Multi-view intrinsic low-rank representation
for robust face recognition and clustering

Zhi-yang Wang | Stanley Ebhohimhen Abhadiomhen | Zhi-feng Liu | Xiang-jun Shen | Wen-yun Gao | Shu-ying Li

1 School of Computer Science and Communication Engineering, JiangSu University, Zhenjiang, JiangSu, China
2 Department of Computer Science, University of Nigeria, Nsukka, Nigeria
3 Jingkou New-Generation Information Technology Industry Institute, Jiangsu University, Zhenjiang, Jiangsu, China
4 Nanjing LES Information Technology Co., LTD, Nanjing, Jiangsu, China
5 College of Computer and Information, Hohai University, Nanjing, Jiangsu, China
6 School of Automation, Xi’an University of Posts & Telecommunications, Xi’an, Shaanxi, China

Abstract
In the last years, subspace-based multi-view face recognition has attracted increasing attention and many related methods have been proposed. However, the most existing methods ignore the specific local structure of different views. This drawback can cause these methods’ discriminating ability to degrade when many noisy samples exist in data. To tackle this problem, a multi-view low-rank representation method is proposed, which exploits both intrinsic relationships and specific local structures of different views simultaneously. It is achieved by hierarchical Bayesian methods that constrain the low-rank representation of each view so that it matches a linear combination of an intrinsic representation matrix and a specific representation matrix to obtain common and specific characteristics of different views. The intrinsic representation matrix holds the consensus information between views, and the specific representation matrices indicate the diversity among views. Furthermore, the model injects a clustering structure into the low-rank representation. This approach allows for adaptive adjustment of the clustering structure while pursuing the optimization of the low-rank representation. Hence, the model can well capture both the relationship between data and the clustering structure explicitly. Extensive experiments on several datasets demonstrated the effectiveness of the proposed method compared to similar state-of-the-art methods in classification and clustering.

1 INTRODUCTION

Multi-view face data has become pervasive in the real-world. For example, several images taken from different angles can represent different views of a person, which can be exploited by methods to improve face recognition accuracy. Therefore, a naive method such as in [1–3] concatenates the features in different views to obtain a common feature representation. However, this strategy ignores the correlation between different views, which then incites the “curse of dimensionality” and result in high computational time. Besides, the methods based on feature concatenation do not comprehensively explore underlying data distribution among different views. So, the concatenated feature representation may not be meaningful. This drawback expedited some more researches, which focused on integrating complementary information through graph construction. In [4], Kumar et al. proposed a method, which found a unified graph by reducing the disagreement between different views.
views. In [5], Xia et al. attempted to separate noise from graphs through a low-rank transition probability matrix. In [6], Huang et al. performed face recognition tasks through an embedding graph learned simultaneously from different views. In [7], Cao et al. proposed a constrained multi-view method, which applied pairwise constraint to facilitate finding a consistent graph. In [8], Li et al. attained high accuracy and low computational complexity by using a bipartite graph to approximate the similarity graphs of different views. Nonetheless, these methods designed based on traditional spectral learning [9–11] did not utilize the self-expressiveness property of data. Therefore, they can be susceptible to noise in the different views.

Recently, multi-view subspace methods based on Low-Rank Representation (LRR) [12] and Sparse Subspace Clustering (SSC) [13] have become increasingly popular. Because both baseline methods utilize the self-expressiveness property of data to represent each data point as a linear combination of others in order to learn a computationally efficient data representation. Even so, some studies such as [14] suggested more that LRR may indeed provide a more robust approach to tackle the above drawback considering that it captures a global structure of data. Thus, many methods, which LRR inspired have been proposed over the years. For instance, Structure-Constrained Low-Rank Representation [15] analyzed the structure of multiple disjoint sub-spaces, which is more general for computer vision data. Latent Low-Rank Representation (LatLRR) [16] was proposed for latent detection in images using subspace segmentation. Robust multi-view subspace learning [17] used dual low-rank decompositions to uncover local manifold structures. Multi-Spectral Low-Rank Structured Dictionary Learning (MLSDL) [18] applied low-rank structural incoherence term to reduce the discrepancy between different views. Low-Rank Graph Optimization for Multi-View Dimensionality Reduction (LRGO-MVDR) [19] used an adaptive non-negative weight vector to capture the complementary structure in different face images. Common Subspace based Low-Rank and Joint Sparse Representation (CSLRRJSR) [20] found common face representation fusing discriminative features in different views through a joint low-rank and sparse methods. Despite these efforts, two main drawbacks limit the overall performance of the subspace methods. First, because subspace methods assume that similar data points stay near each other, they often do not perform well in practical cases when the distribution of data points within a subspace is arbitrary due to excessive noise. This weakness is because two unrelated data points may stay connected, and it is even more feasible knowing that these methods utilize a fixed shared representation. Secondly, most of these methods consider only the consistency of multi-view data or enhance only the diversity in these data similar to [21], so both common structure and intrinsic structure can mix up easily to degrade recognition accuracy.

Regarding the above limitations, we propose a novel method, which integrates clustering structure into low-rank representation to pursue optimization of multi-view intrinsic representation, which is used subsequently for face recognition. To achieve this, first, we follow the intuition of Hierarchical Bayes [22–24] by decomposing the low-rank representation learned for each view into two matrices, namely the intrinsic representation and the specific representation. The intrinsic representation matrix captures the common features of different views. And the specific representation matrix captures specific features of each view through diversity regularization. Then the core of our idea is introduced by imposing the spectral clustering structure into our model in order to guarantee more optimum representation that is devoid of specific redundant information. In other words, an affinity matrix in introduced into our model that captures the clustering structure explicitly, such that the clustering structure is integrated into the low-rank representation to allow our method to find a better intrinsic representation adaptively so that it can be used for face recognition reducing the effect of noisy data. Therefore, our model can be used for face recognition and clustering simultaneously. Figure 1 shows the framework of our proposed method.

The main contributions of our work are:

1. We propose a novel method where a low-rank representation learned for each view is composed of two low-rank components: an intrinsic representation across views and a specific view representation in each view. The intrinsic representation captures the common structure across all views, while the specific representation captures the diversity among views. With these structures, we can discover the correlation and discrepancy in multi-view face data comprehensively.

2. Furthermore, different from the existing methods which utilize a fixed representation of multi-view face data, our method obtains an intrinsic representation of the data flexibly. We achieve this by integrating a clustering structure into the low-rank representation to allow the clustering structure to be adjusted adaptively while pursuing an ideal low-rank face representation that will suppress the effect of noisy data. As a result, our proposed model can be used for face recognition and clustering simultaneously.

3. We conduct experiments on face benchmark datasets, and the experimental results demonstrate the effectiveness of our proposed method.

