



# Discriminative Subspace Learning for Cross-view Classification with Simultaneous Local and Global Alignment

Ao Li<sup>1</sup>(✉), Yu Ding<sup>1</sup>, Deyun Chen<sup>1</sup>, Guanglu Sun<sup>1</sup>, and Hailong Jiang<sup>2</sup>

<sup>1</sup> School of Computer Science and Technology,  
Harbin University of Science and Technology, Harbin 150080, China  
dargonboy@126.com

<sup>2</sup> Department of Computer Science, Kent State University, Kent, USA

**Abstract.** With the wide applications of cross-view data, cross-view Classification tasks draw much attention in recent years. Nevertheless, an intrinsic imperfection existed in cross-view data is that the data of the different views from the same semantic space are further than that within the same view but from different semantic spaces. To solve this special phenomenon, we design a novel discriminative subspace learning model via low-rank representation. The model maps cross-view data into a low-dimensional subspace. The main contributions of the proposed model include three points. 1) A self-representation model based on dual low-rank models is adopted, which can capture the class and view structures, respectively. 2) Two local graphs are designed to enforce the view-specific discriminative constraint for instances in a pair-wise way. 3) The global constraint on the mean vector of different classes is developed for further cross-view alignment. Experimental results on classification tasks with several public datasets prove that our proposed method outperforms other feature learning methods.

**Keywords:** Cross-view classification · Subspace learning · Discriminative analysis · Low-rank constraint.

## 1 Introduction

Cross-view classification is a noteworthy technology in the area of pattern recognition due to the large distribution discrepancy of data from different views. So, a well designed cross-view feature extraction plays a significant role in improving the classification performance. Nevertheless, it is worth noting that the samples in same view space are closer than that in different view spaces but from the same class space. Therefore, how to deal with the large view-variance between data from two views is a meaningful topic for subspace learning.

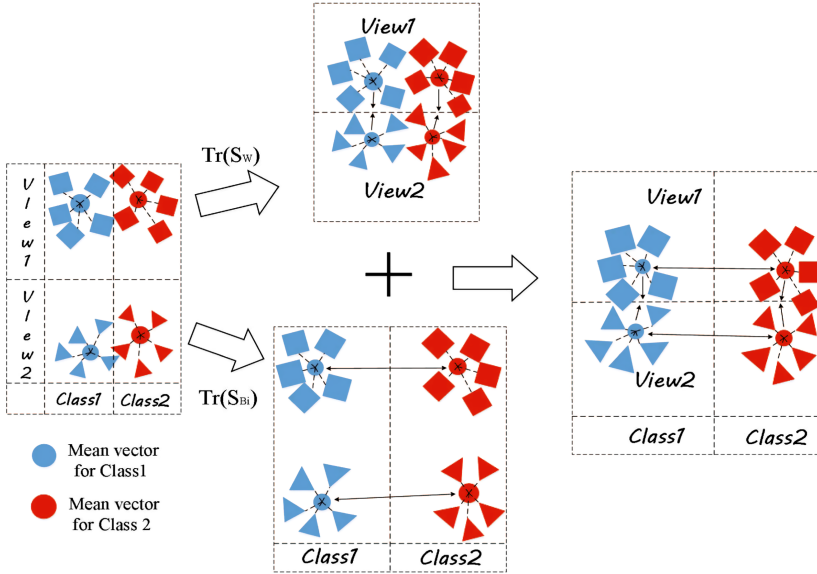
The aim of subspace learning is to find a feature space that can represent the high-dimensional data more effectively, which has been studied in many literatures [4–13]. Principal component analysis (PCA) [1], as a typical method, seeks

a projection subspace by maximizing variance. In order to make the learned subspace discriminative, linear discriminant analysis (LDA) was proposed with the Fisher criterion [2]. The feature subspace learned by LDA makes the intra-class samples more similar, and the inter-class samples less similar. Unfortunately, LDA may lead to overfitting for corrupted data. Recently, low-rank models to solve overfitting problems are very popular. Robust PCA (RPCA) [3] was designed to recover noisy data through the rank-minimization technique. Inspired by RPCA, low-rank representation (LRR) was designed to explore the intrinsic latent manifold structures of data from multiple subspaces. Liu et al. proposed the latent LRR (LatLRR) by considering the latent feature of the data [4]. After that, the supervised regularization-based robust subspace (SRRS) framework built in [5] provides a discriminative feature learning method that unifies the low-rank constraint and discriminant to learning low-dimensional subspace. In [6], an unsupervised robust linear feature extraction method was designed, namely low-rank embedding (LRE). LRE can reduce the negative impact of samples being occluded and corrupted. Ren et al. extended the LRE by introducing the  $l_{2,1}$ -norm term to make it more robust and effective [7].

