\mathbf{Z},t}^i - \mathbf{A}_{\mathbf{L},t}^i + \mathbf{D}_t^{i^T} \mathbf{P}_t^T \mathbf{X}^i) \quad (6)$$

Updating \mathbf{D}^i :

$$\mathbf{D}_{t+1}^i = \arg \min_{\mathbf{D}^i} \|\mathbf{P}_t^T \mathbf{X}^i - \mathbf{D}^i \mathbf{\Gamma}_{t+1}^i\|_F^2 \quad s.t. \quad \|\mathbf{d}_j^i\|_2 \leq 1 \quad (7)$$

Updating \mathbf{W} :

$$\mathbf{W}_{t+1} = \arg \min_{\mathbf{W}} \frac{1}{\alpha_{\mathbf{W},t}} \|\mathbf{W}\|_* + \frac{1}{2} \|\mathbf{W} - (\mathbf{P}_t + \frac{\mathbf{A}_{\mathbf{W},t}}{\alpha_{\mathbf{W},t}})\|_F^2 \quad (8)$$

Updating \mathbf{P} :

$$\mathbf{P}_{t+1} = (\sum_{i=1}^k \mathbf{X}^i \mathbf{X}^{i^T} + \alpha_{\mathbf{W},t})^{-1} (\sum_{i=1}^k \mathbf{X}^i \mathbf{\Gamma}_{t+1}^{i^T} \mathbf{D}_{t+1}^{i^T} - \mathbf{A}_{\mathbf{W},t} + \alpha_{\mathbf{W}} \mathbf{W}_{t+1}) \quad (9)$$

Updating \mathbf{Z} :

$$\mathbf{Z}_{t+1} = \min_{\mathbf{Z}} \frac{\lambda_1}{\alpha_{\mathbf{Z},t}} \|\mathbf{Z}\|_{1,2} + \frac{1}{2} \|\mathbf{Z} - (\mathbf{\Gamma}_{t+1} + \alpha_{\mathbf{Z}}^{-1} \mathbf{A}_{\mathbf{Z},t})\|_F^2 \quad (10)$$

We use Singular Value Thresholding (SVT) [23] to solve Eqs. (5) and (8), and a quadratic problem solver [24] to address Eq. (7). Since the structure of Eq. (10) is separable, we solve it separately with respect to each row of \mathbf{Z} . For each row, we follow the method used in [18] to solve the following sub-problems:

$$\mathbf{z}_{i,t+1} = \min_{\mathbf{z}} \frac{1}{2} \|\mathbf{n} - \mathbf{z}\|_2^2 + \frac{\lambda}{\alpha_{\mathbf{z}}} \|\mathbf{z}\|_2 \quad (11)$$

where $\mathbf{n} = \gamma_{i,t+1} + \alpha_{\mathbf{z}}^{-1} \mathbf{a}_{\mathbf{z},t}$, $\gamma_{i,t+1}$, $\mathbf{a}_{\mathbf{z},t}$, $\mathbf{z}_{i,t+1}$ represent the i -th row of matrix $\mathbf{\Gamma}_{t+1}$, $\mathbf{A}_{\mathbf{Z},t}$ and \mathbf{Z}_{t+1} , respectively.

In conclusion, we present the detailed optimization procedure of CSLRJSR in Algorithm 1, in which $\alpha_{\mathbf{Z}}$, $\alpha_{\mathbf{L}}$ and $\alpha_{\mathbf{W}}$ are set empirically, while tuning the two tradeoff parameters λ_1 and λ_2 via experiments elaborated in the next section. We initialize \mathbf{P} randomly in the same way as [12] and initialize dictionary \mathbf{D} via an online dictionary

learning method that used in [8]. To determine the influence of different initialization methods, we use several traditional methods to initialize \mathbf{P} , which turns out that the final recognition performance tends to be exceedingly similar.