2 RELATED WORK

Many studies have been conducted in search of an efficient technique that can explore the complementary information in multi-view face data to improve recognition accuracy. The preliminary methods such as feature concatenation [1–3], spectral graph [4–8] are not computationally efficient, and they do not provide robustness against noise. Recently, there has been much more attention on the subspace approach because it provides a mechanism to tackle the above problems. So far, LRR [12] and SSC [13] are the two most effective approaches in subspace segmentation, and they both utilize the self-expressiveness property of data [25]. SSC, however, does not capture a global structure of the data. The reason is that it uses the L1 norm to find a sparse representation of data points individually, which means that SSC may be susceptible to noise. LRR, on the other hand, provides an efficient tool to recover an actual manifold structure from corrupted data by using the data itself as the dictionary
FIGURE 1 Framework of our proposed method. For multi-view data matrix $X_i$, we obtain intrinsic low-rank matrix $Z_0$ by diversity constraint and affinity matrix $W$ by k-nearest neighbor graph. $Z_0$ holds the intrinsic information across each view and $Z_i$ preserves difference between different views. Then with a dynamic approach on the affinity matrix and intrinsic low-rank matrix, the local manifold and clustering structure are simultaneously learned (self-expressiveness matrix) to seek a representation matrix with the lowest rank [26]. To illustrate further, let’s assume that we have a set of data points $X = \{x_1, ..., x_n\} \in \mathbb{R}^{d \times n}$ where $n$ is the number of data points and $d$ is the dimensionality of each data point, then LRR will represent each data point by a linear combination of others to capture a global structure of data, that is, $x_i = Xp_i$, $i = 1, ..., n$, where $P = \{p_1, ..., p_n\} \in \mathbb{R}^{d \times n}$ is the self-expressiveness matrix, which can be found in the following way:

$$\min P \text{ Rank}(P) + \|E\|_2 \text{ s.t. } X = XP + E. \quad (1)$$

However, Equation (1) is hard to achieve due to the discrete nature of the rank function. Instead, the nuclear norm provides a good convex surrogate as follows:

$$\min_P \|P\|_* + \|E\|_2 \text{ s.t. } X = XP + E. \quad (2)$$

For the data $x_i$, it is one of the $k$ classes and can be connected by all the data points with weighted value $p_{ij}$, and such a weighted value can be seen easily as the similarity between them. In other words, the degree of correlation between $x_i$ and $x_j$ can be characterized by $p_{ij}$, where closer data points will have a larger weighted value. Therefore, $p_{ij}$ can correctly depict the local manifold structure of data when two similar data points reconstruct one another.

Considering the robustness of LRR, many methods have been proposed recently for multi-view face representation based on its principles. In [18], Jing et al. learn common and separate dictionaries. The main idea is to utilize these dictionaries to explore the correlated information and the complementary information simultaneously to find a robust face representation that will enhance face recognition. In [28], Xu et al. proposed a structured low-rank matrix recovery method to solve the problem of huge divergence in multi-view face data. This approach eliminates graph embedding by introducing a class label matrix to learn a discriminative unified matrix. Yet, it fails to consider the specific local structure of different views so, the problem may persist. In [17], Ding et al. had previously used a dual low-rank decomposition approach to separate local structures from the global structure. However, the dual minimization of singular values may cause the face representation matrix to be of high rank since the rank may not approximate correctly in practice [29]. Alternatively, Zheng et al. [30], Brbic and Kopriva [27], and Wang et al. [20] combine the idea of SSC and LRR to find a joint subspace representation that is robust against large-scale noise. In [31], Meng et al. proposed a Multi-View Low-Rank Preserving Embedding (MvLPE) method, which learns the complimentary features in different views by maximizing the agreement between an individual view and a centroid view. Nonetheless, these methods equally ignored the specific local structure of different views as [28]. Therefore, inspired by [17], we utilize the L1 norm to capture the specific local structure of each view such that a common face representation learned is through a cooperative effort to guarantee a more optimum solution that suppresses the effect of noisy data adaptably.

3 | THE PROPOSED APPROACH

In this section, we present a novel multi-view learning algorithm, which can capture the relationship among data and unfold the clustering structure at the same time. The notations are summarized in Table 1.

3.1 | Structure consistency

LRR is a promising method, which intends to build a good affinity matrix by using self-expression. However, it ignores the important data locality. To better exploit local manifold structure, we follow the intuition of Hierarchical Bayes [22–24] and
assumed that $P$ can be decomposed into two parts; one is the intrinsic structure shared across each view and the other is the diversity structure of each view. We utilize the intrinsic matrix $Z_0$ to capture the general structure across different views. Simultaneously, we use $Z_i$ to capture specific features of each view. We utilize the constraint $P_i = Z_i + Z_0$ to model the assumption above. We envision that the parts unique to each view are relatively sparse structures, so we utilize $l_1$ norm on $Z_i$ and leading to the problem below:

$$
\min_{Z,F} \sum_{i=1}^{v} (\|P_i\|_1 + \|Z_i\|_1) + \sum_{i=1}^{v} \|E_i\|_{2,1},
$$

s.t., $X_i = X_i P_i + E_i$, $P_i = Z_0 + Z_i, i = 1, \ldots, v$. 

### 3.2 Diversity regularization

Multi-view face data contain intrinsic and complementary features and general shared information at the same time. In this way, the intrinsic feature matrices should have difference between each other. Meanwhile, $Z_0$ and $Z_i$ should be different and complementary. In another word, each view should have less similarity and have weak connection with common feature as well. To achieve this, we introduce a diversity regularization formulated as follow:

$$\Theta(Z_0, Z_i) = \text{Tr} \left( Z_i^T Z_i \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} (Z_{i})_{ij} (Z_{i})_{ij},$$

where $Z_i$ and $Z_o$ are different views. According to the equation above, when $\Theta(Z_o, Z_i)$ is minimized, one of $Z_o$ and $Z_i$ value will be large and the other will be small. Therefore, we can utilize this to compute the differences between each view. We formulate the diversity regularization in our model as:

$$\sum_{i=1}^{v} \Theta(Z_i, Z_0) + \sum_{i=1}^{v} \sum_{i \neq j} \Theta(Z_i, Z_j)$$

$$= \sum_{i=1}^{v} \text{Tr} \left( Z_i^T Z_0 \right) + \sum_{i=1}^{v} \sum_{i \neq j} \text{Tr} \left( Z_i^T Z_j \right).$$

### 3.3 Integrating spectral clustering structure

Self-representation is a good method to capture the local manifold structure. The representation coefficient is larger when two different samples are similar while it is smaller when two samples are different. However, in many real-world applications, there are possibilities that similar samples belong to different classes. Thus, only a self-representation of data will mislead the clustering task, for the reason that many samples sharing different labels may be clustered together. Therefore, directly using the self-representation of data to clustering will not get ideal results at all time. Because local manifold captured by self-representation is not an ideal clustering structure. Basically, an ideal graph should ensure that samples with the same label should have strong connection while samples from different classes should have weak connection or even not be connected. Therefore, to exactly separate sample data into $k$ clusters, we need a clustering structure with $k$ connected components. In this way, we can deal with the problem that samples with different labels have strong connection. To this end, we impose a rank constraint on the Laplacian matrix derived from the affinity matrix so that it can be an ideal clustering structure. This rank constraint can guarantee that the affinity matrix has exact the $k$ connected components which is the same as clusters number. Meanwhile, the estimated affinity matrix will dynamically approximate local manifold captured from self-representation to exploit structure information among data. To solve the problem, we define an ideal affinity matrix $W = \{w_1, \ldots, w_v\} \in \mathbb{R}^{m \times n}$ that can not only utilize self-representation structure, but also is an ideal clustering structure.