Recently, a large number of cross-view feature learning algorithms also have achieved satisfied achievement [14–17]. However, some of methods ignore that samples from same view also have valuable discriminative information. To address this problem, Kan et al. designed a multi-view discriminant analysis (MvDA) [14]. MvDA can learn a multi-view feature subspace by optimizing multiple linear transforms from different views. Next, a most recent method was proposed, namely robust cross-view learning (RCVL) [15], which aims to learn an effective discriminative subspace by two discriminative graphs. Nevertheless, the global discriminative information is lost in RCVL.

Inspired by the above mentioned subspace learning methods, we propose a novel feature subspace learning model by exploring the local and global discriminative constraints simultaneously to realize the cross-view alignment. The main values of our methods are presented as follows: (1) The dual low-rank constraints framework is built to describe the two latent structures in cross-view data, namely the view structure and the class structure. Our method reveals the potential manifold of cross-view data so that the learned subspace contains more valuable feature information. (2) A local alignment mechanism based on two local graph constraints is adopted to constrain the neighbor relationships of the samples in the feature subspace. This mechanism can make the two structures in (1) to be separated effectively. (3) We set up a global alignment constraint as a complement to further reduce the effect of view-variance within the classes between views. The illustration of our framework is shown in Fig. 1, which learns a invariant subspace by maximizing the distance of the inter-class within each view and minimizing intra-class between views from both of class and view perspective.

The structure for the remainder of this paper is as follows. Section 2 simply introduces some of the related methods. Section 3 presents the proposed model



**Fig. 1.** Conceptual illustration of the proposed discriminative subspace learning framework

and its solution process. Section 4 shows the results of the comparison experiment and parameter experiment. At last, Section 5 summarizes this paper.

## 2 Related Works

There are two related methods involved in our framework: 1) low-rank representation, 2) linear discriminant analysis. Then, we discuss the role of two related works as follows.

### 2.1 Low-Rank Representation

Low-rank representation can cope well with data from multiple subspaces. Assuming  $X = [X_1, X_2, \dots, X_k]$  is a matrix of natural data from  $k$  categories. LRR can be expressed by

$$\min_{Z, E} \text{rank}(Z) + \lambda \|E\|_1, \text{ s.t. } X = XZ + E \quad (1)$$

in which  $Z$  is a low-rank linear combination coefficient matrix of data  $X$ . Generally, the samples data contains a lot of random noise. In function (1), matrix  $E$  denotes the noisy data, and we use  $l_1$ -norm to get randomness. In this way,  $XZ$  can recover the true data from noise.  $\lambda > 0$  is the balanced parameter. LRR can dig and utilize the self-similar information hidden in data. Therefore, LRR can not only learn the original subspace of the data in noisy environments, but also have the ability to uncover the latent manifold structure of data, which is belief to be feasible and potential for representing the cross-view data.

## 2.2 Linear Discriminant Analysis

The principle of LDA is to find a discriminative subspace with the largest inter-class variance and the smallest intra-class variance. Assuming training data  $\{X, y\} = \{(x_1, y_1), \dots, (x_n, y_n)\}$  is from  $m$  classes, where  $X$  denotes samples and  $y$  means label. In addition,  $\hat{x}$  represents the center of all samples, and  $\hat{x}_i$  represents the center of the  $i$ th class samples. Hence, between-class and within-class scatter matrices are expressed by:

$$S_b = \sum_{i=1}^m n_i (\hat{x}_i - \hat{x})(\hat{x}_i - \hat{x})^T$$

$$S_w = \sum_{i=1}^m \sum_{x \in X_i} (x - \hat{x}_i)(x - \hat{x}_i)^T$$
(2)

where  $n_i$  is the number of samples from the  $i$ th class, and  $X_i$  is the  $i$ th class samples set. Therefore, LDA finds a projection by maximizing the generalized Rayleigh quotient as follows:

$$\max_w \frac{Tr(w^T S_b w)}{Tr(w^T S_w w)}$$
(3)

where  $Tr(\cdot)$  denotes the trace operator and  $w$  is the projection matrix. In addition, the trace-ratio problem is not conducive to the solution of subsequent problems, so we transform function (3) into the trace-difference problem as follows:

$$\max_w Tr(w^T S_b w) - Tr(w^T S_w w)$$
(4)

LDA can reflect differences between samples based on a supervised discriminative constraint. However, its performance is not satisfied for cross-view analysis due to the large discrepancy of data distribution from different views.

