

Partially latent factors based multi-view subspace learning

Run-kun Lu^a, Jian-wei Liu^b, Ze-yu Liu^b, Jin-zhong Chen^a

^aChina Special Equipment Inspection and Research Institute, Beijing 100029, China

^bDepartment of Automation, College of Information Science and Engineering, China University of Petroleum, Beijing, 260 Mailbox, Changping District, Beijing 102249, China

Abstract

Multi-view subspace clustering always performs well in high-dimensional data analysis, but is sensitive to the quality of data representation. To this end, a two stage fusion strategy is proposed to embed representation learning into the process of multi-view subspace clustering. This paper first propose a novel matrix factorization method that can separate the coupling consistent and complementary information from observations of multiple views. Based on the obtained latent representations, we further propose two subspace clustering strategies: feature-level fusion and subspace-level hierarchical strategy. Feature-level method concatenates all kinds of latent representations from multiple views, and the original problem therefore degenerates to a single-view subspace clustering process. Subspace-level hierarchical method performs different self-expressive reconstruction processes on the corresponding complementary and consistent latent representations coming from each view, i.e. the prior constraints imposed on different types of subspace representations are related to the appropriate input factors. Finally, extensive experimental results on real-world datasets demonstrate the superiority of our proposed methods by comparing against some state-of-the-art subspace clustering algorithms.

Keywords: multi-view subspace clustering, matrix factorization, representation learning

1. Introduction

In real-world applications, an object can be observed from multiple perspectives, which are denoted as views or modalities. In particular, readers acquire information from the title, text, and picture while reading the newspaper; or multi-feature learning always generates multiple kinds of features using different approaches, e.g., extracting diverse image features using LBP [1], HOG [2], and SIFT [3] operators. In contrast to typical machine learning approaches, multi-view learning is expert in modeling the relationships between views or modalities [4]. A fundamental assumption behind such a paradigm is that different views share some semantics called consistent or consensus information, and each view also has unique semantic called complementary or view-specific information [5]. Obviously, the above assumptions conform to real-world physical law, e.g., each camera in a multi-camera system has a shared field of view and its own unique viewpoint. Therefore, multi-view learning could leverage both consistency and complementarity to explore more compact and complete information from multiple observations.

Subspace clustering is one of the most attractive research topics in the field of multi-view learning, especially for algorithms based on self-expressiveness [6, 7, 8, 9, 10, 11, 12, 13, 14], which state each data point in a union of subspaces can be represented by a linear combination of all other data points [15]. Existing multi-view subspace clustering usually focus on two perspectives: the prior constraint on subspace representation and the construction

*This work was supported by the Science Foundation of China University of Petroleum, Beijing (No. 2462020YXZZ2023).

**We note that a shorter conference version of this paper appeared in International Joint Conference on Neural Networks (2021). Our initial conference paper only propose the subspace-level hierarchical method MVSC-CSLF. This manuscript propose another feature-level fusion algorithm MVSC-CSLF. Besides, the proposed algorithms are analyzed and described in more details, and the performances are verified on more datasets.

Email addresses: zsy1rk@gmail.com (Run-kun Lu), liujw@cup.edu.cn (Jian-wei Liu), 2020310708@student.cup.edu.cn (Ze-yu Liu), cchenbeiter@aliyun.com (Jin-zhong Chen)

method of latent representation. To the former, different effects can be obtained by introducing various prior constraints on subspace representation, such as [9] imposes sparse and low rank constraints on corresponding subspace representations simultaneously, then fuses all obtained coefficient matrices with an average operation to construct the final adjacency matrix; [13] integrates multiple coefficient matrices into a common subspace representation through learning with Laplacian regularization. In addition, for the construction method of latent representation, researchers often attempt to mine the underlying complementary and consistent information between views to obtain more complete and non-redundant latent representations, and thus improving performance of the learning task. For instance, [16] discovers a latent embedding representation involved complementary information from multiple views; and [14] first concatenates all views together and handles the view-specific corruptions with $L_{2,1}$ norm so that the model can leverage the consistent semantics to improve the clustering performance. Compared to learning compact and comprehensive latent representation from multiple views, most algorithms are more concerned with the prior norms and subspace representation learning process. However, self-expressive subspace clustering requires the data should be well distributed inside the subspace [15], and it is difficult to satisfy the condition with real-world datasets. Therefore, it makes sense to introduce representation learning in the process of multi-view subspace clustering, which not only provides a good representation for subspace clustering, but also can utilize the specific task to guide the algorithm to explore the underlying semantics from multiple observations.

Motivated by this, we propose a two-stage strategy that can link representation learning with specific task (multi-view clustering). Firstly, a novel matrix factorization algorithm is introduced to separate the coupling consistent and complementary information from multiple observations. Among them, the consistent one involves the common information shared by all views, and complementary one is the unique view-specific information corresponding to each view. And based on this, we further propose two self-expressive learning methods: Multi-View Subspace Clustering with Consistent and view-Specific Latent Factors (MVSC-CSLF), and Multi-View Subspace Clustering with Consistent and view-Specific Latent Factors and Subspaces (MVSC-CSLFS). The first method MVSC-CSLF concatenates the learned multiple representations and applies it to least squares regression subspace clustering algorithm [17, 18], which indicates that competitive clustering performance can be obtained even using a simple clustering strategy with a good representation. And the second algorithm MVSC-CSLFS imposes various prior constraints on different kinds of subspace representations, i.e., the self-expressive reconstruction task guides the matrix factorization process to separate the coupled consistent and complementary information, and the appropriate prior terms imposed on them derive better clustering performance. Finally, adaptive weights corresponding to multiple views are introduced to the learning process, which improves the model’s ability to accommodate noisy or unreliable views. Extensive experiments on real-world datasets demonstrate that even the simple method MVSC-CSLF outperforms than many baseline algorithms, and with more suitable subspace clustering strategy, MVSC-CSLFS has achieved superior performance on most datasets.

Our contributions in this study are summarized as follows:

- (1) Our proposed matrix factorization approach helps MVSC-CSLF and MVSC-CSLFS to obtain more compact and comprehensive latent representations by considering consistent and complementary information of multiple views, simultaneously. In particular, the method could separate the coupled consistent and complementary information to make latent representations have more explicit semantics.
- (2) MVSC-CSLFS regularizes subspace representations with different norms according to the types of learned semantics, which makes the appropriate prior constraint applied on suitable subspace representation.
- (3) Our proposed methods both introduce dynamic weights learning process for each view, which eliminates the influence of unreliable views.

The remainder of this paper is organized as follows: section 2 introduces the related work and notation definitions of this study; section 3 first elaborates our proposed MVSC-CSLF and MVSC-CSLFS, then introduces their optimization method in details; section 4 gives detailed introduction and discussion of experimental results; section 5 concludes this study and discusses some future works.