**Theorem 1.** If affinity matrix $W$ is nonnegative, the multiplicity $K$ of the eigenvalue zero of the Laplacian matrix $L$ is equal to the number of connected components in the graph associated with $W$.

Given an affinity matrix $W$, Theory 1 indicates that if $\text{rank}(L) = n - k$, then $W$ is an ideal affinity matrix that can separate the data points into exact $k$ clusters. Moreover we hope $W$ dynamically approximates $Z$ used to capture the local manifold structure. Thus, we impose a $F$-norm to appropriately minimize the loss between $Z_0$ and $W$. Thus, we have the following formulation:

$$\min \|Z_0 - W\|^2_F + \lambda_1 \sum_{i=1}^{v} \left( \|P_i\|_1 + \|Z_i\|_1 \right)$$

$$+ \lambda_2 \sum_{i=1}^{v} \text{Tr} \left( Z_i^T Z_0 \right) + \lambda_2 \sum_{i=1}^{v} \sum_{j \neq i} \text{Tr} \left( Z_i^T Z_j \right)$$

$$+ \sum_{i=1}^{v} \|E_i\|_{2,1},
$$

s.t. $X_i = X_i P_i + E_i$, $P_i = Z_i + Z_0$, $\text{rank}(L) = n - k$.

$W_{ij} \geq 0, \sum_j W_{ij} = 1.$

### Table 1: The notations used in this paper

| Notations | Explanation |
|-----------|-------------|
| $X_i$     | Data matrix of $i$th view |
| $P_i$     | Self-representation matrix of $i$th view |
| $E_i$     | Error matrix of $i$th view |
| $Z_i$     | Intrinsic structure matrix of $i$th view |
| $Z_0$     | Common structure matrix of $i$th view |
| $W$       | Affinity matrix |
| $L$       | Laplacian matrix |
3.4 Optimization

We propose an optimization method to approximatively solve the objective function in Equation (6) and apply the Augmented Lagrange Multiplier (ALM) [32, 33] to iteratively update all the variables. We denote \( \Theta_j \) as the \( j \)-th smallest eigenvalue of \( L \). Because \( L \) is a positive semidefinite matrix, so \( \Theta_j \geq 0 \). In this way, the rank constraint in Equation (6) can be removed after given a large enough \( \lambda \). Therefore, the problem (6) can be rewritten as follows.

\[
\begin{align*}
\min \| Z_0 - W \|_F^2 + \lambda_1 \sum_{i=1}^r (\| P_i \|_* + \| Z_i \|_1) \\
+ \lambda_2 \sum_{i=1}^r \text{Tr} (Z_i^T Z_0) + \lambda_2 \sum_{i=1}^r \sum_{i \neq j} \text{Tr} (Z_i^T Z_j) \\
+ \lambda_3 \sum_{i=1}^k \Theta_i + \sum_{i=1}^r \| E_i \|_{2,1}\end{align*}
\]

s.t. \( X_j = X_j P_j + E_j, P_j = Z_j + Z_0, \ W_{ij} \geq 0, \sum_j W_{ij} = 1. \)

When \( \lambda_3 \) is large enough, the term \( \sum_{i=1}^k \Theta_i \) will equal to 0 and the constraint rank(\( L \)) = \( n-k \) will be satisfied. According to Ky Fan’s Theorem [19], \( \sum_{i=1}^k \Theta_i \) is equal to minimizing \( \text{Tr}(F^T L F) \) subject to \( F^T F = I \). Then Equation (7) can be rewritten as:

\[
\begin{align*}
\min \| Z_0 - W \|_F^2 + \lambda_1 \sum_{i=1}^r (\| P_i \|_* + \| Z_i \|_1) \\
+ \lambda_2 \sum_{i=1}^r \text{tr} (Z_i^T Z_0) + \lambda_2 \sum_{i=1}^r \sum_{i \neq j} \text{tr} (Z_i^T Z_j) \\
+ \lambda_3 \text{Tr} (F^T L F) + \sum_{i=1}^r \| E_i \|_{2,1}\end{align*}
\]

s.t. \( X_j = X_j P_j + E_j, P_j = Z_j + Z_0, F^T F = I, \ W_{ij} \geq 0, \sum_j W_{ij} = 1. \)

We introduce auxiliary variables \( U_j = P_j \) to make the objective function become separable and can be easily solved. Then, we use the Augmented Lagrange Multiplier (ALM) [34] to obtain the following augmented Lagrangian function.

\[
\begin{align*}
\min_{E_i, P_j, Z_0, U_j} \| Z_0 - W \|_F^2 + \lambda_1 \sum_{i=1}^r (\| U_i \|_* + \| Z_i \|_1) \\
+ \lambda_2 \sum_{i=1}^r \text{tr} (Z_i^T Z_0) + \lambda_2 \sum_{i=1}^r \sum_{i \neq j} \text{tr} (Z_i^T Z_j) + \lambda_3 \text{Tr} (F^T L F) \\
+ \sum_{i=1}^r \| E_i \|_{2,1} + \sum_{i=1}^r \langle H_i, X_j - X_j P_j - E_i \rangle + \sum_{i=1}^r \langle P_i - U_i \rangle \\
+ \sum_{i=1}^r \langle K_i, P_j - Z_j - Z_0 \rangle + \frac{\mu}{2} \sum_{i=1}^r (\| X_j - X_j P_j - E_i \|_F^2) \\
+ \frac{\mu}{2} \sum_{i=1}^r (\| P_i - U_j \|_2^2 + \| P_i - Z_i - Z_0 \|_F^2). \quad \text{(9)}
\end{align*}
\]

**Step 1:** Solving \( U \)—We update the variable \( U \) by fixing the other variables and dropping the unconcerned terms. Then, we have the following subproblem with respect to the variable \( U \):

\[
\begin{align*}
\min_U \lambda_1 \| U_i \|_* + \langle H_i, P_i - U_i \rangle + \frac{\mu}{2} \| P_i - U_i \|_F^2 \\
= \lambda_1 \| U_i \|_* + \frac{\mu}{2} \| U_i - \left( P_i + \frac{H_i}{\mu} \right) \|_F^2. \quad \text{(10)}
\end{align*}
\]