## 3 The Proposed Algorithm

This section gives detailed discussions on our framework and develops a numerical scheme to obtain the approximate solutions iteratively.

### 3.1 Notations

Assuming  $X = [X_1, X_2] \in R^{d \times n}$  is a set of cross-view samples with two views from  $c$  classes, where  $n$  means the number of the training samples and  $d$  denotes the dimensionality of natural data. We design two local graph-based constraints to seek two latent view-invariant structures, which are composed of class structure matrix  $Z_c \in R^{n \times n}$  and view structure matrix  $Z_v \in R^{n \times n}$ , respectively.  $E \in R^{d \times n}$  is a matrix of error data designed to obtain a robust subspace from noise.  $P \in R^{d \times p}$  is the low-dimensional projection matrix. In addition,  $V_1, V_2, \hat{V}_1$  and  $\hat{V}_2 \in R^{n \times c}$  are constant coefficient matrices used to construct the discriminative global alignment constraint.

### 3.2 Objective Function

To address the discriminative cross-view analysis, we proposed a novel subspace learning with simultaneous local and global alignments, of which the objective function is as follows:

$$\begin{aligned}
 & \min_{Z_c, Z_v, E, P} \overbrace{\|Z_c\|_* + \|Z_v\|_* + \lambda_1 \|E\|_{2,1}}^{D(Z_c, Z_v, E)} \\
 & + \alpha \overbrace{(Tr(P^T X Z_c L_c (P^T X Z_c)^T) - Tr(P^T X Z_v L_v (P^T X Z_v)^T))}^{U(P, Z_c, Z_v)} \\
 & + \lambda_2 \overbrace{(Tr(S_W(P^T X Z)) - Tr(S_{B_1}(P^T X Z)) - Tr(S_{B_2}(P^T X Z)))}^{G(P, Z)} \\
 & s.t. X = X(Z_c + Z_v) + E, P^T P = I
 \end{aligned} \tag{5}$$

where  $D(Z_c, Z_v, E)$  represents class and view structures of the cross-view space by dual low-rank representations.  $U(P, Z_c, Z_v)$  enforces the view-specific discriminative local neighbor relationship among instances.  $G(P, Z)$  ( $Z = Z_c + Z_v$ ) presents the global discriminative constraint by the mean instance of each class cross different views. In short, we combine the local alignment constraint and the global alignment constraint on a dual low-rank framework to learn cross-view subspaces. In the following, the above terms are illustrated in detail.

**Dual Low-Rank Representations:** In general, we adopt a single rank minimum constraint to learn the latent information of data. But cross-view data contains class information and view information simultaneously. Even data from the same class have a large divergence. Hence, we use two structure matrices  $Z_c$  and  $Z_v$  to solve the specific problem that is the between-view samples from the same class are far away and the within-view samples from the different classes are closer. Thus, we define the first term with dual low-rank representations to strip down the class and view structures as follows:

$$D(Z_c, Z_v, E) = \|Z_c\|_* + \|Z_v\|_* + \lambda_1 \|E\|_{2,1}, s.t. X = X(Z_c + Z_v) + E \tag{6}$$

where  $\|\cdot\|_*$  is the symbol of the nuclear norm, which is an approximate representation of the rank minimum problem, and its solution relatively convenient. We adopt the  $l_{2,1}$ -norm to make matrix  $E$  have the sparsity as the noisy data.  $\lambda_1$  is a positive balance parameter, which can be tuned in experiments.

**Graph-Based Discriminative Local Alignment:** To introduce the local discriminative constraint, two graph-based constraints are constructed on the each pair of synthetic samples with  $Z_c$  and  $Z_v$  from class and view subspaces respectively as follows, which can better cluster intra-class samples and decentralize inter-class ones.

$$\begin{aligned}
 U_c &= \sum_{i,j} (Y_{c,i} - Y_{c,j})^2 W_{i,j}^c \\
 U_v &= \sum_{i,j} (Y_{v,i} - Y_{v,j})^2 W_{i,j}^v
 \end{aligned} \tag{7}$$

where  $Y_{c,i}, Y_{v,i}$  denote the  $i$ th projected sample of cross-view data in the class space  $Y_c = P^T X Z_c$  and view space  $Y_v = P^T X Z_v$ , respectively. Correspondingly,  $Y_{c,j}, Y_{v,j}$  denote the  $j$ th projected sample.  $W_{i,j}^c$  and  $W_{i,j}^v$  denote graph weight matrices that are defined as follows:

$$W_{i,j}^c = \begin{cases} 1, & \text{if } x_i \in N_{k_1}^c(x_j), \text{ and } l_i = l_j, \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$W_{i,j}^v = \begin{cases} 1, & \text{if } x_i \in N_{k_2}^v(x_j), \text{ but } l_i \neq l_j, \\ 0, & \text{otherwise} \end{cases}$$

where  $l_i$  and  $l_j$  are the labels of sample  $x_i, x_j$ , respectively.  $x_i \in N_{k_1}^c(x_j)$  denotes that  $x_i$  belongs to the  $k_1$  adjacent data sets of the same sample  $x_j$ .  $x_i \in N_{k_2}^v(x_j)$  means that  $x_i$  belongs to the  $k_2$  adjacent data sets of the same view sample  $x_j$ . With the help of trace operator, the pair-wise local discriminative constraint  $U(P, Z_c, Z_v)$  can be rewritten with  $U_c$  and  $U_v$  based on Fisher criterion as follows.

$$U(P, Z_c, Z_v) = \alpha(\text{Tr}(P^T X Z_c L_c (P^T X Z_c)^T) - \text{Tr}(P^T X Z_v L_v (P^T X Z_v)^T)) \quad (9)$$

where  $L_c$  and  $L_v$  mean the Laplacian operators of  $W^c$  and  $W^v$ .  $\alpha$  is a balance parameter, which can be tuned in experiments.

**Discriminative Global Alignment:** It is noteworthy that the  $U(P, Z_c, Z_v)$  preserves the discriminant in a local way by focusing on each pair of samples, which is not powerful enough. So, to further improve the proposed our model, we design a global discriminative constraint for cross-view analysis as the third term  $G(P, Z)$ , which is denoted as  $G(P, Z) = \text{Tr}(S_W(P^T X Z)) - \text{Tr}(S_{B1}(P^T X Z)) - \text{Tr}(S_{B2}(P^T X Z))$ . In  $G(P, Z)$ ,  $S_W(P^T X Z)$  is within-class scatter matrix of two views, defined by  $S_W(P^T X Z) = \sum_{j=1}^c (\mu_j^1 - \mu_j^2)(\mu_j^1 - \mu_j^2)^T$ .  $\text{Tr}(S_{B_i}(P^T X Z)) (i = 1, 2)$  is between-class scatter matrices of the  $i$ th view, defined by  $\text{Tr}(S_{B_i}(P^T X Z)) = \sum_{j=1}^c (\mu_j^i - \mu^i)(\mu_j^i - \mu^i)^T$ , where  $\mu_j^i$  denotes the mean projected sample of the  $j$ th class from the  $i$ th view, and  $\mu^i$  denotes the overall mean projected sample from the  $i$ th view. To be computed efficiently, the third term can be designed as:

$$G(P, Z) = \lambda_2(\text{Tr}(S_W(P^T X Z)) - \text{Tr}(S_{B1}(P^T X Z)) - \text{Tr}(S_{B2}(P^T X Z)))$$

$$= \lambda_2(\|P^T X Z(V_1 - V_2)\|_F^2 - \|P^T X Z(V_1 - \hat{V}_1)\|_F^2 - \|P^T X Z(V_2 - \hat{V}_2)\|_F^2) \quad (10)$$

where  $Z = Z_c + Z_v$  denotes the global representation.  $\lambda_2$  is a balance parameter, which can be tuned in experiments.  $V_i$  and  $\hat{V}_i (i = 1, 2)$  are the coefficient matrices of the within-class mean sample of each view and the global mean sample of each view, respectively. In detail,  $V_i(k, m) = (1/n_i^m)$  only if  $x_k$  belongs to the  $m$ th class from the  $i$ th view, where  $n_i^m$  means the number of samples of the  $m$ th class from the  $i$ th view; otherwise,  $V_i(k, m) = 0$ .  $\hat{V}_i(k, m) = (1/n_i)$  only if  $x_k$  belongs to the  $i$ th view, where  $n_i$  denotes the number of samples from the  $i$ th view; otherwise,  $\hat{V}_i(k, m) = 0$ . Equation (10) achieves global alignment by the

mean vectors of joint synthetic samples from global representation and further enforces the view-invariant constraint on the same class.