2. Related Work and Notations

2.1. Related work

Self-expressive subspace clustering has several advantages over traditional subspace clustering methods:

- (1) the subspace dimensionality does not need to be set in advance;
- (2) the global geometric information of the data can be exploited;
- (3) the computational overhead is relatively small;
- (4) the clustering accuracy can be theoretically guaranteed under certain conditions.

And the method has been widely used in real-world applications, such as the Internet-of-Things (IoT) [19], human motion segmentation [20, 21], and face recognition [22, 23, 24]. In particular, Sparse Subspace Clustering (SSC) [25] and Low-Rank Representation (LRR) are foundational studies, which provide a standard learning paradigm:

$$\begin{aligned} \min_{\mathbf{C}} R(\mathbf{C}) \\ \text{s.t. } \mathbf{X} = \mathbf{XC}, \mathbf{C} \in \Omega, \end{aligned}$$

where \mathbf{C} is the subspace representation, and $R(\cdot)$ is the prior constraint on it. Then the adjacency matrix can be obtained using the following formula:

$$\frac{|\mathbf{C}| + |\mathbf{C}^T|}{2},$$

and one can substitute it into standard spectral clustering to obtain clustering result. Based on the above process, plenty of improved algorithms for self-expressive subspace clustering have been proposed recent years [26, 27, 28, 29, 30, 31].

Besides, multi-view subspace clustering has gradually become an important method because it can model the relationships between views, and explore the underlying common and view-specific information of subspace representations corresponding to different views. Specifically, [9, 32] both impose low-rank and sparse constraints on the subspace representations, simultaneously. [8] is the first multi-view method that combines representation learning and subspace clustering. The algorithm assumes that all the observations (views) originate from one common latent representation, and performs the self-expressive reconstruction on it. Besides, [14, 33] impose Laplacian constraint on subspace representations to preserve local manifold structures of each view. Most standard methods cannot deal with highly nonlinear data, therefore some kernel based methods are employed [7, 9, 32, 13, 34], but they also encounter the dilemma of kernel selection. Inspired by [35], researchers begin to explore the network structure of deep multi-view subspace clustering [36, 37, 12], and [38] further introduces attention mechanism in network to derive dynamic weights corresponding multiple views.

2.2. Notations

In this paper, we define matrices and column vectors using bold uppercase and lowercase characters, respectively, and scalars are denoted by italics and not bold characters. Let $[\mathbf{A}, \mathbf{B}]$ and $[\mathbf{A}; \mathbf{B}]$ denote the horizontal and vertical concatenation between matrices \mathbf{A} and \mathbf{B} . Then let $\text{diag}(\mathbf{X})$ denotes the diagonal elements of matrix $\mathbf{X} \in \mathbb{R}^{I \times J}$, and the norms we used throughout this paper are summarized as follows:

$$\text{2-norm: } \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^I x_i};$$

$$\text{Frobenius norm: } \|\mathbf{X}\|_F = \sqrt{\sum_{i=1}^I \sum_{j=1}^J x_{i,j}^2};$$

$$L_{2,1}\text{-norm: } \|\mathbf{X}\|_{2,1} = \sum_{j=1}^J \sqrt{\sum_{i=1}^I x_{i,j}^2};$$

$$\text{Nuclear norm: } \|\mathbf{X}\|_* = \sum_{i=1}^r \sigma_i, \text{ where } \sigma_i \text{ is } i\text{-th non-zero singular value, and } r \text{ is the total number of non-zero singular values } (r \leq \min(I, J)).$$

Since a number of symbols need to be defined to help the description of our proposed algorithms, we have summarized all commonly used symbols in Table 1 for the convenience of reference.

Table 1: Notations

Notation	Size	Description
\mathbf{X}^v	$M^v \times N$	$\mathbf{X}^v = [\mathbf{x}_1^v, \dots, \mathbf{x}_N^v]$, where $\mathbf{x}_n^v \in \mathbb{R}^{M^v}$ is n -th instance of v -th view.
\mathbf{P}_s^v	$M^v \times K_s$	Projection matrix that projects v -th view's specific latent representation into the corresponding observation space.
\mathbf{P}_c^v	$M^v \times K_c$	Projection matrix that projects v -th view's consistent latent representation into the corresponding observation space.
\mathbf{H}_s^v	$K_s \times N$	Specific latent representation of v -th view.
\mathbf{H}_c	$K_c \times N$	Consistent latent representation shared by all views.
\mathbf{H}	$(K_s \times V + K_c) \times N$	Joint latent representation concatenated by $\{\mathbf{H}_s^v\}_{v=1}^V$ and \mathbf{H}_c following the rule: $[\mathbf{H}_s^1; \dots; \mathbf{H}_s^V; \mathbf{H}_c]$.
\mathbf{Z}	$N \times N$	self-expressive matrix for MVSC-CSLF.
\mathbf{Z}^v	$N \times N$	View-specific self-expressive matrix of v -th view for MVSC-CSLFS.
\mathbf{Z}_c	$N \times N$	Consistent self-expressive matrix shared by all views for MVSC-CSLFS.
$\mathbf{D}, \mathbf{D}^v, \mathbf{D}_c$	$N \times N$	Auxiliary variable of \mathbf{Z}, \mathbf{Z}^v , and \mathbf{Z}_c .
\mathbf{E}_r^v	$M^v \times N$	Observation recovery error of v -th view.
\mathbf{E}_s	$(K_s \times V + K_c) \times N$	self-expressive error for MVSC-CSLF.
\mathbf{E}_s^v	$K_s \times N$	View-specific self-expressive error of v -th view for MVSC-CSLFS.
\mathbf{E}_s^c	$K_c \times N$	Consistent self-expressive error shared by all views for MVSC-CSLFS.
$\boldsymbol{\pi}$	$1 \times V$	Dynamic weights collection corresponding to the observation recovery process for MVSC-CSLF.
$\boldsymbol{\pi}_1$	$1 \times V$	Dynamic weights collection corresponding to the observation recovery process for MVSC-CSLFS.
$\boldsymbol{\pi}_2$	$1 \times (V + 1)$	Dynamic weights collection corresponding to the self-expressive reconstruction process for MVSC-CSLFS.

3. Proposed Methods

Suppose $\{\mathbf{X}^v \in \mathbb{R}^{M^v \times N}\}_{v=1}^V$ is the collection of observational data comprising V views, where M^v and N represent instance dimension and number, respectively. We assume that observable views mainly originate from two parts: the consistent latent representation $\mathbf{H}_c \in \mathbb{R}^{K_c \times N}$ shared by all views; the view-specific latent representation $\mathbf{H}_s^v \in \mathbb{R}^{K_s \times N}$ related only to v -th view. Then suppose that \mathbf{X}^v can be recovered by \mathbf{H}_s^v and \mathbf{H}_c according to the following rule:

$$\mathbf{X}^v = \mathbf{P}_s^v \mathbf{H}_s^v + \mathbf{P}_c^v \mathbf{H}_c,$$

where $\mathbf{P}_s^v \in \mathbb{R}^{M^v \times K_s}$ and $\mathbf{P}_c^v \in \mathbb{R}^{M^v \times K_c}$ are projections corresponding to view-specific and consistent latent representations, respectively. Matrix factorization and fusion process of multiple views is illustrated in Fig. 1.