We solve the subproblem in Equation (10) by using Singular Value Thresholding [35] to get the closed-form solution as

\[
U_j = S_{\lambda j/\mu} \left[ D_j + \frac{H_j}{\mu} \right], \quad \text{(11)}
\]

where \( S_{\cdot} [\cdot] \) is a shrinkage thresholding operator which is defined by

\[
S_{\lambda} [M] = U H_{\mu} [\Sigma] V^T, \quad \text{H}_{\mu} [\Sigma] = \min(0, \Sigma + \mu) + \max(0, \Sigma - \mu). \quad \text{(12)}
\]

**Step 2:** Solving \( E \)—Similar to the \( U \) subproblem, the optimization problem with respect to \( E \) can be formulated as:

\[
\min_{E_i} \lambda_1 \| E_i \|_{2,1} + \frac{\mu}{2} \| E_i - \left( X_j - X_j P_j + \frac{H_j}{\mu} \right) \|_F^2. \quad \text{(13)}
\]

Based on lemma 3.3 in [36], we have the optimal solution as follows:

\[
[E_i]_{:,k} = \begin{cases} \| [A_i]_{:,k} \|_2 - 1/\mu & \text{if } [A_i]_{:,k} > 1/\mu \\
0, & \text{else} \end{cases}
\]

\( A_i = X_j - X_j P_j + \frac{H_j}{\mu} \) and \([A_i]_{:,k}\) is the \( k \)-th column of \( A_i \).

**Step 3:** Solving \( P_i \)—Fixing the other variables, the optimization problem concerned with \( P \) can be rewritten as:

\[
\min_P \langle H_i, X_j - X_j P_j - E_i \rangle + \langle H_i, P_j - U_i \rangle \\
+ \langle K_i, P_j - Z_j - Z_0 \rangle + \frac{\mu}{2} \| X_j - X_j P_j - E_i \|_F^2 \\
+ \frac{\mu}{2} \| P_j - U_i \|_2^2 + \| P_j - Z_j - Z_0 \|_F^2. \quad \text{(14)}
\]

With respect to \( P_j \) and setting the derivative of above, we obtain the following solution:

\[
P_j = (X_j^T X_j + 2I)^{-1} \left( X_j^T \left( X_j + \frac{Y_j}{\mu} - E_i \right) + Z_j + Z_0 \right) \\
+ \frac{K_j}{\mu} + U_j - \frac{H_j}{\mu}. \quad \text{(15)}
\]
The solution in Equation (17) is obtained as \( \zeta = H_{\lambda_k} (-a/\mu) \).

Step 5: Solving \( Z_0 \)—By fixing the other variables and removing the irrelevant terms, we rewrite the subproblem as:

\[
\min_{Z_0} \left\| Z_0 - W \right\|_F^2 + \lambda_2 \sum_{i=1}^{r} r(Z_i^T Z_0) \\
+ \sum_{i=1}^{r} (K_i P_i - Z_i - Z_0) + \frac{\mu}{2} \sum_{i=1}^{r} \left\| P_i - Z_i - Z_0 \right\|_F^2.
\] (18)

Setting the derivative of the above with respect to \( Z_0 \) to zero we obtain the following solution:

\[
Z_0 = (2I + 2\lambda_2 L + \mu J)^{-1} \times \left( 2W + \mu \sum_{i=1}^{r} (P_i - K_i) - \lambda_3 \sum_{i=1}^{r} Z_i \right).
\] (19)

Step 6: Solving \( W \)—We update \( W \) by solving the following problem:

\[
\Gamma(W) = \left\| Z_0 - W \right\|_F^2 + \lambda_3 \text{Tr}(F^T LF) \\
s.t., W_{ij} \geq 0, \sum_j W_{ij} = 1.
\] (20)

The problem Equation (20) is independent for different \( i \), therefore we can solve the following problem for each \( i \).

\[
\Gamma(w_{ij}) = \sum_j (Z_{0ij} - w_{ij})^2 + \lambda_3 \sum_j \left\| f_i - f_j \right\|_2^2 w_{ij} \\
s.t., w_{ij} \geq 0, \sum_j w_{ij} = 1.
\] (21)

Denoting \( t_{ij} = \left\| f_i - f_j \right\|_2^2 \), and \( f_i \) as vector with the \( i^{th} \) element equal to \( t_{ij} \), Equation (21) above can be written as [37, 38]:

\[
\min_{w_{ij} \geq 0, 1} \left\| w_i - \left( \frac{\lambda_3}{2} t_i \right) \right\|_2^2.
\] (22)

The Lagrangian function of Equation (22) is:

\[
\Gamma = \frac{1}{2} \left\| w_i - \left( \frac{\lambda_3}{2} t_i \right) \right\|_2^2 - \eta (w_i^T 1 - 1) - \beta_i^T w_i,
\] (23)

where \( \eta \) and \( \beta_i \) are the Lagrangian multipliers. By taking the derivative of Equation (23) with respect to \( w_i \) and setting to zero, we can obtain:

\[
w_i - \left( \frac{\lambda_3}{2} t_i \right) - \eta_1 - \beta_i = 0.
\] (24)

According to the KTT condition, the optimal solution \( w_i \) is

\[
w_{ij} = \left( \frac{\lambda_3}{2} t_{ij} + \eta \right) +.
\] (25)

We use Euclidean Projection onto the Simplex [39] to solve Equation (25). Finally, the whole algorithm is presented in Algorithm 1.

### 3.5 Model complexity

In this subsection, we analyze the complexity of our proposed model. The first step is to solve the problem in Equation (10) which utilizes the shrinkage threshold operator. The time complexity for updating \( U_i \) is \( O(n^2) \). For the second sub-problem \( E_i \), the time complexity is \( O(n^2 d/d) \). The sub-problems \( P_i \) and \( Z_0 \) need matrix inversion and their time complexities of these two sub-problems are \( O(n^3) \). The last problem of \( W \) is a Euclidean projection problem on the simplex space. We need \( O(n) \) to compute. The time complexity of sub-problem \( Z_i \) and the update process of the multipliers and \( \mu \) can be omitted in comparison with the other sub-problems. For \( v \) different views, the total complexity of our method in each iteration is \( O(n^2 d + 2n^3 + n) \), where \( d = d_1 + \cdots + d_v \).

### 4 EXPERIMENTS

In this section, we first describe three face datasets, six similar state-of-the-art compared methods, and general experimental settings. Secondly, we show experimental results with analysis for face recognition and clustering. The final clustering labels can be obtained directly from the affinity matrix \( W \) obtained from Algorithm 1 because each connected component belongs to one cluster, according to the strongly connected component algorithm.

#### 4.1 Datasets

- **ORL\(^1\)**: This face dataset contains 400 images of 40 distinct subjects. For each category, images were taken at different views for different people with different lighting, facial expressions, and facial details. We use the labeled ORL face dataset for experiments.