### 3.3 Optimization Scheme

To facilitate the solution of  $Z_c$  and  $Z_v$  in Eq. (5), we add two auxiliary variables  $J_c$  and  $J_v$ . Then, Eq. (5) can be transformed into the following form:

$$\begin{aligned} \min_{Z_c, Z_v, E, P} & \|J_c\|_* + \|J_v\|_* + \lambda_1 \|E\|_{2,1} + U(P, Z_c, Z_v) + G(P, Z) \\ \text{s.t.} & X = X(Z_c + Z_v) + E, P^T P = I, J_c = Z_c, J_v = Z_v \end{aligned} \quad (11)$$

Then, we transform the function (11) to the Augmented Lagrangian form, and the result is as follows:

$$\begin{aligned} \min_{Z_c, Z_v, E, P} & \|J_c\|_* + \|J_v\|_* + \lambda_1 \|E\|_{2,1} + U(P, Z_c, Z_v) + G(P, Z) \\ & + Tr(Y_1^T (X - X(Z_c + Z_v) - E)) + Tr(Y_2^T (J_c - Z_c)) + Tr(Y_3^T (J_v - Z_v)) \\ & + \frac{\eta}{2} (\|X - X(Z_c + Z_v) - E\|_F^2 + \|J_c - Z_c\|_F^2 + \|J_v - Z_v\|_F^2) \\ \text{s.t.} & P^T P = I \end{aligned} \quad (12)$$

where  $Y_1, Y_2, Y_3$  are the Lagrange multipliers and  $\eta > 0$  is the penalty parameter. We use an alternating solution to iteratively optimize all variables. We define the left-bottom of the variable plus  $t$  as the  $t$ -th solution.

First, we solve the projection matrix  $P_t$  one by one, because  $P_t$  is an orthogonal matrix. Hence, enforcing the derived function to be zero, the objective function with respect to  $P$  is:

$$\begin{aligned} & (XZ_t((V_1 - \hat{V}_1)(V_1 - \hat{V}_1)^T + (V_2 - \hat{V}_2)(V_2 - \hat{V}_2)^T - (V_1 - V_2)(V_1 - V_2)^T)Z_t^T X^T \\ & - \alpha(X(Z_{c,t}L_cZ_{c,t} - Z_{v,t}L_vZ_{v,t})X^T))P_{i,t} = \epsilon_{i,t}P_{i,t} \end{aligned} \quad (13)$$

Updating  $J_c$  and  $J_v$ :

$$J_{c,t+1} = \min_{J_{c,t}} \frac{1}{\eta_t} \|J_{c,t}\|_* + \frac{1}{2} \|J_{c,t} - (Z_{c,t} + (Y_{2,t}/\eta_t))\|_F^2 \quad (14)$$

$$J_{v,t+1} = \min_{J_{v,t}} \frac{1}{\eta_t} \|J_{v,t}\|_* + \frac{1}{2} \|J_{v,t} - (Z_{v,t} + (Y_{3,t}/\eta_t))\|_F^2 \quad (15)$$

The singular value thresholding is an approximate method to solve the above two kernel norm minimization equations[18].

Then, we ignore the other variables except  $Z_c$  or  $Z_v$  and make the function to be zero.

$$\begin{aligned} & Z_{c,t+1}((\lambda_2 H + \alpha L_c)/\eta_t) + X_N(I + X^T X)Z_{c,t+1} \\ & = X_N(X^T (X - XZ_{v,t} - E_t)) + J_{c,t} + ((X^T Y_{1,t} - Y_{2,t})/\eta_t) - \frac{\lambda_2}{\eta_t} Z_{v,t} H \end{aligned} \quad (16)$$

$$\begin{aligned}
 & Z_{v,t+1}((\lambda_2 H - \alpha L_v)/\eta_t) + X_N(I + X^T X)Z_{v,t+1} \\
 & = X_N(X^T(X - XZ_{c,t} - E_t)) + J_{v,t} + ((X^T Y_{1,t} - Y_{3,t})/\eta_t) - \frac{\lambda_2}{\eta_t} Z_{c,t} H \quad (17)
 \end{aligned}$$

where  $H = V_1 \hat{V}_1^T + V_2 \hat{V}_2^T + \hat{V}_1 V_1^T + \hat{V}_2 V_2^T - V_1 V_2^T - V_2 V_1^T - \hat{V}_1 \hat{V}_1^T - \hat{V}_2 \hat{V}_2^T$  and  $X_N = (X^T P_t P_t^T X)^{-1}$ . It is obvious that Eq. (16) and Eq. (17) are two standard Sylvester equations, which can be easily solved.

Updating  $E$ :

$$E_{t+1} = \min_{E_t} \frac{\lambda_1}{\eta_t} \|E_t\|_{2,1} + \frac{1}{2} \|E_t - (X - X(Z_{c,t} + Z_{v,t}) + Y_1/\eta_t)\|_F^2 \quad (18)$$

The above equation is a  $l_{2,1}$ -norm minimization problem whose solution is shown in [19].