We construct a pipeline that combines latent representation learning process with subspace clustering, so that the specific task could guide the algorithm to extract feature, and the representation learning process also provides subspace clustering with compact and comprehensive factors. Therefore, our proposed two algorithms all mainly contain two learning goals, one is observation information recovery, the other is self-expressive reconstruction:

$$\mathcal{I}(\mathbf{X}^v, \mathbf{H}_s^v, \mathbf{H}_c; \boldsymbol{\Theta}_1) + \lambda \mathcal{S}(\mathbf{H}_s^v, \mathbf{H}_c; \boldsymbol{\Theta}_2), \quad (1)$$

where $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ are the underlying parameters related to the two terms, and λ is the trade-off parameter to balance multiple terms. The first term $\mathcal{I}(\cdot, \cdot)$ of equation (1) is to measure the recovery ability from latent representations to observations, and the specific form in this paper is the same for two algorithms we proposed, which is formulated as follow:

$$\begin{aligned} \mathcal{I}(\mathbf{X}^v, \mathbf{H}_s^v, \mathbf{H}_c; \boldsymbol{\Theta}_1) &= \min_{\boldsymbol{\Theta}_1, \mathbf{H}_s^v, \mathbf{H}_c} \sum_{v=1}^V \|\mathbf{E}_r^v\|_{2,1} \\ \text{s.t. } (\mathbf{P}_s^v)^T \mathbf{P}_s^v &= \mathbf{I}, (\mathbf{P}_c^v)^T \mathbf{P}_c^v = \mathbf{I}, \mathbf{X}^v = \mathbf{P}_s^v \mathbf{H}_s^v + \mathbf{P}_c^v \mathbf{H}_c + \mathbf{E}_r^v, \end{aligned} \quad (2)$$

where $\mathbf{E}_r^v \in \mathbb{R}^{M^v \times N}$ is the recovery error, and $\boldsymbol{\Theta}_1 = \{\mathbf{P}_s^v, \mathbf{P}_c^v, \mathbf{E}_r^v\}$. \mathbf{P}_s^v and \mathbf{P}_c^v are constrained to prevent the value of \mathbf{H}_s^v and \mathbf{H}_c from being imposed to 0.

The second term of equation (1) measures the self-expressive ability. The strategies employed by our proposed two methods are different, and we will discuss the details in the next two subsections.

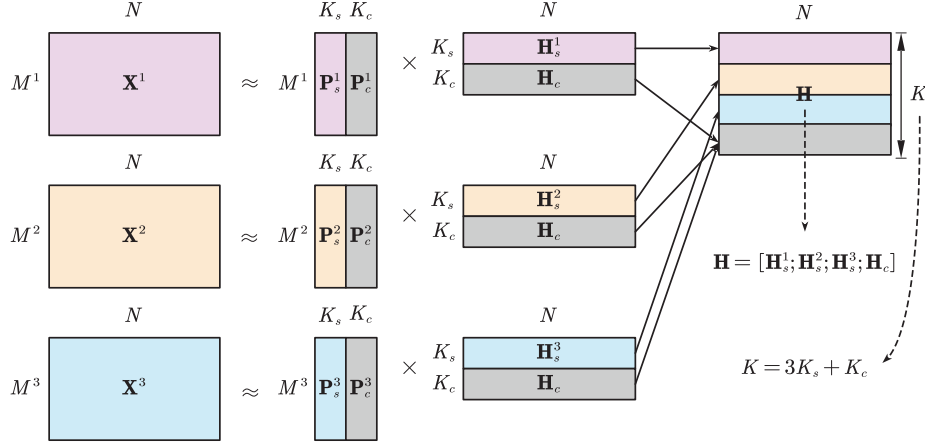


Figure 1: **Illustration of matrix factorization and fusion:** there are three views in this example, and each of them approximately equals to the product of projection matrix and latent representation matrix. Both of the matrices can be divided into view-specific and consistent parts. Take view 1 as an example, we concatenate \mathbf{P}_s^1 and \mathbf{P}_c^1 horizontally to construct projection matrix, and concatenate \mathbf{H}_s^1 and \mathbf{H}_c^1 vertically to construct latent representation. And we can finally obtain a unified joint latent representation \mathbf{H} through vertical concatenation of \mathbf{H}_s^1 , \mathbf{H}_s^2 , \mathbf{H}_s^3 , and \mathbf{H}_c .

3.1. MVSC-CSLF

In this subsection, we propose a novel framework called Multi-View Subspace Clustering with Consistent and view-Specific Latent Factors (MVSC-CSLF). Firstly, refer to equation (2) and Fig. 1, a joint latent representation can be defined as the vertical concatenation of view-specific and consistent representations: $\mathbf{H} \in \mathbb{R}^{K \times N} = [\mathbf{H}_s^1; \dots; \mathbf{H}_s^V; \mathbf{H}_c]$, where $K = K_s \times V + K_c$. Obviously, the representation concludes multiple views' complementary and consistent information, simultaneously. Finally, Self-expressive reconstruction is applied on \mathbf{H} . We state such a learning process as feature-level fusion strategy, i.e. fusing different factors as a joint latent representation, and make the downstream subspace clustering task degenerate to a single-view problem. Consequently, the second term of equation (1) is formulated as follow:

$$\begin{aligned} \mathcal{S}(\mathbf{H}_s^v, \mathbf{H}_c; \Theta_2) &= \min_{\Theta_2, \mathbf{H}_s^v, \mathbf{H}_c} \lambda_1 \|\mathbf{E}_s\|_{2,1} + \frac{\lambda_2}{2} \|\mathbf{Z}\|_F^2 \\ \text{s.t. } \mathbf{H} &= \mathbf{H}\mathbf{Z} + \mathbf{E}_s, \text{diag}(\mathbf{Z}) = 0, \end{aligned} \quad (3)$$

where \mathbf{E}_s is self-expressive error, \mathbf{Z} is self-expressive matrix, λ_1 and λ_2 are trade-off parameters, and $\Theta_2 = \{\mathbf{E}_s, \mathbf{Z}\}$. Then formulas (2) and (3) are combined to construct a joint optimization problem by introducing dynamic weight π and auxiliary variable \mathbf{D} :

