Journal of Computers Vol. 33 No. 4, August 2022, pp. 121-131
doi: 10.53106/199115992022083304010

Multi-view Re-weighted Sparse Subspace Clustering with Intact Low-rank Space Learning

Hengdong Zhu[1, 2], Ting Yang[1], Yingcang Ma[1*], Xiaofei Yang[1]

[1] School of Science, Xi'an Polytechnic University, Shaanxi Xian 710600, China
ZhuHengDong1997@163.com, mayingcang@126.com, 15929121393@126.com, yangxiaofei2002@163.com
[2] Hunan Transportation Engineering Institute, Hunan Hengyang 421001, China

Received 21 September 2021; Revised 2 February 2022; Accepted 2 March 2022

Abstract. In this paper, we propose a new Multi-view Re-weighted Sparse Subspace Clustering with Intact Low-rank Space Learning (ILrS-MRSSC) method, trying to find a sparse representation of the complete space of information. Specifically, this method integrates the complementary information inherent in multiple angles of the data, learns a complete space of potential low-rank representation, and constructs a sparse information matrix to reconstruct the data. The correlation between multi-view learning and subspace clustering is strengthened to the greatest extent, so that the subspace representation is more intuitive and accurate. The optimal solution of the model is solved by the augmented lagrangian multiplier (ALM) method of alternating direction minimal. Experiments on multiple benchmark data sets verify the effectiveness of this method.

Keywords: intact space learning, low-rank, multi-view, subspace clustering

1 Introduction

The increasing development of information technology has caused rapid growth of data on the internet, and the complexity of large-scale data has also brought certain difficulties to data analysis. Subspace clustering is an effective means to achieve large-scale data clustering, and is widely used in machine learning, computer vision, image processing and other fields [1-4].

The main idea of the subspace clustering method is to assume that high-dimensional data can be linearly represented by multiple independent low-dimensional subspace data, and find the coefficient representation matrix to achieve subspace segmentation. For example, Sparse Subspace Clustering (SSC) [5] analyzes the similarity relationship of each data point individually, imposes $l_1$norm constraints on the optimization objective, and obtains a sparse coefficient matrix. Low Rank Representation (LRR) [6] imposes a nuclear norm constraint on the optimization objective, aiming to jointly find the lowest rank representation of a vector set, which can better capture the global structure of the data. Non-negative Low Rank and Sparse (NNLRS) [7] adds a non-negative constraint on the coefficient matrix, and strengthens the connection between subspaces from the global perspective of the data. Lu et al [8] added a mandatory block diagonal condition for the coefficient matrix. Ming et al [9] introduced the riemannian geometric structure to enhance the anti-noise ability of the algorithm. The above methods are all unsupervised subspace methods, which only look for similar relationships between data points from the data itself. However, the data itself may carry some prior information, and the commonly used prior information includes label information [10] and pairwise constraint information [11]. Therefore, introducing a semi-supervised framework into subspace clustering and using prior information to guide the construction of the coefficient matrix can improve the subspace learning performance. Huang et al [12] introduced spatial information and label information to improve the connectivity of similarity matrix. Zhang et al [13] restricted the weight coefficient between unconnected data to 0 through pairwise constraint information, which enhanced the algorithm's anti-noise ability and robustness.

At present, the research on subspace clustering mainly focuses on the design of data regularization term, the design of fast algorithm, the introduction and practical application of semi-supervised learning and so on. In fact, for single-view data, the subspace methods described above can usually yield good clustering results. Single-view data usually describes the object from a single or fixed angle, and cannot obtain comprehensive information about the object. However, in practical applications, real data is often composed of information from different perspectives, that is, multi-view data. Unlike single view, multi-view data can describe the same object from different perspectives. For example, web page data may contain textual information, visual information, and hyperlink

* Corresponding Author

information. For example, if we photograph the same person from multiple angles, we can get image data from different angles. The front image can obtain clear face information relative to the back image or the side image. The single-view subspace method performs clustering analysis from the single-view data itself, without considering that the complete data information of the object may exist in different view data, which makes the clustering effect of the single-view subspace method on multi-view data often insignificant. In order to solve the problem of poor clustering performance of multi-view data subspace methods, the introduction of multi-view learning framework into subspace methods is a new solution in recent years.