Algorithm 1 Solving Problem (CSLRJSR) by ADMM

Input: $\mathbf{X}^i (i = 1, \dots, k), \lambda_1, \lambda_2$

Initialization: $\mathbf{L}_0 = 0, \mathbf{Z}_0 = \mathbf{\Gamma}_0 = 0, \mathbf{W}_0 = 0, \mathbf{A}_{\mathbf{Z},0} = 0, \mathbf{A}_{\mathbf{L},0} = 0, \mathbf{A}_{\mathbf{W},0} = 0$

$\alpha_{\mathbf{Z}} = \alpha_{\mathbf{L}} = \alpha_{\mathbf{W}} = 0.9 \times 10^{-4}, \rho = 1.1, \varepsilon = 10^{-5}, t = 0$

While not converged **do**

1. Fix other variables and optimize \mathbf{L}^i through Eq.(5).
2. Fix other variables and optimize $\mathbf{\Gamma}^i$ through Eq.(6).
3. Fix other variables and optimize \mathbf{D}^i through Eq.(7).
4. Fix other variables and optimize \mathbf{W} through Eq.(8).
5. Fix other variables and optimize \mathbf{P} through Eq.(9).
6. Fix other variables and optimize \mathbf{Z} through Eq.(10).
7. Optimize the multipliers $\mathbf{A}_{\mathbf{Z},t+1}, \mathbf{A}_{\mathbf{L},t+1}, \mathbf{A}_{\mathbf{W},t+1}$ through

$$\mathbf{A}_{\mathbf{Z},t+1} = \mathbf{A}_{\mathbf{Z},t} + \alpha_{\mathbf{Z}}(\mathbf{\Gamma}_{t+1} - \mathbf{Z}_{t+1});$$

$$\mathbf{A}_{\mathbf{L},t+1}^i = \mathbf{A}_{\mathbf{L},t}^i + \alpha_{\mathbf{L}}(\mathbf{\Gamma}_{t+1}^i - \mathbf{Z}_{t+1}^i);$$

$$\mathbf{A}_{\mathbf{W},t+1} = \mathbf{A}_{\mathbf{W},t} + \alpha_{\mathbf{W}}(\mathbf{P}_{t+1} - \mathbf{W}_{t+1}).$$

8. Optimize the penalty parameter $\alpha_{\mathbf{Z},t+1}, \alpha_{\mathbf{L},t+1}, \alpha_{\mathbf{W},t+1}$ through:

$$\alpha_{\lambda,t+1} = \rho \alpha_{\lambda,t} \quad (\lambda = \mathbf{Z}, \mathbf{L}, \mathbf{W})$$

9. Check the convergence conditions

$$\|\mathbf{\Lambda}_{t+1} - \mathbf{Z}_{t+1}\|_{\infty} < \varepsilon,$$

$$\|\mathbf{P}_{t+1} - \mathbf{W}_{t+1}\|_{\infty} < \varepsilon.$$

10. $t = t + 1$

11. **end while**

output: $\mathbf{L}^i, \mathbf{\Gamma}, \mathbf{D}^i, \mathbf{W}, \mathbf{P}, \mathbf{Z}$

3 Experiments

In this section, a public multi-view dataset named CMU-PIE face dataset and experimental protocols are first introduced. Secondly, we demonstrate the comparison results of our proposed algorithm and the state-of-the-art algorithms. Lastly, for a comprehensive evaluation, several properties of our CSLRJSR approach are evaluated.

3.1 Dataset and Experimental Setting

The CMU-PIE Face dataset [25] consists of 68 subjects, each of which has multiple poses. Examples for each subject have 21 various illumination variations. In the experiments, we use face images of 7 poses (C02, C05, C07, C09, C14, C27, C29), where there are large differences between every subject in different poses (Fig. 2). Different numbers of poses are selected to build multiple evaluation subsets.

Furthermore, face images are cropped into size of 64×64 and only the raw features are used as the input. We randomly choose 10 samples from each subject each pose to construct the training set, whilst the remaining samples are used for testing.

3.2 Comparison Results

To demonstrate the effectiveness of our approach, we compare the proposed method with sparse representation based methods and subspace learning based methods. All of these methods directly or indirectly exploit the inherent characteristics of the relationship between multi-view data to improve recognition performance. Sparse representation based methods include SCDL [6], CDL [7], GCDL1&GCDL2 [8]; and subspace learning based methods include PCA [10], SRRS [11], LRCS [12] and CLRS [13]. Actually, SCDL, CDL, GCDL1 and GCDL2 are specially designed for two-view cases, thus, only the recognition performances of two-view cases are presented. While PCA, SRRS, LRCS and CLRS are used to process multi-view cases by seeking a robust subspace.