$$\begin{aligned} \min_{\Theta} \quad & \sum_{v=1}^V \pi^v \|\mathbf{E}_r^v\|_{2,1} + \lambda_1 \|\mathbf{E}_s\|_{2,1} + \frac{\lambda_2}{2} \|\mathbf{D}\|_F^2 + \frac{\lambda_3}{2} \|\pi\|_2^2 \\ \text{s.t. } \quad & (\mathbf{P}_s^v)^T \mathbf{P}_s^v = \mathbf{I}, (\mathbf{P}_c^v)^T \mathbf{P}_c^v = \mathbf{I}, \\ & \mathbf{X}^v = \mathbf{P}_s^v \mathbf{H}_s^v + \mathbf{P}_c^v \mathbf{H}_c + \mathbf{E}_r^v, \\ & \mathbf{H} = \mathbf{H}\mathbf{Z} + \mathbf{E}_s, \\ & \mathbf{D} = \mathbf{Z} - \text{diag}(\mathbf{Z}), \\ & \pi \geq 0, \sum_{v=1}^V \pi^v = 1, \end{aligned} \quad (4)$$

where $\Theta = \{\mathbf{E}_r^v, \mathbf{E}_s, \mathbf{P}_s^v, \mathbf{P}_c^v, \mathbf{H}_s^v, \mathbf{H}_c, \mathbf{D}, \mathbf{Z}, \pi\}$, and $\lambda_1, \lambda_2, \lambda_3$ are trade-off parameters, $\pi = \{\pi^v\}_{v=1}^V$ weights the importance of each view. The regularization on π is to prevent the model from setting the weight value of the best view to 1. The optimization process will be introduced in subsection 3.3 together with MVSC-CSLFS.

3.2. MVSC-CSLFS

In this subsection, we further propose a improved framework based on MVSC-CSLF named Multi-View Subspace Clustering with Consistent and view-Specific Latent Factors and Subspaces (MVSC-CSLFS). Compare with MVSC-CSLF, a subspace-level hierarchical strategy is employed to preform self-expressive reconstruction on all views'

specific and consistent latent representations, respectively. Such strategy imposes more appropriate prior constraints on the corresponding representation with different attributes. Therefore, the second term of equation (1) can be formulated as:

$$\begin{aligned} S(\mathbf{H}_s^v, \mathbf{H}_c; \Theta_2) &= \min_{\Theta_2, \mathbf{H}_s^v, \mathbf{H}_c} \lambda_1 \left(\sum_{v=1}^V \|\mathbf{E}_s^v\|_{2,1} + \|\mathbf{E}_s^c\|_{2,1} \right) + \lambda_2 \left(\sum_{v=1}^V \frac{1}{2} \|\mathbf{Z}^v\|_F^2 + \|\mathbf{Z}_c\|_* \right) \\ s.t. \quad \mathbf{H}_s^v &= \mathbf{H}_s^v \mathbf{Z}^v + \mathbf{E}_s^v, \quad \mathbf{H}_c = \mathbf{H}_c \mathbf{Z}_c + \mathbf{E}_s^c, \\ \text{diag}(\mathbf{Z}^v) &= 0, \quad \text{diag}(\mathbf{Z}_c) = 0, \end{aligned} \quad (5)$$

where \mathbf{E}_s^v and \mathbf{E}_s^c are reconstruction errors related to \mathbf{H}_s^v and \mathbf{H}_c , \mathbf{Z}^v and \mathbf{Z}_c are subspace representations related to \mathbf{H}_s^v and \mathbf{H}_c , λ_1 and λ_2 are trade-off parameters, and $\Theta_2 = \{\mathbf{E}_s^v, \mathbf{E}_s^c, \mathbf{Z}^v, \mathbf{Z}_c\}$. Then formulas (2) and (5) are combined to construct a joint optimization problem by introducing dynamic weights π_1 , π_2 , and auxiliary variable \mathbf{D}^v , \mathbf{D}_c :

$$\begin{aligned} \min_{\Theta} \quad & \sum_{v=1}^V \pi_1^v \|\mathbf{E}_r^v\|_{2,1} + \lambda_1 \left(\sum_{v=1}^V \pi_2^v \|\mathbf{E}_s^v\|_{2,1} + \pi_2^{V+1} \|\mathbf{E}_s^c\|_{2,1} \right) + \lambda_2 \left(\sum_{v=1}^V \frac{\pi_2^v}{2} \|\mathbf{D}^v\|_F^2 + \pi_2^{V+1} \|\mathbf{D}_c\|_* \right) + \frac{\lambda_3}{2} (\|\pi_1\|_2^2 + \|\pi_2\|_2^2) \\ s.t. \quad & (\mathbf{P}_s^v)^T \mathbf{P}_s^v = \mathbf{I}, \quad (\mathbf{P}_c^v)^T \mathbf{P}_c^v = \mathbf{I}, \\ & \mathbf{X}^v = \mathbf{P}_s^v \mathbf{H}_s^v + \mathbf{P}_c^v \mathbf{H}_c + \mathbf{E}_r^v, \\ & \mathbf{H}_s^v = \mathbf{H}_s^v \mathbf{Z}^v + \mathbf{E}_s^v, \quad \mathbf{H}_c = \mathbf{H}_c \mathbf{Z}_c + \mathbf{E}_s^c, \\ & \mathbf{D}^v = \mathbf{Z}^v - \text{diag}(\mathbf{Z}^v), \quad \mathbf{D}_c = \mathbf{Z}_c - \text{diag}(\mathbf{Z}_c), \\ & \pi_1, \pi_2 \geq 0, \quad \sum_{v=1}^V \pi_1^v = 1, \quad \sum_{v=1}^{V+1} \pi_2^v = 1, \end{aligned} \quad (6)$$

where $\Theta = \{\mathbf{E}_r^v, \mathbf{E}_s^v, \mathbf{E}_s^c, \mathbf{P}_s^v, \mathbf{P}_c^v, \mathbf{H}_s^v, \mathbf{H}_c, \mathbf{D}^v, \mathbf{D}_c, \mathbf{Z}^v, \mathbf{Z}_c, \pi_1, \pi_2\}$; $\lambda_1, \lambda_2, \lambda_3$ are trade-off parameters; $\pi_1 = \{\pi_1^v\}_{v=1}^V$ weights the recovery ability of each view; $\pi_2 = \{\pi_2^v\}_{v=1}^{V+1}$ weights the self-expressive ability of view-specific and consistent latent representations from multiple views. Specifically, the regularization terms on π_1 and π_2 are to prevent the model from setting the weight value of the best view to one. The optimization process will be introduced in subsection 3.3 together with MVSC-CSLF.