Multi-view learning mainly explores the complementary information between different views and tries to find the underlying complete information [14]. According to the work of literature [15-17], we can know that multi-view learning can be divided into three types, including late integration, intermediate integration and early integration. Late integration is to analyze the clustering algorithm for each view separately, and then combine the results. Greene et al [18] integrated the clustering results of different views into a matrix and then passed non-negative matrix factorization to obtain the final result. Xia et al [19] recover a shared low-rank transition probability matrix from a single view. Compared with late integration, intermediate integration calculates the similarity matrix on different views and generates a fused pairwise representation, that is, using multiple single-view data to learn a common data graph information. However, there is a problem. For clustering tasks, due to the different information carried between each view, the weight coefficients assigned to different views should also be different [20-22]. Parameter Free Auto Weighted Multiple Graph Learning (AMGL) [23] automatically learns the weight coefficients of a single view by re-adjusting the standard spectrum learning mode. Self-weighted multi-view clustering (SwMC) [24] learns the similarity matrix of individual views to automatically assign weight coefficients and learn a common similarity matrix. Multi-View Clustering and Semi-Supervised Classification with Adaptive Neighbors (MLAN) [25] learns the local manifold structure of a single view and integrates all views by automatically assigning weight coefficients. Exclusivity Consistent Regularized Multi-View Subspace Clustering (ECMC) [26] introduces multi-view learning into sparse subspace methods that fully consider multi-view complementary information between.

In this research, we focus on multi-view learning based on early integration. The purpose of early integration is to find a complete representation of multi-view data, that is, to fuse each view information into a comprehensive view and perform data analysis on the comprehensive view. One of the representative works is Multi-view Intact Space Learning (MISL) [27]. MISL integrates the encoded complementary information into all views to restore the complete latent representation (as shown in Fig. 1). This is a new idea, which assumes that the multi-view data is derived from the projection data from different perspectives of the complete latent representation. Multi-view Subspace Clustering with Intactness-Aware Similarity (MSC_IAS) [28] learn an intact space by integrating encoded complementary information, which enforces the constructed similarity to have maximum dependence with its latent intact points by adopting the Hilbert Schmidt Independence Criterion (HSIC). Latent Multi-view Subspace Clustering (LMSC) [29] introduces the learned latent complete space into the subspace method, and performs data reconstruction on the basis of the latent space.

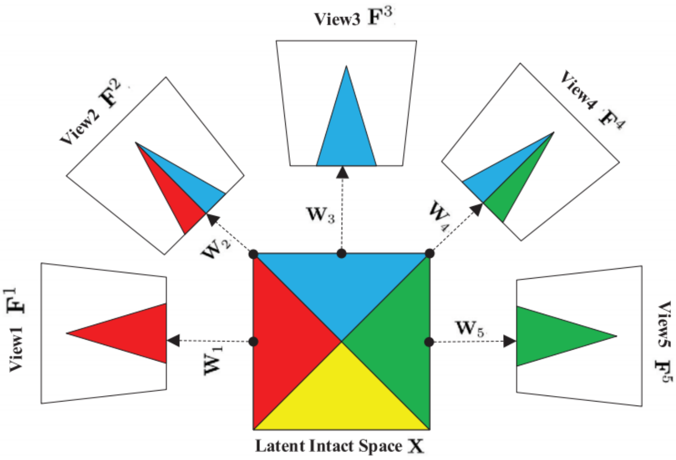


Fig. 1. Intact space learning schematic diagram [27]
(Mining a potential complete space through the complementary information of multi-view data can describe the underlying structure of multi-view data more comprehensively.)

In this paper, we propose a new model to explore the relationship between multi-view data and subspace methods. Assuming that the single-view data is the projection data of the potentially complete data under different viewing angles, there is complementary information between multiple views. Different from other methods, we believe that there may be redundant information between different views, and achieve data reconstruction by imposing different constraints, eliminating redundant data information, and simultaneously finding sparse representations of potentially complete data. Specifically, our main contributions can be summarized as follows:

(1) We propose a novel multi-view subspace clustering model, namely Multi-view Re-weighted Sparse Subspace Clustering with Intact Low-rank Space Learning. Unify subspace methods and multi-view learning in one framework, and jointly optimize to improve clustering performance.

(2) We fully consider the information correlation between multiple views. Through the low-rank property of the nuclear norm, we minimize redundant information between multiple views, and at the same time reconstruct the data through a subspace approach, ensuring that the recovered potentially complete data is clean and noise-free as much as possible.

(3) Through experiments on multiple benchmark datasets, we demonstrate the effectiveness of our proposed framework and outperform the state-of-the-art alternatives.

The rest of this paper is organized as follows. In Section 2, we introduce classical models with related backgrounds. In Section 3, we introduce the proposed Multi-view Re-weighted Sparse Subspace Clustering with Intact Low-rank Space Learning (ILrS-MRSSC) algorithm for clustering, providing a new ALM-based optimization scheme. In Section 4, we conduct experimental analysis on benchmark datasets to evaluate the performance of the algorithm. Finally, Section 5 presents the conclusion of this paper.

2 Related Work

This section aims to introduce the classic model with relevant background, mainly introduces the optimization model of complete space learning and reweighted subspace method.