For all comparison algorithms, we adopt the nearest neighbor classifier to evaluate the final recognition performance. We conduct five random selections and average the results. Table 1 depicts recognition performance of different methods on CMU-PIE datasets, where Case1: {C02, C14}, Case2: {C05, C07}, Case3: {C05, C29}, Case4: {C07, C09}, Case5: {C09, C27}, Case6: {C09, C29}, Case7: {C05, C07, C29}, Case8: {C05, C27, C29}, Case9: {C07, C09, C27}, Case10: {C07, C09, C29}, and Case11: {C05, C07, C09, C29}.

From the experimental results, we obtain the following observations. (1) All methods based on sparse representation show good performance. SCDL, CDL, GCDL1 and GCDL2 are to learn the dictionary for the two-view cases in a coupled manner, so that for the corresponding classes from the two views, the sparse coefficients of them are maximally correlated in a certain transformed space. However, these methods are only applicable to two-view cases. While our method is completely compatible for multi-view cases, which applies low-rank common subspace constraint to different views to mitigate the differences between views. And the recognition performance of our proposed algorithm is better than the above algorithms in most case. (2) Comparing with other subspace learning methods, we apply low-rank and joint sparse constraints to learn and fuse discriminative features so that the recognition performance is improved.

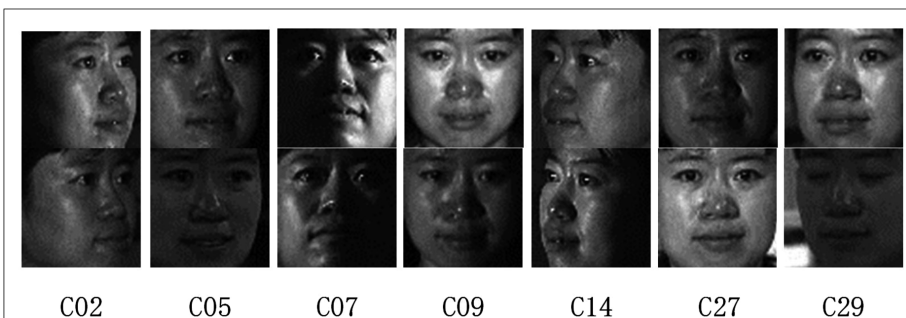


Fig. 2. Face samples from different views of one subject in CMU-PIE dataset. We can notice the differences across different views of the same subject.

Table 1. Comparison results between 9 algorithms on CMU-PIE dataset

Methods	SCDL	CDL	GCDL1	GCDL2	PCA	SRRS	LRCS	CLRS	Ours
Case1	85.18	84.43	84.38	84.54	66.47	74.52	86.35	86.03	87.69
Case2	90.35	88.97	88.91	88.43	63.22	61.14	88.27	88.06	90.14
Case3	89.71	88.81	88.75	88.17	64.93	74.36	88.70	76.33	91.68
Case4	90.41	89.66	88.97	88.75	68.28	76.39	90.72	80.06	91.10
Case5	90.09	89.13	88.49	88.75	73.13	78.31	90.19	87.26	90.41
Case6	89.93	89.45	88.86	89.07	69.08	71.96	91.10	86.57	91.95
Case7	/	/	/	/	64.39	72.96	85.18	75.41	91.22
Case8	/	/	/	/	64.53	72.67	83.01	75.27	91.26
Case9	/	/	/	/	73.77	72.03	87.74	88.91	90.41
Case10	/	/	/	/	68.12	66.56	87.63	88.77	91.65
Case11	/	/	/	/	66.07	63.09	84.81	75.88	91.68

3.3 Convergence Analysis and Parameters Analysis

In this part, we analyze several properties of our proposed method, i.e., convergence and parameter influence.

First, we carry out server experiments on convergence curve and recognition performance of different iterations. Specifically, we evaluate on two-view case {C02, C14} and the performance is presented in Fig. 3. From the results, we can see our algorithm converges pretty well. Also, we find that the recognition performance goes up quickly and keeps a relatively stable point.