3.3. Optimization

In this subsection, we will derive the updating formulas of MVSC-CSLF and MVSC-CSLFS by constructing augmented lagrangian dual functions and using alternating direction minimization strategy. Firstly, the lagrangian dual forms for equations (4) and (6) are formulated as:

$$\begin{aligned} \min_{\Theta} \quad & \sum_{v=1}^V \pi^v \|\mathbf{E}_r^v\|_{2,1} + \lambda_1 \|\mathbf{E}_s\|_{2,1} + \frac{\lambda_2}{2} \|\mathbf{D}\|_F^2 + \frac{\lambda_3}{2} \|\pi\|_2^2 \\ & + \sum_{v=1}^V \frac{\mu}{2} \left\| \frac{1}{\mu} \mathbf{A}_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right\|_F^2 \\ & + \frac{\mu}{2} \left\| \frac{1}{\mu} \mathbf{A}_2 + \mathbf{H} - \mathbf{H} \mathbf{Z} - \mathbf{E}_s \right\|_F^2 \\ & + \frac{\mu}{2} \left\| \frac{1}{\mu} \mathbf{A}_3 + \mathbf{D} - \mathbf{Z} + \text{diag}(\mathbf{Z}) \right\|_F^2 \\ & - \sum_{v=1}^V \frac{1}{2\mu} \|\mathbf{A}_1^v\|_F^2 - \frac{1}{2\mu} \|\mathbf{A}_2\|_F^2 - \frac{1}{2\mu} \|\mathbf{A}_3\|_F^2 \\ s.t. \quad & (\mathbf{P}_s^v)^T \mathbf{P}_s^v = \mathbf{I}, \quad (\mathbf{P}_c^v)^T \mathbf{P}_c^v = \mathbf{I}, \\ & \pi \geq 0, \quad \sum_{v=1}^V \pi^v = 1, \end{aligned} \quad (7)$$

where Λ_1^v , Λ_2 , and Λ_3 are multipliers, μ is a positive penalty scalar.

$$\begin{aligned}
\min_{\Theta} \quad & \sum_{v=1}^V \pi_1^v \|\mathbf{E}_r^v\|_{2,1} + \lambda_1 \left(\sum_{v=1}^V \pi_2^v \|\mathbf{E}_s^v\|_{2,1} + \pi_2^{V+1} \|\mathbf{E}_s^c\|_{2,1} \right) \\
& + \lambda_2 \left(\sum_{v=1}^V \frac{\pi_2^v}{2} \|\mathbf{D}^v\|_F^2 + \pi_2^{V+1} \|\mathbf{D}_c\|_* \right) + \frac{\lambda_3}{2} (\|\boldsymbol{\pi}_1\|_2^2 + \|\boldsymbol{\pi}_2\|_2^2) \\
& + \sum_{v=1}^V \frac{\mu}{2} \left\| \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right\|_F^2 \\
& + \sum_{v=1}^V \frac{\mu}{2} \left\| \frac{1}{\mu} \Lambda_2^v + \mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z}^v - \mathbf{E}_s^v \right\|_F^2 \\
& + \frac{\mu}{2} \left\| \frac{1}{\mu} \Lambda_3 + \mathbf{H}_c - \mathbf{H}_c \mathbf{Z}_c - \mathbf{E}_s^c \right\|_F^2 \\
& + \sum_{v=1}^V \frac{\mu}{2} \left\| \frac{1}{\mu} \Lambda_4^v + \mathbf{D}^v - \mathbf{Z}^v + \text{diag}(\mathbf{Z}^v) \right\|_F^2 \\
& + \frac{\mu}{2} \left\| \frac{1}{\mu} \Lambda_5 + \mathbf{D}_c - \mathbf{Z}_c + \text{diag}(\mathbf{Z}_c) \right\|_F^2 \\
& - \frac{1}{2\mu} \left(\sum_{v=1}^V (\|\Lambda_1^v\|_F^2 + \|\Lambda_2^v\|_F^2 + \|\Lambda_4^v\|_F^2) + \|\Lambda_3\|_F^2 + \|\Lambda_5\|_F^2 \right) \\
s.t. \quad & (\mathbf{P}_s^v)^T \mathbf{P}_s^v = \mathbf{I}, \quad (\mathbf{P}_c^v)^T \mathbf{P}_c^v = \mathbf{I}, \\
& \boldsymbol{\pi}_1, \boldsymbol{\pi}_2 \geq 0, \quad \sum_{v=1}^V \pi_1^v = 1, \quad \sum_{v=1}^{V+1} \pi_2^v = 1,
\end{aligned} \tag{8}$$

where Λ_1^v , Λ_2^v , Λ_3 , Λ_4^v and Λ_5 are lagrangian multipliers, μ is a positive penalty scalar. We can decompose the original problem into several sub-optimization problems, and derive their updating rules, respectively.

3.3.1. \mathbf{P} sub-problems

The sub-problems w.r.t \mathbf{P}_s^v and \mathbf{P}_c^v are the same for MVSC-CSLF and MCSC-CSLFS, and they can be formulated as:

$$\begin{aligned}
\mathbf{P}_s^{v*} &= \arg \min_{\mathbf{P}_s^v} \left\| (\mathbf{Y}_1)^T - (\mathbf{H}_s^v)^T (\mathbf{P}_s^v)^T \right\|_F^2 \\
s.t. \quad & (\mathbf{P}_s^v)^T \mathbf{P}_s^v = \mathbf{I},
\end{aligned} \tag{9}$$

where $\mathbf{Y}_1 = \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v$.

$$\begin{aligned}
\mathbf{P}_c^{v*} &= \arg \min_{\mathbf{P}_c^v} \left\| (\mathbf{Y}_2)^T - (\mathbf{H}_c)^T (\mathbf{P}_c^v)^T \right\|_F^2 \\
s.t. \quad & (\mathbf{P}_c^v)^T \mathbf{P}_c^v = \mathbf{I},
\end{aligned} \tag{10}$$

where $\mathbf{Y}_2 = \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{E}_r^v$. Equation (9) and (10) perform the *Wahba problem* [39], and according to [40], such problem can be solved by introducing Theorem 1.

Theorem 1. Given the optimal problem as follow:

$$\begin{aligned}
& \arg \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{W}\mathbf{X}\|_F^2 \\
s.t. \quad & \mathbf{X}^T \mathbf{X} = \mathbf{X} \mathbf{X}^T = \mathbf{I}
\end{aligned}$$

The optimal solution is $\mathbf{X} = \mathbf{U}\mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are left and right singular value matrices of $\mathbf{W}^T \mathbf{Y}$.

Based on Theorem 1, the optimal solutions w.r.t equation (9) and (10) are:

$$(\mathbf{P}_s^{v*})^T = \mathbf{U}_s \mathbf{V}_s^T, \tag{11}$$

$$(\mathbf{P}_c^{v*})^T = \mathbf{U}_c \mathbf{V}_c^T, \tag{12}$$

where \mathbf{U}_s and \mathbf{V}_s are the left and right singular value matrices of $\mathbf{H}_s^v (\mathbf{Y}_1)^T$, \mathbf{U}_c and \mathbf{V}_c are the left and right singular value matrices of $\mathbf{H}_c (\mathbf{Y}_2)^T$.