2.1 Intact Space Learning

As mentioned in the work [27-28], in the complete space learning, assuming that there is a latent complete space $X \in R^{d \times n}$, the information of a single view $F^v \in R^{dv \times n}$ can be obtained through the projection matrix $W^v \in R^{dv \times d}$ of different perspectives (as shown in Fig. 1), note $v = 1, 2...V$, dv represents the dimension of the view F^v, d represents the dimension of the space X, and n is the number of data. At the same time, in order to stabilize the model and improve the accuracy, regularization constraints on the complete space X and the projection matrix X are added. Specifically, the square of the Frobenius norm is used for the complete space X, and the additional constraint $\|W_i^v\|_2 \leq 1$ is used for the projection matrix W^v, so the complete the spatial learning model:

$$\begin{aligned} &\min_{X, W^v} \frac{1}{V} \sum_{v=1}^{V} \|F^v - W^v X\|_F^2 + \gamma \|X\|_F^2 \\ &\text{s.t.} \|W_i^v\|_2 \leq 1 \end{aligned} . \quad \textbf{(1)}$$

Where γ is the regularization parameter.

2.2 Re-weighted Sparse Optimization Framework

The SSC algorithm use the sparse representation of vectors lying on a union of subspaces to cluster the data into separate subspaces. In order to obtain the sparse representation of each data point, the re-weighted l_1 norm minimization is used to perform convex relaxation. At the same time, in practical problems, data points are often mixed with sparse singular values and noise. In addition, the data are often distributed on the union of affine subspaces rather than linear subspaces. So a re-weighted sparse [30] optimization framework is established as follows:

$$\begin{aligned}&\min_{Z\,E}\|W\odot A\|_1+\lambda\|E\|_1\\&\text{s.t.}X=XA+E,[A]_{ii}=0\end{aligned}\quad. \tag{2}$$

Where $W\in R^{n\times n}$ is a re-weighted diagonal matrix. In reference [16], re-weighted matrix updating formula is as follows $w_i^k=\dfrac{1}{|a_i|+\varepsilon}$.

3 Research of this Paper

3.1 Proposed Approach

In this work, assume that different views all come from a common latent representation, as shown in Fig. 1. We can observe that view F^1 contains part of the same information and different information (complementary information) compared to view F^2, we call this part of the same information redundant information. In complete spatial learning, we should retain complementary information as much as possible and eliminate redundant information. Therefore, we use the low-rank characteristics of the kernel norm to eliminate the redundant features of the data, and at the same time we use more standard regularization constraints on the projection matrix. Our complete spatial learning model can be simplified to:

$$\begin{aligned}&\min_{X,W^v}\|X\|_*+\frac{\lambda}{V}\sum_{v=1}^{V}\|F^v-W^vX\|_F^2\\&\text{s.t.}W^{v\mathrm{T}}W^v=I\end{aligned}\quad. \tag{3}$$

We know that due to the complementarity of multiple views, the potential complete space describes the data more comprehensively than a single view. At the same time, considering that learning a sparse coefficient matrix can usually obtain good performance and reduce the computational burden of subsequent processing [31], we choose to perform re-weighted sparse subspace learning on *X*. On the other hand, we consider that the above subspace model [5-11] is usually solved by the Augmented Lagrange Multiplier (ALM) method, which requires a large number of iterations and high complexity. Therefore, we use the subspace of the quadratic programming to carry out the data of the potential complete space refactoring. Multi-view Re-weighted Sparse Subspace Clustering with Intact Low-rank Space Learning can be described as:

$$\begin{aligned}&\min\|X\|_*+\frac{\lambda_1}{V}\sum_{v=1}^{V}\|F^v-W^vX\|_F^2+\lambda_2\|H\odot A\|_1+\lambda_3\|X-XA\|_F^2\\&st.W^{v\mathrm{T}}W^v=I,A_{ii}=0\end{aligned}\quad. \tag{4}$$

Fig. 2 shows the main framework of the proposed method.

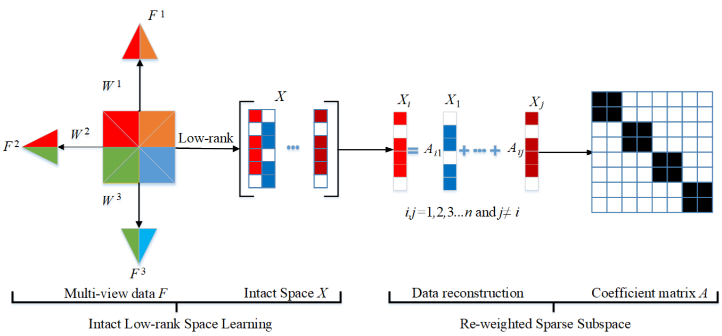


Fig. 2. Multi-view re-weighted sparse subspace clustering with intact low-rank space learning framework diagram (The model integrates all multi-view information $\{F^1, F^2, \ldots, F^v\}$ to learn Intact Low-rank Space Learning *X*, and completes data reconstruction through Re-weighted Sparse Subspace to form a unified sparse coefficient matrix *A*, and then uses normalized cut and other spectral clustering algorithms to The learned similarity is clustered, and the final clustering result is obtained.)