---

### Algorithm 1

---

**Input:** data matrix  $X$ , parameters  $\lambda_1, \lambda_2, \alpha$

**Initialize:**  $\rho = 1.3$ ,  $\theta = 10^{-9}$ ,  $t = 0$ ,  $t_{max} = 200$ ,  $\eta_0 = 0.1$ ,  $\eta_{max} = 10^{10}$ ;

**while not converged or  $t \leq t_{max}$  do**

1. Calculate  $P_{t+1}$  using Eq.(13);

2. Calculate  $J_{c,t+1}$  using Eq.(14);

3. Calculate  $J_{v,t+1}$  using Eq.(15);

4. Calculate  $Z_{c,t+1}$  using Eq.(16);

5. Calculate  $Z_{v,t+1}$  using Eq.(17);

6. Calculate  $E_{t+1}$  using Eq.(18);

7. Calculate  $Y_{1,t+1}$  by

$$Y_{1,t+1} = Y_{1,t} + \eta_t(X - X(Z_{c,t+1} + Z_{v,t+1}) - E_{t+1})$$

8. Calculate  $Y_{2,t+1}$  by  $Y_{2,t+1} = Y_{2,t} + \eta_t(J_{c,t+1} - Z_{c,t+1})$

9. Calculate  $Y_{3,t+1}$  by  $Y_{3,t+1} = Y_{3,t} + \eta_t(J_{v,t+1} - Z_{v,t+1})$

10. Calculate the parameter  $\eta_{t+1}$  by  $\eta_{t+1} = \min(\eta_{max}, \rho\eta_t)$ ;

11. Check convergence by

$$\begin{aligned}
 & \max(\|X - X(Z_{c,t+1} + Z_{v,t+1}) - E_{t+1}\|_\infty, \\
 & \|J_{c,t+1} - Z_{c,t+1}\|_\infty, \|J_{v,t+1} - Z_{v,t+1}\|_\infty) < \theta;
 \end{aligned}$$

12.  $t = t + 1$ .

**end while**

**Output:**  $J_c, J_v, Z_c, Z_v, E, P$

---

The solution of all variables in objective function (12) is shown in Algorithm 1. Where these parameters  $\rho, \theta, t_{max}, \eta, \eta_{max}$  are set by experience. Moreover, the trade-off parameters  $\alpha, \lambda_1, \lambda_2$  are tuned by the experiments and we initialize these matrices  $Z_c, Z_v, E, Y_1, Y_2, Y_3$  as 0.

## 4 Experiments

In this section, we compare the proposed algorithm with excellent feature subspace learning algorithms. The data are mapped by the low-dimensional subspace as the features, on which the kNN classifier is implemented to evaluate the performance.



## 4.1 Experimental Datasets

**CMU-PIE Face dataset** is composed of face pictures of 68 different people. Everyone has 21 different illumination conditions and 9 different poses. We adopt 4 poses Pose05, Pose09, Pose27 and Pose29. We divide the dataset equally to set different cross-view training and testing subsets. **Wikipedia dataset** is an image-text bimodal data set, consisting of 2866 image-text samples from 10 classes. Due to the inconsistency of dimensionality of the two feature, we use PCA to adjust the image dimensions for the next experiment. **COIL-20 object dataset** is composed of 20 objects from a level 360-degree view. There are 5 degrees between every two adjacent images, so each category has 72 samples. We divide the 72 images into two groups G1 and G2. In addition, G1 is composed of samples from  $[0^\circ, 85^\circ]$  and V2  $[185^\circ, 265^\circ]$ . Similarly, G2 is composed of samples from  $[90^\circ, 175^\circ]$  and  $[270^\circ, 355^\circ]$ . **COIL-100 object dataset** is an extension of the COIL-20. The only difference is that the COIL-100 is composed of 20 objects from a level 360-degree view. Therefore, the set of the COIL-100 database is similar to the COIL-20 dataset.

## 4.2 Experimental Results and Analysis

In experiments, we select PCA, LDA, locality preserving projections (LPP) [20], LatLRR, SRRS, and RCVL as several comparison methods. To CMU-PIE dataset, we randomly select two poses to form a cross-view experiment set, where C1: {Pose05, Pose09}, C2: {Pose05, Pose27}, C3: {Pose05, Pose29}, C4: {Pose09, Pose 27}, C5: {Pose09, Pose29}, C6: {Pose27, Pose29}. We show the classification results of all experimental algorithms in Table 1. For COIL-20 and COIL-100 object databases, we select two sets of samples from G1 and G2 as a cross-view training set, and the others set as a test set. Figure 2 displays the classification results of four experimental groups from COIL-20 and COIL-100 datasets. Moreover, Fig. 3 displays the results of comparison experiments on Wikipedia.