\(^1\) http://cam-orl.co.uk/facedatabase.html
TABLE 2  Dataset information

| Dataset     | Size | View | Classes |
|-------------|------|------|---------|
| ORL         | 400  | 3    | 40      |
| Extended Yale B | 2414 | 3    | 38      |
| AR          | 4000 | 3    | 126     |

*Figure 2: Example pictures of ORL and Yale B faces datasets*

- ORL: This dataset contains 400 frontal-face images of 40 individuals. Each class has 10 images.
- Extended Yale B: This dataset contains 2414 frontal-face images of 38 individuals. The images were captured under various laboratory-controlled lighting conditions. Each class has 64 images.
- AR: This database consists of over 4000 frontal images from 126 individuals. Specifically, there are 26 face images available for each person, with images taken under different variations, including illumination, expression, and facial occlusion/disguise within two separate sessions.

For all three datasets, we extract three types of features to represent different views. These features are Intensity, LBP, and Gabor. See Table 2 for a summary of the datasets, while Figure 2 gives example images of ORL and Extended Yale B datasets.

### 4.2 Compared methods

We compare our method with six similar state-of-the-art ones, including Diversity-Induced Multi-View Subspace Clustering (DiMSC) [21], Multi-View Low-Rank Sparse Subspace Clustering (MLRSSC) [27], Co-regularized multi-view spectral clustering (CoReg) [4], Nuclear Norm Based Matrix Regression (NMR) [26], Iterative Re-Constrained Group Sparse Face Recognition (IRGSC) [40], and RRC [41]. The following is a brief description of the different methods. For some methods that mainly focus on representation learning, such as NMR, IRGSC, and RRC, we used k-means on their optimal data representation to obtain clustering labels.

- DiMSC: This method utilizes the Hilbert Schmidt Independence Criterion as a diversity term to enhance the diversity between views.
- MLRSSC: This method combines the idea of SSC and LRR to find a joint subspace representation.
- CoReg: This method maximizes the agreement between different views by regularizing the eigenvectors of specific graph Laplacian matrices.
- NMR: This method uses the minimal nuclear norm of representation error image as a criterion to find an ideal face representation.
- IRGSC: This method uses a group sparse representation classification (GSRC) approach where the weighted features and groups are captured cooperatively so as to encode more discriminative information and structure information.
- RRC: This method uses a unified sparse weight learning approach to allow random noises and occlusions in images to be resolved simultaneously.

### 4.3 Experimental settings

Without loss of generality, we downloaded the source code from the authors’ website and followed the experimental settings and the parameter tuning steps described in their paper. We ran each method ten times, and the mean performance is reported. The tuned parameters from corresponding papers range from 0.001, 0.01, 0.1, 1, 10, 100, 1000 and the best parameters are fine-tuned using the grid-search strategy. In our experiments, the parameter $\mu$ is a Lagrangian penalty variable. Referring to [42], we initialize $\mu$ by 0.01 and set its maximum value as $10^8$. For the parameter $\rho$, we set its value as 1.1. According to parameters analysis in subsection 4.6, the parameters of proposed method are set as follow: $\lambda_1 = 0.01$, $\lambda_2 = 10$, $\lambda_3 = 1$. We initialize all the matrices with a zero matrix except $E_i$ with a sparse matrix. Besides, we performed the experiments for NMR, IRGSC, and RRC methods separately on each view, and the results for the view with the best performance is reported.

In experiments for face recognition, we used 60% of the data point as the training set ($X$) and 40% as the testing set ($X_{\text{text}}$). Then for each of DiMSC, and MLRSSC methods, we obtained representation matrices for $X$ and $X_{\text{text}}$, respectively, while the intrinsic low-rank representation matrix $Z_0$ of $X$ and ($Z_{\text{text}}$) of $X_{\text{text}}$ in our experiment is obtained by solving Equation (6). Finally, we used the support vector machine (SVM) algorithm [43] to measure the classification performance of our method, DiMSC, and MLRSSC. Note that steps 2 and 3 above are integrated into NMR, IRGSC, and RRC algorithms because they are designed mainly for face recognition.

### 4.4 Face recognition

In this section, we evaluate the recognition accuracy of the different state-of-the-art methods. First, we look at the recognition effect of various embedding dimensionalities in Section 4.4.1. Then, in Section 4.4.2, we investigate the robustness of each method to noise and occlusion.

2 http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html
3 https://www2.ece.ohio-state.edu/~aleix/ARdatabase.html
4.4.1 Recognition rate on different embedding dimension

Here, six instances of experiments are performed on each of ORL and Extended Yale B datasets, while seven instances of experiments were performed on the AR dataset. In other words, we explored different embedding dimensionalities in all three datasets to evaluate the face recognition accuracy of the different methods, and we show the performances of each method in Figure 3 and Tables 3–Table 5 in which the bold numbers denote the best results, respectively. Hence, we see clearly that all different methods have significant improvement by increasing the dimensionality in all three datasets. For example, in Table 3, our method has an accuracy rate of 82.32% and 91.28% with a dimension of 25 and 50, respectively. However, by projecting all data points from 50 to 200 dimensions, the performance improved significantly by 5.04%, which is better than RRC, DiMSC, and NMR, which has 4.7%, 4.5%, and 4.3%, respectively. Furthermore, we can observe that the single view methods RRC, NMR, and IRGSC have relatively better performances than DiMSC and MLRSSC, which are multi-view methods. The reason is easy to establish since the trio of RRC, NMR, and IRGSC methods are designed solely for face recognition.

---

**TABLE 3**  
Face recognition accuracy of different methods on ORL dataset using different embedding dimensionalities

| Method  | 25   | 50   | 75   | 100  | 150  | 200  |
|---------|------|------|------|------|------|------|
| NMR     | 90.52| 94.51| 95.12| 95.33| 96.07| 96.27|
| DiMSC   | 84.58| 91.11| 92.77| 93.83| 94.41| 94.5 |
| MLRSSC  | 85.83| 90.44| 92.53| 93.43| 93.99| 94.08|
| IRGSC   | 87.45| 92.58| 93.51| 93.96| 94.97| 95.16|
| RRC     | 88.25| 93.12| 93.91| 94.36| 95.29| 95.76|
| Proposed| 89.12| 94.12| 95.23| 95.89| 96.36| 96.41|

---

**TABLE 4**  
Face recognition accuracy of different methods on Extended Yale B dataset using different embedding dimensionalities

| Method  | 25   | 50   | 75   | 100  | 200  | 300  |
|---------|------|------|------|------|------|------|
| NMR     | 63.31| 69.16| 74.78| 77.84| 80.29| 80.84|
| DiMSC   | 60.21| 64.63| 69.21| 73.18| 76.83| 77.21|
| MLRSSC  | 61.51| 65.21| 70.56| 74.73| 77.15| 77.91|
| IRGSC   | 66.83| 69.59| 73.39| 74.81| 77.88| 78.88|
| RRC     | 68.37| 71.02| 73.88| 75.83| 78.58| 79.32|
| Proposed| 62.15| 75.13| 79.83| 82.37| 82.68| 82.85|