3.3.2. \mathbf{H}_s^v sub-problems

The sub-problems w.r.t \mathbf{H}_s^v are different for MVSC-CSLF and MVSC-CSLFS, which can be formulated as follows, respectively:

$$\mathbf{H}_s^{v*} = \arg \min_{\mathbf{H}_s^v} \left\| \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right\|_F^2 + \left\| \frac{1}{\mu} \bar{\Lambda}_2 + \mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z} - \bar{\mathbf{E}}_s \right\|_F^2, \quad (13)$$

where $\bar{\Lambda}_2 = \Lambda_2^{((v-1)K_s, vK_s, \cdot)}$, and $\bar{\mathbf{E}}_s = \mathbf{E}_s^{((v-1)K_s, vK_s, \cdot)}$.

$$\mathbf{H}_s^{v*} = \arg \min_{\mathbf{H}_s^v} \left\| \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right\|_F^2 + \left\| \frac{1}{\mu} \Lambda_2^v + \mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z}^v - \mathbf{E}_s^v \right\|_F^2. \quad (14)$$

Then let the derivatives of equations (13) and (14) w.r.t \mathbf{H}_s^v equal to zero and we obtain the formulas as follows:

$$\begin{aligned} \mathbf{A} \mathbf{H}_s^v + \mathbf{H}_s^v \mathbf{B} &= \mathbf{C} \\ \text{with } \mathbf{A} &= (\mathbf{P}_s^v)^T \mathbf{P}_s^v, \mathbf{B} = (\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})^T \\ \mathbf{C} &= (\mathbf{P}_s^v)^T \left(\frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right) \\ &\quad - \left(\frac{1}{\mu} \bar{\Lambda}_2 - \bar{\mathbf{E}}_s \right) (\mathbf{I} - \mathbf{Z})^T, \end{aligned} \quad (15)$$

$$\begin{aligned} \mathbf{A} \mathbf{H}_s^v + \mathbf{H}_s^v \mathbf{B} &= \mathbf{C} \\ \text{with } \mathbf{A} &= (\mathbf{P}_s^v)^T \mathbf{P}_s^v, \mathbf{B} = (\mathbf{I} - \mathbf{Z}^v)(\mathbf{I} - \mathbf{Z}^v)^T \\ \mathbf{C} &= (\mathbf{P}_s^v)^T \left(\frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right) \\ &\quad - \left(\frac{1}{\mu} \Lambda_2^v - \mathbf{E}_s^v \right) (\mathbf{I} - \mathbf{Z}^v)^T. \end{aligned} \quad (16)$$

Equation (15) and (16) are sylvester equations [41], which can be solved using *lyap*¹ function of a python package *control*².

3.3.3. \mathbf{H}_c sub-problems

The sub-problems w.r.t \mathbf{H}_c are different for MVSC-CSLF and MVSC-CSLFS, which can be formulated as follows:

$$\mathbf{H}_c^* = \arg \min_{\mathbf{H}_c} \sum_{v=1}^V \left\| \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right\|_F^2 + \left\| \frac{1}{\mu} \bar{\Lambda}_2 + \mathbf{H}_c - \mathbf{H}_c \mathbf{Z} - \bar{\mathbf{E}}_s \right\|_F^2, \quad (17)$$

where $\bar{\Lambda}_2 = \Lambda_2^{(VK_s+1:, \cdot)}$, $\bar{\mathbf{E}}_s = \mathbf{E}_s^{(VK_s+1:, \cdot)}$.

$$\mathbf{H}_c^* = \arg \min_{\mathbf{H}_c} \sum_{v=1}^V \left\| \frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v \right\|_F^2 + \left\| \frac{1}{\mu} \Lambda_3 + \mathbf{H}_c - \mathbf{H}_c \mathbf{Z}_c - \mathbf{E}_s^c \right\|_F^2. \quad (18)$$

Then let the derivatives of equations (17) and (18) w.r.t \mathbf{H}_c equal to zero and we obtain the formulas as follows:

$$\begin{aligned} \mathbf{A} \mathbf{H}_c + \mathbf{H}_c \mathbf{B} &= \mathbf{C} \\ \text{with } \mathbf{A} &= \sum_{v=1}^V (\mathbf{P}_c^v)^T \mathbf{P}_c^v, \mathbf{B} = (\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})^T \\ \mathbf{C} &= \sum_{v=1}^V (\mathbf{P}_c^v)^T \left(\frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{E}_r^v \right) \\ &\quad - \left(\frac{1}{\mu} \bar{\Lambda}_2 - \bar{\mathbf{E}}_s \right) (\mathbf{I} - \mathbf{Z})^T, \end{aligned} \quad (19)$$

$$\begin{aligned} \mathbf{A} \mathbf{H}_c + \mathbf{H}_c \mathbf{B} &= \mathbf{C} \\ \text{with } \mathbf{A} &= \sum_{v=1}^V (\mathbf{P}_c^v)^T \mathbf{P}_c^v, \mathbf{B} = (\mathbf{I} - \mathbf{Z}_c)(\mathbf{I} - \mathbf{Z}_c)^T \\ \mathbf{C} &= \sum_{v=1}^V (\mathbf{P}_c^v)^T \left(\frac{1}{\mu} \Lambda_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{E}_r^v \right) \\ &\quad - \left(\frac{1}{\mu} \Lambda_3 - \mathbf{E}_s^c \right) (\mathbf{I} - \mathbf{Z}_c)^T. \end{aligned} \quad (20)$$

Equation (19) and (20) are sylvester equations [41], which can be solved using *lyap* function of a python package *control*.

¹github.com/python-control/python-control/blob/master/control/mateqn.py

²github.com/python-control/python-control

3.3.4. \mathbf{Z} sub-problems

The sub-problems w.r.t \mathbf{Z} , \mathbf{Z}^v , and \mathbf{Z}_c for MVSC-CSLF and MVSC-CSLFS are formulated as:

$$\begin{aligned}\mathbf{Z}^* &= \arg \min_{\mathbf{Z}} \left\| \frac{1}{\mu} \mathbf{\Lambda}_2 + \mathbf{H} - \mathbf{H}\mathbf{Z} - \mathbf{E}_s \right\|_F^2 + \left\| \frac{1}{\mu} \mathbf{\Lambda}_3 + \mathbf{D} - \mathbf{Z} \right\|_F^2, \\ \mathbf{Z}^* &= \mathbf{Z}^* - \text{diag}(\mathbf{Z}^*),\end{aligned}\quad (21)$$

$$\begin{aligned}\mathbf{Z}^{v*} &= \arg \min_{\mathbf{Z}^v} \sum_{v=1}^V \left\| \frac{1}{\mu} \mathbf{\Lambda}_2^v + \mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z}^v - \mathbf{E}_s^v \right\|_F^2 + \sum_{v=1}^V \left\| \frac{1}{\mu} \mathbf{\Lambda}_4^v + \mathbf{D}^v - \mathbf{Z}^v \right\|_F^2, \\ \mathbf{Z}^{v*} &= \mathbf{Z}^{v*} - \text{diag}(\mathbf{Z}^{v*}),\end{aligned}\quad (22)$$

$$\begin{aligned}\mathbf{Z}_c^* &= \arg \min_{\mathbf{Z}_c} \left\| \frac{1}{\mu} \mathbf{\Lambda}_3 + \mathbf{H}_c - \mathbf{H}_c \mathbf{Z}_c - \mathbf{E}_s^c \right\|_F^2 + \left\| \frac{1}{\mu} \mathbf{\Lambda}_5 + \mathbf{D}_c - \mathbf{Z}_c \right\|_F^2, \\ \mathbf{Z}_c^* &= \mathbf{Z}_c^* - \text{diag}(\mathbf{Z}_c^*).\end{aligned}\quad (23)$$