3.2 ILrS-MRSSC Solve by ALM

Rewrite eq. (4) as a lagrangian multiplier function through the ALM method, note *X*=*Z*, *A*=*J*:

$$\begin{aligned}\min_{X,Z,A,J,W^v,Y_1,Y_2} & \|Z\|_* + \frac{\lambda_1}{V}\sum_{v=1}^{V}\left\|F^v - W^v X\right\|_F^2 + \lambda_2\|H \odot J\|_1 + \lambda_3\|X - XA\|_F^2 \\ & + \mathrm{tr}\left[Y_1^{\mathrm{T}}(X-Z)\right] + \mathrm{tr}\left[Y_2^{\mathrm{T}}(A-J)\right] + \frac{\mu}{2}\left(\|X-Z\|_F^2 + \|A-J\|_F^2\right)\end{aligned} \quad (5)$$

1. *X*-**subproblem:** By leaving only terms in (5) that depend on *X*, we obtain:

$$\min_X \frac{\lambda_1}{V}\sum_{v=1}^{V}\left\|F^v - W^v X\right\|_F^2 + \mathrm{tr}\left[Y_1^{\mathrm{T}}(X-Z)\right] + \frac{\mu}{2}\|X-Z\|_F^2 + \lambda_3\|X - XA\|_F^2 . \quad (6)$$

Taking the derivative with respect to *X* and setting it to zero, we get:

$$\begin{aligned} & PX + XB = C \\ \text{with } P = & \left(\frac{2\lambda_1}{V}\sum_{v=1}^{V} W^{v\mathrm{T}} W^v + \mu I\right), \ B = 2\lambda_3 (I-A)^2 \\ C = & \left(\frac{2\lambda_1}{V} W^{v\mathrm{T}} F^v\right) - Y_1 + \mu Z \end{aligned} . \quad (7)$$

The above equation is a Sylvester equation, work [32] gives the solution method.

2. *Z*-**subproblem:** By leaving only terms in (5) that depend on *Z*, we obtain:

$$\min_Z \|Z\|_* + \mathrm{tr}\left[Y_1^T(X-Z)\right] + \frac{\mu}{2}\|X-Z\|_F^2 . \quad (8)$$

According to work [6], we get:

$$Z = \arg\min \frac{1}{\mu}\|Z\|_* + \frac{1}{2}\left\|Z - (X + \frac{Y_1}{\mu})\right\|_F^2 . \quad (9)$$

3. *A*-**subproblem:** By leaving only terms in (5) that depend on *A*, we obtain:

$$\min_A \lambda_3\|X - XA\|_F^2 + \mathrm{tr}\left[Y_2^{\mathrm{T}}(A-J)\right] + \frac{\mu}{2}\|A-J\|_F^2 . \quad (10)$$

Taking the derivative with respect to A and setting it to zero, we get:

$$A = \left(2\lambda_3 X^T X + \mu I\right)^{-1}\left(\mu J + 2\lambda_3 X^T X - Y_2\right). \quad (11)$$

4. *J*-**subproblem:** By leaving only terms in (5) that depend on *J*, we obtain:

$$\min_J \lambda_2\|H \odot J\|_1 + \mathrm{tr}\left[Y_2^{\mathrm{T}}(A-J)\right] + \frac{\mu}{2}\|A-J\|_F^2 . \quad (12)$$

Get the analytical formula according to the soft threshold function:

$$J = \text{sgn}\left(A + \frac{Y_2}{\mu}\right)\max\left(\left|A + \frac{Y_2}{\mu} - \frac{\lambda_2}{\mu}H\right|, 0\right). \tag{13}$$

5. W-**subproblem:** By leaving only terms in (5) that depend on W, we obtain:

$$\begin{aligned} &\min_{W^v} \lambda \left\| F^v - W^v X \right\|_F^2 \\ &\text{s.t.} W^{vT} W^v = I \end{aligned}. \tag{14}$$

According to work [33], the optimal solution is $W^{v\mathrm{T}} = UV^{\mathrm{T}}$, where U and V are left and right singular values of SVD decomposition of $XF^{v\mathrm{T}}$.

6. **Updating Multipliers:** We update the multipliers by:

$$\begin{cases} Y_1^{k+1} = Y_1^k + \mu(X - Z) \\ Y_2^{k+1} = Y_2^k + \mu(A - J) \\ \mu^{k+1} = \max(\mu^k \rho, \mu_{\max}) \end{cases}. \tag{15}$$

7. **Updating** H: We update the multipliers by:

$$H = \frac{1}{|A| + \varepsilon}, \tag{16}$$

note $\varepsilon \geq 0$.

Algorithm 1. Eq.(4) solve by ALM

Input: Data F^v, Parameter $\lambda_1, \lambda_2, \lambda_3, \varepsilon$, intact space feature d

initialization $\begin{aligned} &Y_1 = 0, Y_2 = 0, A = 0, J = 0, \mu_0 = 10^{-5} \\ &\rho = 1.2, \mu_{\max} = 10^5, X = rand(d, n) \end{aligned}$.