---

**TABLE 5**  
Face recognition accuracy of different methods on AR dataset using different embedding dimensionalities

| Method  | 25   | 50   | 75   | 100  | 200  | 300  | 500  |
|---------|------|------|------|------|------|------|------|
| NMR     | 63.31| 69.16| 74.78| 77.84| 80.29| 80.84| 81.05|
| DiMSC   | 60.21| 64.63| 69.21| 73.18| 76.83| 77.21| 77.33|
| MLRSSC  | 61.51| 65.21| 70.56| 74.73| 77.15| 77.91| 78.11|
| IRGSC   | 66.83| 69.59| 73.39| 74.81| 77.88| 78.88| 79.19|
| RRC     | 68.37| 71.02| 73.88| 75.83| 78.58| 79.32| 79.81|
| Proposed| 62.15| 75.13| 79.83| 82.37| 82.68| 82.85| 82.95|
Algorithm 1 Algorithm of our proposed method

1: Input: The Multi-view matrix $X$, parameters $\lambda_1, \lambda_2, \lambda_3$, Cluster number $k$
2: Initialize $H_i = 0, J_i = 0, K_i = 0$; $Z$ and $W$ are similarity matrices based on $k$ nearest neighbor graph; $F$ is formed by the $k$ eigenvectors of $L$ corresponding to the $k$ smallest eigenvalues; $\rho = 1.01$.
3: while not converged do
4:   Update $U_i$ by Equation (10).
5:   Update $E_i$ by Equation (13).
6:   Update $P_i$ by Equation (15).
7:   Update $Z_i$ by Equation (17).
8:   Update $Z_0$ by Equation (19).
9:   Update $W$ by Equation (25).
10:  $H_i = H_i + \mu(X_i - X_i P_i - E_i)$
11:  $J_i = J_i + \mu(P_i - U_i)$
12:  $K_i = K_i + \mu(P_i - Z_i - Z_0)$
13:  Update $\mu$ by $\mu = \min(\rho \mu, \max(\mu))$.
14: end while
15: Output: $W$, $Z_0$, $Z_i$, $P_i$.

Table 6 Face recognition accuracy of different methods on ORL dataset with different levels of noise

| Method   | 0%  | 10% | 20% | 30%  | 40%  | 50%  |
|----------|-----|-----|-----|------|------|------|
| NMR      | 94.77 | 81.69 | 73.13 | 61.98 | 50.35 | 44.66 |
| DiMSC    | 95.73 | 80.32 | 69.93 | 60.16 | 49.15 | 43.16 |
| MLRSSC   | 94.41 | 76.25 | 64.32 | 54.63 | 45.42 | 37.83 |
| IRGSC    | 93.09 | 79.03 | 70.13 | 59.03 | 47.92 | 39.79 |
| RRC      | 95.56 | 78.04 | 68.27 | 54.91 | 46.65 | 39.18 |
| Proposed | 96.32 | 83.54 | 74.32 | 65.33 | 53.18 | 46.32 |

4.4.2 Robustness to Noise and Occlusion

To investigate the robustness of different methods to noise and occlusion, we use two separate experimental settings. Firstly, by randomly injecting different levels of pixel corruption into the testing data of ORL. And secondly, by randomly injecting different levels of block occlusions into that of Extended Yale B. Therefore, Tables 6 and 7 display the face recognition accuracy of different methods in which the bold numbers denote the best results. We can see that our method significantly outperforms most of the compared methods with the increasing level of pixel corruption or block occlusion. Notably, in Table 7, we can observe that our method only outperforms NMR and RRC by 0.58% and 0.62%, respectively, with a 20% random block occlusion. However, by increasing the level to 70%, our method shows more robustness than NMR and RRC with a difference of 6.4% and 22.46%, respectively. Besides, one may notice equally in Table 5 that our method only performs better than DiMSC by 0.59% with 0% noise corruption. Yet, our method outperforms DiMSC more significantly by increasing the level of pixel corruption. Overall, Tables 6 and 7, with the aid of Figure 4, show clearly that our proposed method is generally more robust to noise and occlusion than the compared state-of-the-art methods. This further buttress the effectiveness of our intrinsic low-rank representation.

Table 7 Face recognition accuracy of different methods on Extended Yale B dataset with different levels of occlusion

| Method   | 20% | 30% | 40% | 50% | 60% | 70% |
|----------|-----|-----|-----|-----|-----|-----|
| NMR      | 95.73 | 93.16 | 92.63 | 91.32 | 78.93 | 55.24 |
| DiMSC    | 94.41 | 92.16 | 90.61 | 86.16 | 70.93 | 43.16 |
| MLRSSC   | 94.11 | 91.22 | 88.31 | 87.32 | 65.19 | 37.83 |
| IRGSC    | 94.42 | 91.13 | 89.29 | 88.51 | 67.19 | 39.79 |
| RRC      | 95.69 | 90.16 | 86.13 | 86.13 | 64.69 | 39.18 |
| Proposed | 96.31 | 94.16 | 93.61 | 92.16 | 81.35 | 61.64 |

4.5 Face clustering

We utilize six kinds of evaluation metrics to verify the effectiveness of the proposed method, which are clustering accuracy...
WANG ET AL.

### TABLE 8 Clustering performance on ORL dataset

| Method  | ACC   | NMI   | AR    | F-score | Precision | Recall |
|---------|-------|-------|-------|---------|-----------|--------|
| DiMSC   | 78.87 | 91.19 | 73.55 | 74.19   | 69.48     | 79.61  |
| MLRSSC  | 71.95 | 87.22 | 64.18 | 65.06   | 60.25     | 70.81  |
| CoReg   | 70.51 | 85.32 | 60.25 | 61.57   | 56.74     | 66.7   |
| NMR     | 76.56 | 85.45 | 72.22 | 73.19   | 70.02     | 79.26  |
| IRGSC   | 72.87 | 86.08 | 70.38 | 73.49   | 66.77     | 76.47  |
| RRC     | 76.44 | 85.16 | 71.88 | 72.75   | 69.93     | 79.09  |
| Proposed| **80.51** | **92.41** | **76.14** | **74.51** | **70.43** | **82.14** |

### TABLE 9 Clustering performance on Extended Yale B dataset

| Method  | ACC   | NMI   | AR    | F-score | Precision | Recall |
|---------|-------|-------|-------|---------|-----------|--------|
| DiMSC   | 66.06 | 67.34 | 47.57 | 50.87   | 49.36     | 52.48  |
| MLRSSC  | 64.31 | 68.18 | 47.59 | 50.94   | 48.68     | 53.46  |
| CoReg   | 56.42 | 64.82 | 43.64 | 46.68   | 45.54     | 49.13  |
| NMR     | 67.65 | 67.86 | 45.65 | 49.08   | 43.67     | 57.39  |
| IRGSC   | 66.65 | 67.75 | 45.62 | 48.94   | 43.59     | 56.58  |
| RRC     | 67.62 | 66.29 | 44.61 | 49.01   | 44.12     | 57.12  |
| Proposed| **68.72** | **71.54** | **49.67** | **51.06** | **50.41** | **58.41** |