**Table 1.** Classification results (%) of all methods on CMU-PIE dataset.

Methods	C1	C2	C3	C4	C5	C6
LDA	62.96	66.76	62.16	61.50	56.54	61.83
PCA	48.28	50.50	49.07	48.43	45.51	49.68
LatLRR	65.10	66.61	62.47	63.09	61.04	60.42
LPP	62.40	60.17	61.97	62.13	58.34	60.72
SRRS	95.35	91.66	95.82	90.22	96.04	87.16
RCVL	97.14	93.70	97.26	92.99	97.55	88.60
Ours	98.43	93.47	98.23	91.52	98.53	89.42

The results of experiments prove that our method achieves the persistent higher classification results than other methods. Another result also can be found that

the classification results of most methods based on LRR are better than that of other comparisons. It is due to that low-rank constraint can retain more valid latent structural information hidden in cross-view data. Nevertheless, among the comparison methods on different datasets, our proposed method is prominent and competitive consistently by considering more cross-view requirements on both local and global alignments.

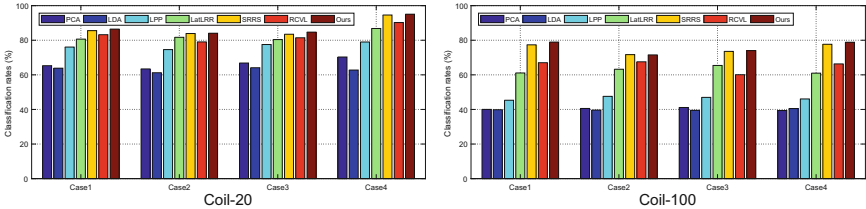


Fig. 2. Recognition results of different experiments on COIL-20 and COIL-100.

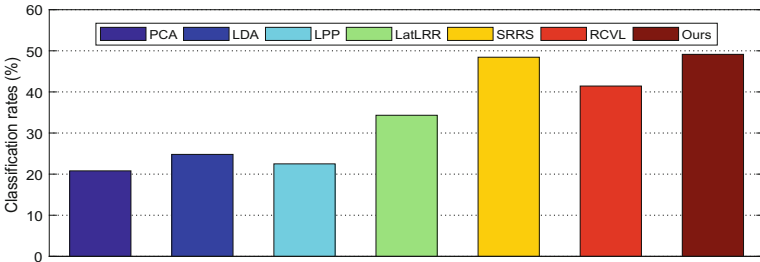
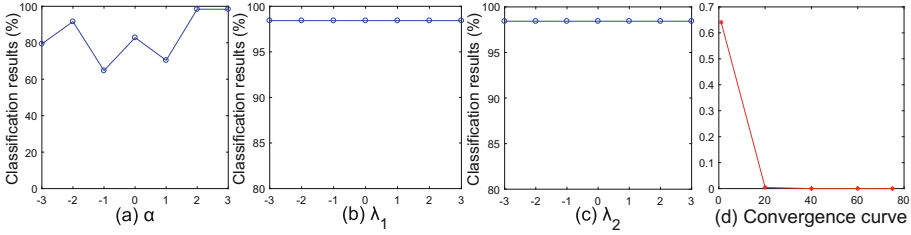


Fig. 3. Classification results of comparison experiments on Wikipedia dataset.

### 4.3 Performance Evaluations

In this section, we first point out how the parameters affect the classification performance. Then, the convergence analysis of the numerical scheme is also reported.

Our framework has three main parameters  $\alpha, \lambda_1, \lambda_2$ . We select CMU-PIE C1 to test their influence. The classification accuracy with variational parameters is shown in Fig.4. When the parameters  $\lambda_1$  and  $\lambda_2$  changes within a certain range, the performance fluctuation on classification is slight. For  $\alpha$ , the best performance happens at  $\alpha = 10^2$ . The results show that the parameters are not quite sensitive to the performance. More stable classification results can be obtained with the learned feature subspace of our proposed algorithm. In addition, the maximum value of  $\|X - X(Z_{c,t+1} + Z_{v,t+1}) - E_{t+1}\|_\infty, \|J_c - Z_c\|_\infty$



**Fig. 4.** Classification results with different values of (a)  $\alpha$ , (b)  $\lambda_1$  and (c)  $\lambda_2$ , where the parameters  $\alpha, \lambda_1, \lambda_2$  value from  $-3$  to  $3$  denotes  $[10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 10^2, 10^3]$ . (d) The convergence curve with increasing iterative step.

and  $\|J_v - Z_v\|_\infty$  is taken as the criteria to evaluate the convergence in each iteration. The convergence curve with increasing iterative step is shown in Fig. 4(d), from which it is confirmed that our method converges fast within a few of iterations.