1. **while** not converged **do**
2. Z update by Eq.(9)
3. H update by Eq.(16)
4. J update by Eq.(13)
5. A update by Eq.(11)
6. X update by Eq.(7)
7. W update by Eq.(14)
8. Multipliers update by Eq.(15)
9. Check convergence conditions:

$$\|X - Z\|_\infty \leq \alpha \text{ and } \|A - J\|_\infty \leq \alpha$$

10. end while

Output: Optimal solution X and A

3.3 Complexity Analysis

The ILrS-MRSSC method is solved by the ALM method, and the complexity of the algorithm is mainly concentrated on the iterative process of alternating related variables. The complexity of the variable (Z, X, Y_1) is $O(3tdn)$, where d and n represent the dimensions of the variable, and t is the number of iterations; the complexity of the variable (A, J, Y_2, H) is $O(4tn^2)$; the complexity of the variable W^v is $\sum_v^V tdvd$, where dv represents the dimension of the variable, and V represents the number of views number; the complexity of the parameter μ is t . So put together, the complexity of the overall algorithm is $O\left(3tdn+4tn^2+\sum_{v}^{V} tdvd+t\right)$.

4 Experiments

In this section, we conduct some experiments to compare the performance of the algorithm through common real data sets, and choose two mainstream evaluation indicators, Clustering Accuracy (ACC) [14] and Normalized Mutual Information (NMI) [14] to observe performance of this algorithm.

We selected five real data for the experiment. The specific information is shown in Table 1. These data NUE-WIDE, MSRC-v1, 3-Sources, Yale and COIL-20 come from the literature [14, 28]. In terms of comparison methods, we chose AMGL [23], SwMC [24], ECMSC [26], MSC_IAS [28], LMSC [29], and the parameters of the comparison algorithm are set to the optimal parameters mentioned in the literature.

Table 1. Data set information

Data set	numbers	classes	views
NUS-WIDE	1600	8	6
MSRC-v1	210	7	5
3-Sources	169	6	3
Yale	165	15	3
COIL-20	1440	20	3

4.1 Performance Comparison Experiment

In this section, we conduct algorithm performance comparison experiments. It is set to run all algorithms 10 times, and the average value is used as the clustering result. Table 2 shows the parameter settings of the algorithm in this paper.

Table 2. Parameter settings for this method

Data set	d	λ_1	λ_2	λ_3	ε
NUS-WIDE	9	15	0.01	0.01	1.6
MSRC-v1	30	10	0.01	0.01	1.6
3-Sources	5	10	0.01	0.01	1.6
Yale	15	1.5	0.01	0.01	3.5
COIL-20	13	15	0.01	0.01	3.5

Table 3 and Table 4 report the performance results of different algorithms on different datasets, where the best results have been bolded and underlined. It is confirmed that our method achieves good performance in most scenarios. This is because the complete low-rank space learning can capture the complementary information of a single view and remove the redundant information, thereby obtaining a clean, noise-free latent complete space that satisfies the requirements of subspace learning. At the same time, the reweighted sparse subspace learning is unified into a complete spatial learning model, which strengthens the connection between multi-view data and subspace segmentation. The obtained coefficient matrix can describe the data information more comprehensively and improve the accuracy of subspace clustering.

Table 3. Performance results of multiple algorithms based on ACC indicators

Data set	AMGL	SwMC	MSC_IAS	ECMSC	LMSC	Ours
NUS-WIDE	0.2469	0.2256	0.2687	0.3040	**0.3106**	**<u>0.3177</u>**
MSRC-v1	0.7476	**<u>0.8514</u>**	0.6481	0.7910	0.7971	**0.8105**
3-Sources	0.3354	0.3513	**0.6491**	0.3395	0.6136	**<u>0.6574</u>**
Yale	0.4138	0.3646	**0.8369**	0.7108	0.5446	**<u>0.8691</u>**
COIL-20	**<u>0.8825</u>**	0.7216	0.8401	0.8605	0.8010	**0.8664**

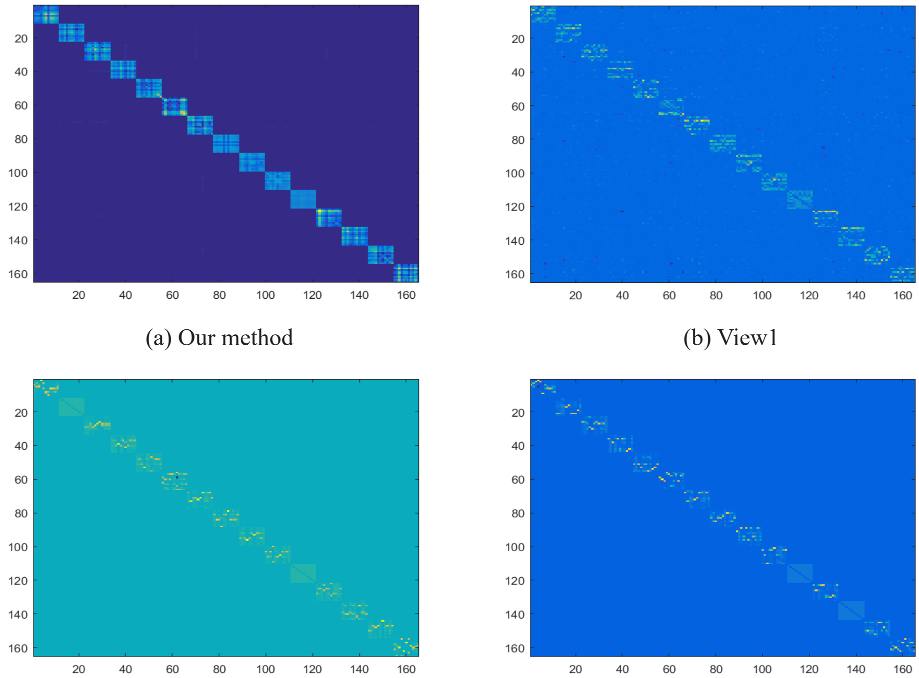
Table 4. Performance results of multiple algorithms based on NMI indicators