Then let the derivatives of equations (21), (22), and (23) w.r.t \mathbf{Z} , \mathbf{Z}^v , and \mathbf{Z}_c equal to zero, and we obtain the corresponding updating formulas:

$$\begin{aligned}\mathbf{Z}^* &= (\mathbf{I} + \mathbf{H}^T \mathbf{H})^{-1} \cdot \left(\frac{1}{\mu} \mathbf{\Lambda}_3 + \frac{1}{\mu} \mathbf{H}^T \mathbf{\Lambda}_2 + \mathbf{D} - \mathbf{H}^T \mathbf{E}_s + \mathbf{H}^T \mathbf{H} \right) \\ \mathbf{Z}^* &= \mathbf{Z}^* - \text{diag}(\mathbf{Z}^*)\end{aligned}\quad (24)$$

$$\begin{aligned}\mathbf{Z}^{v*} &= \left[(\mathbf{H}_s^v)^T \mathbf{H}_s^v + \mathbf{I} \right]^{-1} \cdot \left(\frac{1}{\mu} \mathbf{\Lambda}_4^v + \frac{1}{\mu} (\mathbf{H}_s^v)^T \mathbf{\Lambda}_2^v + \mathbf{D}^v - (\mathbf{H}_s^v)^T \mathbf{E}_s^v + (\mathbf{H}_s^v)^T \mathbf{H}_s^v \right) \\ \mathbf{Z}^{v*} &= \mathbf{Z}^{v*} - \text{diag}(\mathbf{Z}^{v*})\end{aligned}\quad (25)$$

$$\begin{aligned}\mathbf{Z}_c^* &= (\mathbf{H}_c^T \mathbf{H}_c + \mathbf{I})^{-1} \cdot \left(\frac{1}{\mu} \mathbf{\Lambda}_5 + \frac{1}{\mu} \mathbf{H}_c^T \mathbf{\Lambda}_3 + \mathbf{D}_c - \mathbf{H}_c^T \mathbf{E}_s^c + \mathbf{H}_c^T \mathbf{H}_c \right) \\ \mathbf{Z}_c^* &= \mathbf{Z}_c^* - \text{diag}(\mathbf{Z}_c^*)\end{aligned}\quad (26)$$

3.3.5. \mathbf{D} sub-problems

We first give a brief account of the updating rules of \mathbf{D} and \mathbf{D}^v related to equations (7) and (8), respectively:

$$\mathbf{D}^* = \arg \min_{\mathbf{D}} \lambda_2 \|\mathbf{D}\|_F^2 + \mu \left\| \frac{1}{\mu} \mathbf{\Lambda}_3 + \mathbf{D} - \mathbf{Z} \right\|_F^2, \quad (27)$$

$$\mathbf{D}^{v*} = \arg \min_{\mathbf{D}^v} \lambda_2 \pi_2^v \|\mathbf{D}^v\|_F^2 + \mu \left\| \frac{1}{\mu} \mathbf{\Lambda}_4^v + \mathbf{D}^v - \mathbf{Z}^v \right\|_F^2. \quad (28)$$

Then let the derivatives of equations (27) and (28) w.r.t \mathbf{D} , \mathbf{D}^v equal to zero, and we obtain the corresponding updating formulas:

$$\mathbf{D}^* = \frac{\mu \mathbf{Z} - \mathbf{\Lambda}_3}{\lambda_2 + \mu}, \quad (29)$$

$$\mathbf{D}^{v*} = \frac{\mu \mathbf{Z}^v - \mathbf{\Lambda}_4^v}{\lambda_2 \pi_2^v + \mu}. \quad (30)$$

And for \mathbf{D}_c , the optimal problem is:

$$\mathbf{D}_c^* = \arg \min_{\mathbf{D}_c} \bar{\mu} \|\mathbf{D}_c\|_* + \frac{1}{2} \|\mathbf{D}_c - \mathbf{G}\|_F^2,$$

where $\bar{\mu} = \frac{\lambda_2 \pi_2^{V+1}}{\mu}$. Obviously, this kind of low-rank problem can be solved by singular value thresholding algorithm [42]:

$$\mathbf{D}_c^* = \Pi_{\bar{\mu}}(\mathbf{Z}_c - \frac{1}{\mu} \mathbf{\Lambda}_5) = \mathbf{U} \pi_{\bar{\mu}}(\mathbf{\Sigma}) \mathbf{V}^T, \quad (31)$$

where $\mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ is the skinny SVD of $\mathbf{Z}_c - \frac{1}{\mu} \mathbf{\Lambda}_5$. Soft-thresholding operation $\pi_{\tau}(\mathbf{\Sigma})$ is defined as $\pi_{\tau}(\mathbf{\Sigma}) = (|\mathbf{\Sigma}| - \tau)_+ \text{sgn}(\mathbf{\Sigma})$ and $(t)_+ = \max(0, t)$.

3.3.6. \mathbf{E} sub-problems

The updating rules of \mathbf{E}_r^v are the same for MVSC-CSLF and MCSC-CSLFS because $\boldsymbol{\pi} = \boldsymbol{\pi}_1$, which are replaced by $\boldsymbol{\pi}_{temp} = \{\pi_{temp}^v\}_{v=1}^V$ for convenience when discuss \mathbf{E}_r^v . Therefore, the optimal sub-problems w.r.t \mathbf{E}_r^v , \mathbf{E}_s , \mathbf{E}_s^v , and \mathbf{E}_s^c for MVSC-CSLF and MCSC-CSLFS are listed as follows:

$$\mathbf{E}_r^{v*} = \arg \min_{\mathbf{E}_r^v} \bar{\mu} \|\mathbf{E}_r^v\|_{2,1} + \frac{1}{2} \|\mathbf{E}_r^v - \mathbf{G}_r^v\|_F^2, \quad (32)$$

where $\bar{\mu}_r^v = \frac{\pi_{temp}^v}{\mu}$, and $\mathbf{G}_r^v = \frac{1}{\mu} \mathbf{\Lambda}_1^v + \mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c$.

$$\mathbf{E}_s^* = \arg \min_{\mathbf{E}_s} \bar{\mu} \|\mathbf{E}_s\|_{2,1} + \frac{1}{2} \|\mathbf{E}_s - \mathbf{G}_s\|_F^2, \quad (33)$$

where $\bar{\mu}_s = \frac{\lambda_1}{\mu}$, $\mathbf{G}_s = \frac{1}{\mu} \mathbf{\Lambda}_2 + \mathbf{H} - \mathbf{H}\mathbf{Z}$.