### TABLE 10 Clustering performance on AR dataset

| Method  | ACC   | NMI   | AR    | F-score | Precision | Recall |
|---------|-------|-------|-------|---------|-----------|--------|
| DiMSC   | 75.34 | 72.39 | 56.23 | 57.23   | 58.16     | 58.88  |
| MLRSSC  | 72.06 | 74.04 | 55.52 | 55.57   | 49.92     | 56.34  |
| CoReg   | 64.69 | 66.18 | 51.97 | 50.27   | 50.35     | 55.41  |
| NMR     | 68.7  | 73.01 | 55.41 | 57.95   | 43.59     | 61.96  |
| IRGSC   | 69.36 | 76.47 | 53.55 | 56.86   | 47.83     | 67.24  |
| RRC     | 75.25 | 72.17 | 47.09 | 52.05   | 46.23     | 57.2   |
| Proposed| **78.08** | **76.01** | **65.07** | **68.72** | **59.97** | **61.08** |

(ACC), normalized mutual information (NMI), adjusted rand index (AR), F-score, Precision, and Recall. And these metrics are calculated by comparing the obtained label of each sample with the ground-truth labels provided in datasets, where a larger value will indicate the better clustering performance. Specifically, we conducted experiments on three widely used face datasets to evaluate different methods’ clustering performance. The results are shown in Tables 8–10 in which the bold numbers denote the best results. Table 8 shows the clustering results on ORL dataset, and one can observe that our method outperforms most of the compared methods by over 4% in ACC, 5% in NMI, and 4% in AR. In particular, the performance improvements over DiMSC on the ORL dataset in terms of ACC, NMI, and F-score are 1.64%, 1.22%, 0.32%, respectively. The reason can be attributed to the large variation of illumination of this dataset. However, it can be seen clearly that our approach still achieves significant improvements, which are better when compared to other methods. And the quantitative result fully demonstrates the superiority of our approach since it better captures the general low-rank structure of the data space. Moreover, the proposed method can obtain a better graph from the general low-rank structure, so, the results are better than others. For instance, the face images which belong to the same track are remarkably placed similar to each other which implies significant low-rank property. Therefore, the proposed method under the low-rank assumption and integrated ideal global graph guarantees better results. Table 9 shows the clustering results on Extended Yale B dataset, yet, our method still outperforms all the alternative methods. For instance, our method outperforms MLRSSC significantly with above 4% in terms of accuracy and Recall. In fact, on this dataset, the results show not only the influence of local connectivity in constructing the similarity graph but also the superiority of the unified optimization. Table 10 shows the clustering results on AR dataset. On this dataset, some compared methods achieve promising performance. For example, RRC, which is a single-view method has better performance in ACC than CoReg and MLRSSC, which finds complementary information in multi-view face data in a shared affinity matrix of all views. Furthermore, we observe that DiMSC has a better performance than CoReg and MLRSSC as well. This is not unusual because DiMSC enhances the diversity among different views. Notwithstanding, our method improves DiMSC by over 2%, 3%, 10% in terms of ACC, NMI, and F-score, respectively, which demonstrates the importance of combined general low-rank structure and clustering structure.

### 4.6 Ablation experiments

We also conducted ablation experiment on single-view and multi-view datasets to demonstrate the performance of proposed method. In this experiment, we utilize the first view of each dataset as the only input data matrix. The results are shown in Table 11 in which the bold numbers denote the best results. From Table 11, we can observe that results on all multi-view datasets outperform results on single-view dataset, respectively. Especially in ACC on multi-view AR dataset, which improves single-view dataset by over 4%. This further buttress the effectiveness of our intrinsic low-rank representation.

### 4.7 Parameter analysis

We analyze the sensitivity of parameters in Figures 5 and 6 to show the importance of the parameters $\lambda_1$, $\lambda_2$ and $\lambda_3$. The base
of the logarithm is 10. In Figure 5, we show the NMI and ACC results on the Extended Yale B dataset when $\lambda_3$ is fixed to 1. Figure 6 shows the results when $\lambda_2$ is fixed to 1. In particular, these figures show that our method is not sensitive to $\lambda_2$ when $\lambda_1$ and $\lambda_3$ is fixed. Besides, we can observe that $\lambda_1$ plays an important role in improving the performance. And it is not surprising since $\lambda_1$ controls the learning of low-rank affinity matrix $P_i$. Hence, a suitable $\lambda_1$ tends to achieve a better performance, which gives a good explanation for these figures.

4.8 Convergence analysis

The Augmented Lagrange Multiplier method was used to update all the matrices in our optimization problem iteratively, and the convexity of the Lagrange function guarantees the effectiveness of our proposed method to some degree. Figure 7 shows the objective function values on the ORL and AR datasets. It can be seen that on both real-world datasets, the algorithm converges quickly within 20 iterations.

5 CONCLUSION

In this paper, we proposed a novel multi-view low-rank representation method. Our method follows hierarchical Bayesian methods and learns intrinsic and specific representation among each view via a consistent structure and diversity regularization. The intrinsic representation holds the consistent information between different views while the specific representation unfolds the diversity information across views. Meanwhile, our model combines the clustering structure and low-rank local manifold to learn both the relationship in data and the clustering structure. Furthermore, by adjusting the clustering structure adaptively during optimization, our method can achieve a better performance on face recognition and clustering simultaneously. Several experiments on face datasets, in comparison with the state-of-the-art algorithms, show the effectiveness of our method. In the future, we will try to extend this method to handle incomplete face data.

ACKNOWLEDGEMENTS

This research was funded in part by the National Key Research & Development Program of China, Grant No. 2020YFC1511800, Primary Research & Development Plan of Jiangsu Province (BE201862), and Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX20_1417).