Data set	AMGL	SwMC	MSC_IAS	ECMSC	LMSC	Ours
NUS-WIDE	0.1360	0.1538	0.1605	**0.1897**	0.1595	**<u>0.2037</u>**
MSRC-v1	0.6676	**0.7629**	0.6471	0.7229	0.6133	**<u>0.8105</u>**
3-Sources	0.0635	0.0693	0.4308	0.0725	**0.4433**	**<u>0.5765</u>**
Yale	0.3757	0.3516	**0.8343**	0.7276	0.5095	**<u>0.9630</u>**
COIL-20	0.7652	0.7406	**0.9501**	0.8102	0.9100	**<u>0.9569</u>**

4.2 Visualization Experiment

In the work [12, 24], we know that the data sets Yale, MSRC-v1 and COIL-20 can directly display the clustering performance through visualization methods. In order to verify the performance of the method in this paper, we choose to perform re-weighted sparse subspace learning on a single view of the data, and visually display the obtained coefficient matrix and the coefficient matrix learned in this paper.

Fig. 3 to Fig. 5 show the similarity matrices of our method and the single-view subspace method on the Yale, MSRC-v1, and COIL-20 data sets. We can see from Fig. 3(b), Fig. 3(c) and Fig. 3(d) that the single-view data obtained from a better fixed angle can reflect the comprehensive information of the data to a certain extent. At this time, the single-view subspace, the method can learn a coefficient matrix with a relatively clear structure. However, (a) shows that the ILrS-MRSSC method can synthesize information from different views, making the block-diagonal structure of the coefficient matrix clearer.

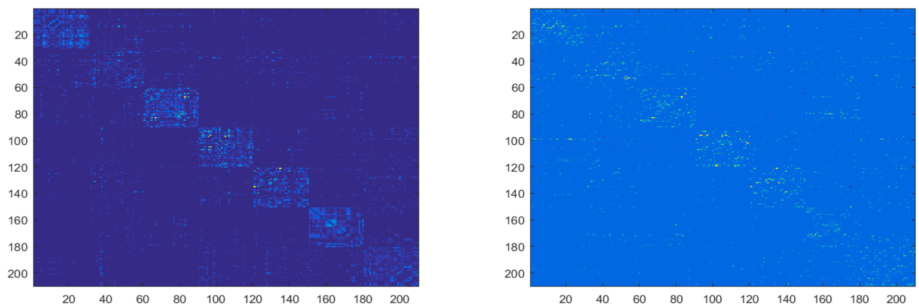


(a) Our method (b) View1

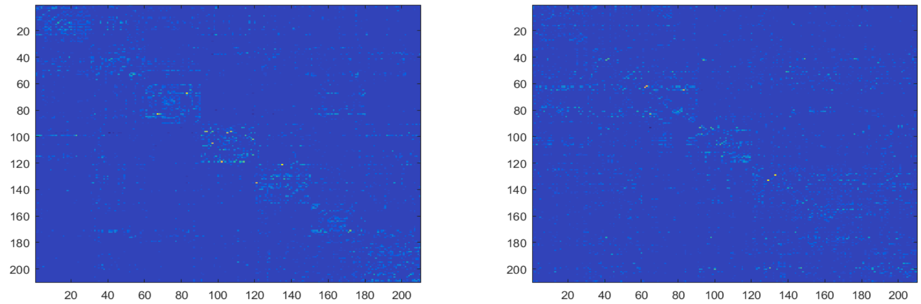
(c) View2 (d) View3

Fig. 3. Visualize the coefficient matrix based on Yale data

We can see from Fig. 4(b), Fig. 4(c) and Fig. 4(d) that when the single-view data has less information, the diagonal structure of the coefficient matrix block learned by the single-view subspace method is not obvious, it is difficult to describe the similarity relationship of the data. From the results in Fig. 4(a), it can be known that the ILrS-MRSSC method can restore a potential complete space by utilizing the complementary information of different views, thereby enhancing the block-diagonal structure stability and distinguishability of the coefficient matrix.