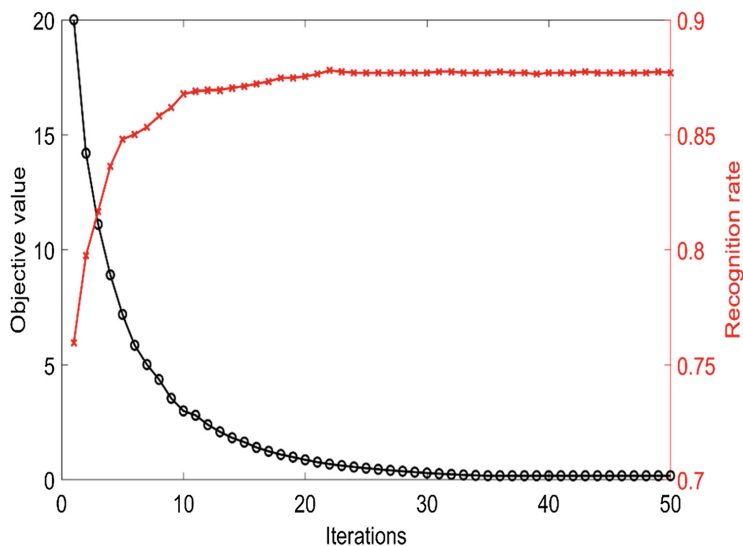


Fig. 3. Convergence curve (black ‘o’) and recognition curve (red ‘x’) of the proposed algorithm in two-view Case1 (C02&C14), where the dimensionality is set to 200 and the parameter values λ_1 and λ_2 to 10 and 0.1, respectively. (Color figure online)

Second, because there are two tradeoff parameters λ_1 , λ_2 in our method, we simultaneously analyze them on Case1 (C02&C14) and the results are presented in Fig. 4.

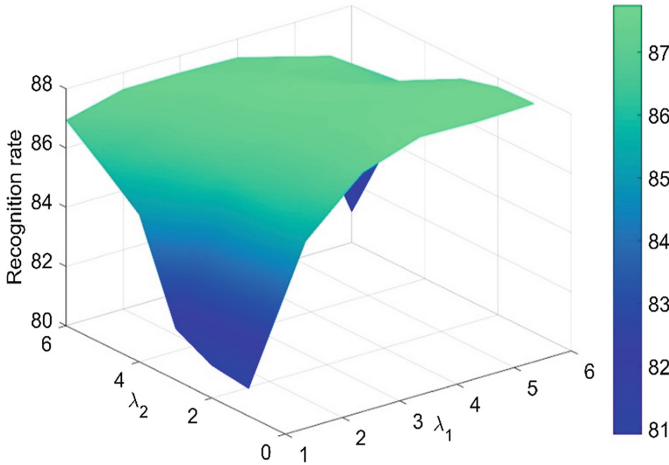


Fig. 4. We evaluate the recognition performance of our proposed algorithm on the two tradeoff parameters influence $\{\lambda_1, \lambda_2\}$ using two-view case (C02&C14). The value from 0 to 6 denotes $[0.1, 0.5, 1, 5, 10, 20]$, respectively.

According to the results, we can see that when λ_1 and λ_2 both are set small or large at the same time, the recognition performance is poor. On the contrary, we notice when $\lambda_1 \in [5, 15]$, it provides a much better performance. Therefore, we set $\lambda_1 = 10$ and $\lambda_2 = 0.1$ during the whole experiments.

4 Conclusion

In this paper, a common subspace based low-rank and joint sparse representation for multi-view face recognition method is proposed. Specifically, we apply low-rank common subspace projection to multi-view data to reduce differences between views of the same class, so that the discriminative ability of learnt features is improved. Furthermore, joint sparse is imposed to get a representation consistent across all the views, and then low-rank representation uncovers the global structure of data and further improves the discriminative power of the features. And finally, the discriminative features are learnt and efficiently fused. Experimental results on multi-view dataset, demonstrate the effectiveness and accuracy of the proposed algorithm, compared to several state-of-the-art algorithms.

Our future work will mainly focus on evaluating our method on more different multi-view datasets and improving the accuracy and robustness of our method on noisy datasets which are more common in real-world applications.