ORCID

Zhi-yang Wang https://orcid.org/0000-0001-6471-7797

REFERENCES

1. Li, Y., et al.: Support vector machine based multi-view face detection and recognition. Image Vis. Comput. 22(5), 413–427 (2004)
2. Fukui, K., Yamaguchi, O.: Face recognition using multi-viewpoint patterns for robot vision. In: Robotics Research: The Eleventh International Symposium, pp. 192-201. Springer, Berlin, Heidelberg (2005)
3. Rattani, A., Tistarelli, M.: Robust multi-modal and multi-unit feature level fusion of face and iris biometrics. In: International Conference on Biometrics, pp. 960-969. Springer, Berlin, Heidelberg (2009)
4. Kumar, A., et al.: Co-regularized multi-view spectral clustering. In: Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, NIPS 2011. ACM, New York (2011)
5. Xia, R., et al.: Robust multi-view spectral clustering via low-rank and sparse decomposition. In: Proceedings of the National Conference on Artificial Intelligence. AAAI Press, Palo Alto (2014)
6. Huang, L., et al.: Co-learned multi-view spectral clustering for face recognition based on image sets. IEEE Signal Process. Lett. 21(7), 875–879 (2014)
7. Cao, X., et al.: Constrained multi-view video face clustering. In: IEEE Trans. Image Process. 24(11), 4381–4393 (2015). https://doi.org/10.1109/TIP.2015.2463223
8. Li, Y., et al.: Large-scale multi-view spectral clustering via bipartite graph. In: Proceedings of the National Conference on Artificial Intelligence, vol. 4, pp. 2750-2756. AAAI Press, Palo Alto (2015)
9. Von Luxburg, U.: A tutorial on spectral clustering. Stat. Comput. (2007), https://doi.org/10.1007/s11222-007-9033-z
10. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. (2000), https://doi.org/10.1109/34.868688
11. Ng, A.Y., et al.: On spectral clustering: Analysis and an algorithm. In: Advances in Neural Information Processing Systems. ACM, New York (2002)
12. Liu, G., et al.: Robust subspace segmentation by low-rank representation. In: ICML 2010 - Proceedings, 27th International Conference on Machine Learning. IEEE, Piscataway (2010)
13. Elhamifar, E., Vidal, R.: Sparse subspace clustering. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 2790-2797. IEEE, Piscataway (2009)
14. Liu, G., et al.: Robust recovery of subspace structures by low-rank representation. IEEE Trans. Pattern Anal. Mach. Intell. 35(1), 171–184 (2013)
15. Tang, K., et al.: Structure-constrained low-rank representation. In: IEEE Transactions on Neural Networks and Learning Systems. IEEE, Piscataway (2014), https://doi.org/10.1109/TNNLS.2014.2306063
16. Zhang, H., et al.: Robust latent low rank representation for subspace clustering. Neurocomputing 145, 369-373 (2014)
17. Ding, Z., Fu, Y.: Robust multi-view subspace learning through dual low-rank decompositions. 30th AAAI Conference on Artificial Intelligence, AAAI 2016, pp. 1181–1187. AAAI Press, Palo Alto (2016)
18. Jing, X.Y., et al.: Multi-spectral low-rank structured dictionary learning for face recognition. Pattern Recognit. 59, 14-25 (2016)
19. Qian, Y., et al.: Low-rank graph optimization for multi-view dimensionality reduction. PLOS One 14(12), e0225987 (2019)
20. Wang, Z.-y., et al.: Multi-view Low-Rank Preserving Embedding: A novel method for multi-view representation. Eng. Appl. Artif. Intell. 99, 104140 (2020)
21. Cao, X., et al.: Diversity-induced Multi-view Subspace Clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2790-2797. IEEE, Piscataway (2009)
22. Allenby, G.M., Rossi, P.: Marketing models of consumer heterogeneity. Journal of Econometrics 89(1-2), 57–78 (1998)
23. Arora, N., et al.: A hierarchical Bayes model of primary and secondary demand. Marketing Science 17(1), 29–44 (1998), https://doi.org/10.1287/mkse.17.1.29
24. Heskes, T.: Empirical Bayes for Learning to Learn. In: Proceedings of the Seventeenth International Conference on Machine Learning, pp. 367–374. IEEE, Piscataway (2000)
25. Gao, H., et al.: Multi-view subspace clustering. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4238-4246. IEEE, Piscataway (2015)
26. Yang, J., et al.: Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes. IEEE Trans. Pattern Anal. Mach. Intell. 39(1), 156-171, 2016
27. Brbic, M., Kopriva, I.: Multi-view low-rank sparse subspace clustering. Pattern Recognit. 73, 247-258 (2018), https://doi.org/10.1016/j.patcog.2018.08.024
28. Xu, J., et al.: Robust Multi-view Subspace Learning Through Structured Low-Rank Matrix Recovery. In: Lai, J.H., et al. (eds.) Pattern Recognition and Computer Vision. PRCV 2018. Lecture Notes in Computer Science, vol 11258. Springer, Cham
29. Hu, Y., et al.: Fast and accurate matrix completion via truncated nuclear norm regularization. IEEE Trans. Pattern Anal. Mach. Intell. 35(9), 2117-2130 (2012)
30. Zheng, C.H., et al.: Improved sparse representation with low-rank representation for robust face recognition. Neurocomputing 198, 114–124 (2016)
31. Meng, X., et al.: Multi-view Low-Rank Preserving Embedding: A novel method for multi-view representation. Eng. Appl. Artif. Intell. 99, 104140 (2020)
32. Lin, Z., et al.: The augmented Lagrange multiplier method for exact recovery of corrupted low-Rank matrices (2010), https://doi.org/10.1016/j.jsp.2012.10.010
33. Boyd, S., et al.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning (2010), https://doi.org/10.1561/2200000016
34. Fan, K.: On a Theorem of Weyl Concerning Eigenvalues of Linear Transformations. Proc. Natl. Acad. Sci. USA 36(1), 31–35 (1950)
35. Cai, JF.: A singular value thresholding algorithm for matrix completion, SIAM J. Optim. (2010), https://doi.org/10.1137/080738970
36. Yang, J., et al.: A fast algorithm for edge-preserving variational multi-channel image restoration. SIAM J. Imag. Sci. 2(2), 569–592 (2009), https://doi.org/10.1137/080730421
37. Wen, W., et al.: Incomplete multiview spectral clustering with adaptive graph learning. In: IEEE Transactions on Cybernetics, 50(4), 1418–1429 (2018), https://doi.org/10.1109/TCSY.2018.2884715
38. Duchi, J., et al.: Efficient projections onto the l1-ball for learning in high dimensions. In: Proceedings of the 25th International Conference on Machine Learning, pp. 272–279. IEEE, Piscataway (2008), https://doi.org/10.1145/1390156.1390191
39. Pock, T., Sabach, S.: Inertial proximal alternating linearized minimization (iPALM) for nonconvex and nonsmooth problems. SIAM J. Imag. Sci. 9(4), 1756–1786 (2016), https://doi.org/10.1137/15M1060646
40. Zhang, J., et al.: Iterative re-constrained group sparse face recognition with adaptive weights learning. In: IEEE Trans. Image Process. 26(8), 2408–2423 (2017), https://doi.org/10.1109/TIP.2017.2681841
41. Yang, M., et al.: Regularized robust coding for face recognition. IEEE Trans. Image Process. 22(5), 1753–1766 (2012)
42. Wang, X., et al.: Multi-view subspace clustering with inexactness-aware similarity. Pattern Recognit. 88, 50–63 (2019), https://doi.org/10.1016/j.patcog.2018.09.009
43. Weston, J., Watkins, C.: Multi-class support vector machines, pp. 98–04. Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London (1998)