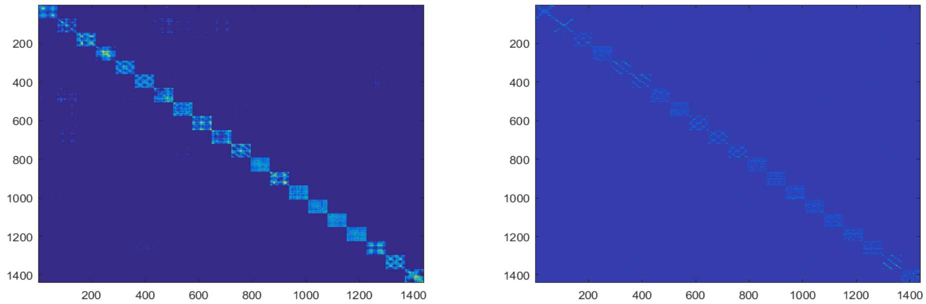
(a) Our method (b) View1



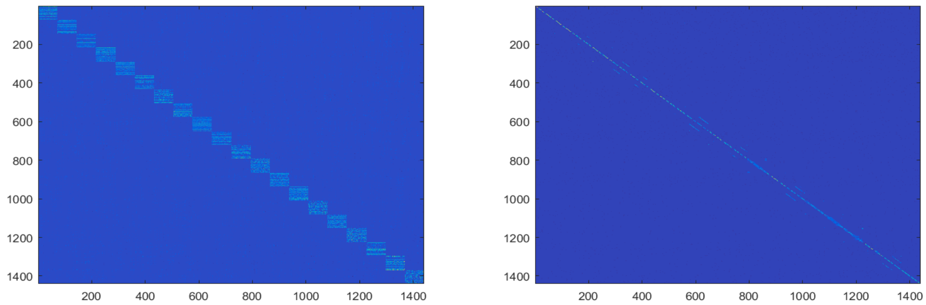
(c) View4 (d) View5

Fig. 4. Visualize the coefficient matrix based on MSRC-v1 data

The results in Fig. 5 further verify our conclusion. Since the data information of a single view is not comprehensive enough, the diagonal structure of the coefficient matrix is not obvious, and it is difficult to reflect the overall structure of the real data. The complete low-rank space learning can avoid the problem of poor subspace learning performance caused by insufficient single-view information, and our method can well reveal the underlying diagonal block structure, which is also required for clustering methods based on re-weighted sparse subspaces, which further validates the advantages of our unified model.



(a) Our method (b) View1



(c) View2 (d) View3

Fig. 5. Visualize the coefficient matrix based on COIL-20 data

5 Conclusion

In this paper, we propose a novel multi-view subspace learning method. Unlike most subspace learning methods, our method takes full advantage of the complementary information between multiple views in subspace learning. Specifically, we first reconstruct the latent complete space using multi-view information and low-rank features of nuclear norm, and then sparsely reorganize the data based on the latent complete space. The learned sparse coefficient matrix describes the underlying structure of the data more comprehensively, and has stronger discrimination and reference value. Finally, we formulate the whole problem as a unified optimization framework and jointly optimize by ALM method. Experimental results on benchmark datasets verify the superior performance of our proposed scheme. In the next stage, we consider how to adaptively determine the parameters based on the data structure to reduce tuning work.

6 Acknowledgement

This work is supported by the Natural Science Foundation of China (approval number: 61976130), Natural Science Foundation of Shaanxi Province (plan number: 2020JQ-923), key research and development projects of Shaanxi Province (plan number: 2018KW-021).