Acknowledgments. The work was supported by National Natural Science Foundation of China (No. 61803075). We thank the anonymous reviewers for their comments and suggestions which make the paper much improved.

References

1. Yuan, X., Liu, X., Yan, S.: Visual classification with multitask joint sparse representation. *IEEE Trans. Image Process.* **21**(10), 4349–4360 (2012)
2. Cao, L., Luo, J., Liang F., Huang, T.: Heterogeneous feature machines for visual recognition. In: *IEEE International Conference on Computer Vision (2009)*
3. Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009)
4. Jiang, X., Lai, J.: Sparse and dense hybrid representation via dictionary decomposition for face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(5), 1067–1079 (2015)
5. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution as sparse representation of raw image patches. In: *2008 Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8 (2008)
6. Wang, S., Zhang, L., Liang, Y., Pan, Q.: Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2216–2223 (2012)
7. Huang D., Wang, Y.: Coupled dictionary and feature space learning with applications to cross-domain image synthesis and recognition. In: *IEEE International Conference on Computer Vision*, pp. 2496–2503 (2013)
8. Mandal, D., Biswas, S.: Generalized coupled dictionary learning approach with applications to cross-modal matching. *IEEE Trans. Image Process.* **25**(8), 3826–3837 (2016)
9. Ouyang, S., Hospedales, T., Song, Y., et al.: A survey on heterogeneous face recognition: sketch, infra-red, 3D and low-resolution. *Image Vis. Comput.* **56**, 28–48 (2016)
10. Turk, M., Pentland, A.: Eigenfaces for recognition. *Cogn. Neurosci.* **3**(1), 71–86 (1991)
11. Li, S., Fu, Y.: Learning robust and discriminative subspace with low-rank constraints. *IEEE Trans. Neural Netw. Learn. Syst.* **27**(11), 2160–2173 (2016)
12. Ding, Z., Fu, Y.: Low-rank common subspace for multi-view learning. In: *Proceedings of the IEEE International Conference on Data Mining*, pp. 110–119 (2014)
13. Ding, Z., Fu, Y.: Robust multiview data analysis through collective low-rank subspace. *IEEE Trans. Neural Netw. Learn. Syst.* **29**(5), 1986–1997 (2018)
14. Li, J., Kong, Y., Zhao, H., Yang, J., Fu, Y.: Learning fast low-rank projection for image classification. *IEEE Trans. Image Process.* **25**(10), 4803–4814 (2016)
15. Li, L., Li, S., Fu, Y.: Learning low-rank and discriminative dictionary for image classification. *Image Vis. Comput.* **32**(10), 814–823 (2014)
16. Zhang, H., Patel, V., Chellappa, R.: Robust multimodal recognition via multitask multivariate low-rank representations. In: *11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, pp. 1–8 (2015)
17. Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., Ma, Y.: Robust recovery of subspace structures by low-rank representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(1), 171–184 (2013)
18. Shekhar, S., Patel, V., Nasrabadi, N., Chellappa, R.: Joint sparse representation for robust multimodal biometrics recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(1), 113–126 (2014)
19. Kan, M., Shan, S., Zhang, H., Lao, S., Chen, X.: Multi-view discriminant analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(1), 188–194 (2016)

20. Li, S., Fu, Y.: Robust subspace discovery through supervised low-rank constraints. In: Proceedings of SIAM International Conference on Data Mining, pp. 163–171 (2014)
21. Yang, J., Zhang, Y.: Alternating direction algorithms for l1 problems in compressive sensing. *SIAM J. Sci. Comput.* **33**(1), 250–278 (2011)
22. Afonso, M., Bioucas-Dias, J., Figueiredo, M.: An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.* **20**(3), 681–695 (2011)
23. Cai, J., Candès, E., Shen, Z.: A singular value Thresholding algorithm for matrix completion. *SIAM J. Optimiz.* **20**(4), 1956–1982 (2010)
24. Schölkopf, B., Platt, J., Hofmann, T.: Efficient sparse coding algorithms. In: Advances in Neural Information Processing Systems, pp. 801–808 (2007)
25. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination, and expression (PIE) database of human faces. In: Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, pp. 53–58 (2002)