## 5 Conclusion

In this paper, a discriminative subspace learning model based low-rank constraint is provided to apply the cross-view classification. The proposed method learns a projection subspace by finding and separating two potential manifold structures with dual low-rank representations. To further enhance the discriminant and adaptability, both of the local and global discriminants are utilized for cross-view alignment, which is validated to be helpful for learning view-invariant subspace. Moreover, to solve the proposed model, a reliable optimization scheme is also designed to ensure convergence. Extensive results prove that our method is superior to existing conventional methods.

**Acknowledgement.** This work was supported in part by National Natural Science Foundation of China under Grant 61501147, University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province under Grant UNPYSCT-2018203, Natural Science Foundation of Heilongjiang Province under Grant YQ2019F011, Fundamental Research Foundation for University of Heilongjiang Province under Grant LGYC2018JQ013, and Postdoctoral Foundation of Heilongjiang Province under Grant LBH-Q19112.

## References

1. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**(1), 71–86 (1991)
2. Belhumeur, P.-N., Hespanha, J.-P., Kriegman, D.-J.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 711–720 (1997)

3. Cands, E.-J., Li, X., Ma, Y., et al.: Robust principal component analysis? *J. ACM* **58**(3), 1–37 (2011)
4. Liu, G., Yan, S.: Latent low-rank representation for subspace segmentation and feature extraction. In: *IEEE International Conference on Computer Vision (ICCV)*, pp. 1615–1622. (2011)
5. Li, S., Fu, Y.: Learning robust and discriminative subspace with low-rank constraints. *IEEE Trans. Neural Networks Learn. Syst.* **27**(11), 2160–2173 (2016)
6. Wong, W.-K., Lai, Z., Wen, J., et al.: Low-rank embedding for robust image feature extraction. *IEEE Trans. Image Process.* **26**(6), 2905–2917 (2017)
7. Ren, Z., Sun, Q., Wu, B., et al.: Learning latent low-rank and sparse embedding for robust image feature extraction. *IEEE Trans. Image Process.* **29**, 2094–2107 (2020)
8. Li, A., Chen, D., Wu, Z., et al.: Self-supervised sparse coding scheme for image classification based on low-rank representation. *PLoS ONE* **13**(6), 1–15 (2018)
9. Li, A., Liu, X., Wang, Y., et al.: Subspace structural constraint-based discriminative feature learning via nonnegative low-rank representation. *PLoS ONE* **14**(5), 1–19 (2019)
10. Zhang, Z., Li, F., Zhao, M., et al.: Joint low-rank and sparse principal feature coding for enhanced robust representation and visual classification. *IEEE Trans. Image Process.* **25**(6), 2429–2443 (2016)
11. Zhang, Z., Ren, J., Li, S., et al.: Robust subspace discovery by block-diagonal adaptive locality-constrained representation. In: *ACM International Conference on Multimedia*, pp. 1569–1577 (2019)
12. Zhang, Z., Wang, L., Li, S., et al.: Adaptive structure-constrained robust latent low-rank coding for image recovery. In: *IEEE International Conference on Data Mining (ICDM)*, pp. 846–855 (2019)
13. Zhang, Z., Zhang, Y., Liu, G., et al.: Joint label prediction based semi-supervised adaptive concept factorization for robust data Representation. *IEEE Trans. Knowl. Data Eng.* **32**(5), 952–970 (2020)
14. Kan, M., Shan, S., Zhang, H., et al.: Multi-view discriminant analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(1), 188–194 (2016)
15. Ding, Z., Fu, Y.: Dual low-rank decompositions for robust cross-view learning. *IEEE Trans. Image Process.* **28**(1), 194–204 (2019)
16. Li, A., Wu, Z., Lu, H., et al.: Collaborative self-regression method with nonlinear feature based on multi-task learning for image classification. *IEEE Access* **6**, 43513–43525 (2018)
17. Zhang, Y., Zhang, Z., Qin, J., et al.: Semi-supervised local multi-manifold isomap by linear embedding for feature extraction. *Pattern Recogn.* **76**, 662–678 (2018)
18. Cai, J.-F., Cands, E.-J., Shen, Z.: A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **20**(4), 1956–1982 (2010)
19. Yang, J., Yin, W., Zhang, Y., et al.: A fast algorithm for edge-preserving variational multichannel image restoration. *SIAM J. Imaging Sci.* **2**(2), 569–592 (2009)
20. He, X., Niyogi, P.: Locality preserving projections. In: *Advances in Neural Information Processing Systems*, pp. 153–160 (2004)