References

[1] R. Xia, Y. Pan, L. Du, J. Yin, Robust multi-view spectral clustering via low-rank and sparse decomposition, in: Proc. Twenty-Eighth AAAI Conference on Artificial Intelligence, 2014.
[2] X. Cao, C. Zhang, H. Fu, S. Liu, H. Zhang, Diversity-induced multi-view subspace clustering, in: Proc. Computer Vision & Pattern Recognition, 2015.
[3] H. Gao, F. Nie, X. Li, H. Huang, Multi-view subspace clustering, in: Proc. IEEE International Conference on Computer Vision, 2016.
[4] C. Zhang, H. Fu, S. Liu, G. Liu, X. Cao, Low-rank tensor constrained multiview subspace clustering, in: Proc. IEEE International Conference on Computer Vision, 2015.
[5] E. Elhamifar, R. Vidal, Sparse subspace clustering, in: Proc. IEEE Conference on Computer Vision & Pattern Recognition, 2009.
[6] G. Liu, Z. Lin, Y. Yu, Robust subspace segmentation by low-rank representation, in: Proc. International Conference on Machine Learning, 2010.
[7] L. Zhuang, H. Gao, Z. Lin, Y. Ma, X. Zhang, N. Yu, Non-negative low rank and sparse graph for semi-supervised learning, in: Proc. Computer Vision & Pattern Recognition, 2012.
[8] C. Lu, H. Min, Z. Zhao, L. Zhu, D. Huang, S. Yan, Robust and efficient subspace segmentation via least squares regression, in: Proc. European Conference on Computer Vision, 2012.
[9] M. Yin, X. Fang, S. Xie, Semi-supervised sparse subspace clustering on symmetric positive definite manifolds, in: Proc. Chinese Conference on Pattern Recognition, 2016.
[10]S. Basu, A. Banerjee, R. Mooney, Semi-supervised clustering by seeding, in: Proc. 19th International Conference on Machine Learning, 2002.
[11]K. Wagstaff, C. Cardie, Clustering with Instance-level Constraints, in: Proc. 17th International Conference on Machine Learning, 2000.
[12]S. Huang, H. Zhang, A. Pižurica, Semisupervised sparse subspace clustering method with a joint sparsity constraint for hyperspectral remote sensing images, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12(3)(2019) 989-999.
[13]L. Zhuang, Z. Zhou, S. Gao, J. Yin, Z. Lin, Y. Ma, Label information guided graph construction for semi-supervised learning, IEEE Transactions on Image Processing 26(9)(2017) 4182-4192.
[14]L. Fu, P. Lin, A.V. Vasilakos, S. Wang, An overview of recent multi-view clustering, Neurocomputing 402(2020) 148-161.
[15]P. Pavlidis, J. Weston, J. Cai, W.S. Noble, Learning gene functional classifications from multiple data types, Journal of Computational Biology 9(2)(2002) 401-411.
[16]Y. Zhang, X. Hu, X. Jiang, Multi-view clustering of microbiome samples by robust similarity network fusion and spectral clustering, IEEE/ACM Transactions on Computational Biology and Bioinformatics 14(2)(2017) 264-271.
[17]M. Zitnik, B. Zupan, Data fusion by matrix factorization, IEEE Transactions on Pattern Analysis and Machine Intelligence 37(1)(2015) 41-53.
[18]D. Greene, P. Cunningham, A Matrix Factorization Approach for Integrating Multiple Data Views, in: Proc. Machine Learning and Knowledge Discovery in Databases, 2009.
[19]R. Xia, Y. Pan, L. Du, J. Yin, Robust multi-view spectral clustering via low-rank and sparse decomposition, Proceedings of the AAAI Conference on Artificial Intelligence 28(1)(2014) 2149-2155.
[20]S.M. Kakade, D.P. Foster, Multi-view regression via canonical correlation analysis, in: Proc. Conference on Learning Theory, 2007.
[21]D. Niu, J.G. Dy, M.I. Jordan, Multiple non-redundant spectral clustering views, in: Proc. 27th International Conference on Machine Learning, 2010.
[22]X. Cai, F. Nie, W. Cai, H. Huang, Heterogeneous image features integration via multi-modal semi-supervised learning model, in: Proc. IEEE International Conference on Computer Vision, 2013.
[23]F. Nie, J. Li, X. Li, Parameter-free auto-weighted multiple graph learning: A framework for multiview clustering and semi-supervised classification, in: Proc. 25th International Joint Conference on Artificial Intelligence, 2016.
[24]F. Nie, J. Li, X. Li, Self-weighted multiview clustering with multiple graphs, in: Proc. 26th International Joint Conference

on Artificial Intelligence, 2017.
[25]F. Nie, G. Cai, X. Li, Multi-view clustering and semi-supervised classification with adaptive neighbours, in: Proc. 31th AAAI Conference on Artificial Intelligence, 2017.
[26]X. Wang, X. Guo, Z. Lei, C. Zhang, S.Z. Li, Exclusivity-consistency regularized multi-view subspace clustering, in: Proc. Computer Vision and Pattern Recognition, 2017.
[27]C. Xu, D. Tao, C. Xu, Multi-view intact space learning, IEEE Transactions on Pattern Analysis and Machine Intelligence 37(12)(2015) 2531-2544.
[28]X. Wang, Z. Lei, X. Guo, C. Zhang, H. Shi, S.Z. Li, Multi-view subspace clustering with intactness-aware similarity, Pattern Recognition 88(2019) 50-63.
[29]C. Zhang, Q. Hu, H. Fu, P. Zhu, X. Cao, Latent multi-view subspace clustering, in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[30]J. Xu, K. Xu, K. Chen, J. Ruan, Reweighted sparse subspace clustering, Computer Vision and Image Understanding 138(2015) 25-37.
[31]F. Nie, X. Wang, H. Huang, Clustering and projected clustering with adaptive neighbors, in: Proc. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2014.
[32]R.H. Bartels, G.W. Stewart, Solution of the matrix equation AX+XB=C, Communications of the ACM 15(9)(1972) 820-826.
[33]J. Huang, F. Nie, H. Huang, Spectral rotation versus k-means in spectral clustering, in: Proc. 27th AAAI Conference on Artificial Intelligence, 2013.