$$\mathbf{E}_s^{v*} = \arg \min_{\mathbf{E}_s^v} \bar{\mu} \|\mathbf{E}_s^v\|_{2,1} + \frac{1}{2} \|\mathbf{E}_s^v - \mathbf{G}_s^v\|_F^2, \quad (34)$$

where $\bar{\mu}_s^v = \frac{\lambda_1 \pi_2^v}{\mu}$, $\mathbf{G}_s^v = \frac{1}{\mu} \mathbf{\Lambda}_2^v + \mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z}^v$.

$$\mathbf{E}_s^{c*} = \arg \min_{\mathbf{E}_s^c} \bar{\mu} \|\mathbf{E}_s^c\|_{2,1} + \frac{1}{2} \|\mathbf{E}_s^c - \mathbf{G}_s^c\|_F^2, \quad (35)$$

where $\bar{\mu}_s^c = \frac{\lambda_1 \pi_2^{V+1}}{\mu}$, $\mathbf{G}_s^c = \frac{1}{\mu} \mathbf{\Lambda}_3 + \mathbf{H}_c - \mathbf{H}_c \mathbf{Z}_c$.

Equations (32) - (35) can be solved according to lemma 4.1 of [43], and their solutions are:

$$\mathbf{E}_r^{v*}(:,i) = \begin{cases} \frac{\|\mathbf{G}_r^{v(:,i)}\|_2 - \bar{\mu}_r^v}{\|\mathbf{G}_r^{v(:,i)}\|_2} \mathbf{G}_r^{v(:,i)}, & \text{if } \|\mathbf{G}_r^{v(:,i)}\|_2 > \bar{\mu}_r^v; \\ 0, & \text{otherwise.} \end{cases} \quad (36)$$

$$\mathbf{E}_s^*(:,i) = \begin{cases} \frac{\|\mathbf{G}_s^{(:,i)}\|_2 - \bar{\mu}_s}{\|\mathbf{G}_s^{(:,i)}\|_2} \mathbf{G}_s^{(:,i)}, & \text{if } \|\mathbf{G}_s^{(:,i)}\|_2 > \bar{\mu}_s; \\ 0, & \text{otherwise.} \end{cases} \quad (37)$$

$$\mathbf{E}_s^{v*}(:,i) = \begin{cases} \frac{\|\mathbf{G}_s^{v(:,i)}\|_2 - \bar{\mu}_s^v}{\|\mathbf{G}_s^{v(:,i)}\|_2} \mathbf{G}_s^{v(:,i)}, & \text{if } \|\mathbf{G}_s^{v(:,i)}\|_2 > \bar{\mu}_s^v; \\ 0, & \text{otherwise.} \end{cases} \quad (38)$$

$$\mathbf{E}_s^{c*}(:,i) = \begin{cases} \frac{\|\mathbf{G}_s^{c(:,i)}\|_2 - \bar{\mu}_s^c}{\|\mathbf{G}_s^{c(:,i)}\|_2} \mathbf{G}_s^{c(:,i)}, & \text{if } \|\mathbf{G}_s^{c(:,i)}\|_2 > \bar{\mu}_s^c; \\ 0, & \text{otherwise.} \end{cases} \quad (39)$$

3.3.7. $\boldsymbol{\pi}$ sub-problems

For MVSC-CSLF and MVSC-CSLFS, the optimal sub-problems w.r.t $\boldsymbol{\pi}$, $\boldsymbol{\pi}_1$, and $\boldsymbol{\pi}_2$ are listed as follows:

$$\begin{aligned} & \min \sum_{v=1}^V \pi^v \|\mathbf{E}_r^v\|_{2,1} + \frac{\lambda_3}{2} \|\boldsymbol{\pi}\|_2^2 \\ & s.t. \boldsymbol{\pi} \geq 0, \sum_{v=1}^V \pi^v = 1, \end{aligned} \quad (40)$$

$$\begin{aligned} & \min \sum_{v=1}^V \pi_1^v \|\mathbf{E}_r^v\|_{2,1} + \frac{\lambda_4}{2} \|\boldsymbol{\pi}_1\|_2^2 \\ & s.t. \boldsymbol{\pi}_1 \geq 0, \sum_{v=1}^V \pi_1^v = 1, \end{aligned} \quad (41)$$

$$\begin{aligned} & \min \sum_{v=1}^V \pi_2^v (\lambda_1 \|\mathbf{E}_s^v\|_{2,1} + \frac{\lambda_2}{2} \|\mathbf{D}^v\|_F^2) + \pi_2^{V+1} (\lambda_1 \|\mathbf{E}_s^c\|_{2,1} + \lambda_2 \|\mathbf{D}_c\|_*) + \frac{\lambda_4}{2} \|\boldsymbol{\pi}_2\|_2^2 \\ & s.t. \boldsymbol{\pi}_2 \geq 0, \sum_{v=1}^{V+1} \pi_2^v = 1. \end{aligned} \quad (42)$$

They are standard convex optimization problems, which can be solved by the python package CVXPY³.

³www.cvxpy.org/

Algorithm 1 MVSC-CSLF

Input:

$$\{\mathbf{X}^v\}_{v=1}^V, K_s, K_c, \lambda_1, \lambda_2, \lambda_3, \mu, \mu_{max}, \rho, \epsilon.$$

Output:

\mathbf{Z} and \mathbf{H} .

- 1: **Initialize:** Set the initial values of $\{\mathbf{P}_s^v\}_{v=1}^V, \{\mathbf{P}_c^v\}_{v=1}^V, \{\mathbf{H}_s^v\}_{v=1}^V, \mathbf{H}_c, \mathbf{Z}, \mathbf{D}, \{\mathbf{E}_r^v\}_{v=1}^V, \mathbf{E}_s$ equal to 0, and let $\{\pi^v = \frac{1}{V}\}_{v=1}^V$.
 - 2: **while** not converged **do**
 - 3: fix others and update $\{\mathbf{P}_s^v\}_{v=1}^V, \{\mathbf{P}_c^v\}_{v=1}^V, \{\mathbf{H}_s^v\}_{v=1}^V, \mathbf{H}_c, \mathbf{Z}, \mathbf{D}, \{\mathbf{E}_r^v\}_{v=1}^V, \{\mathbf{E}_s\}_{v=1}^V, \pi$, and lagrangian multipliers by solving (11), (12), (15), (19), (24), (29), (36), (37), (40), and (43) in turn;
 - 4: update $\mu = \min(\rho\mu, \mu_{max})$
 - 5: check convergence condition:

$$\frac{1}{V} \sum_{v=1}^V \|\mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v\|_{\infty} < \epsilon,$$

$$\|\mathbf{H} - \mathbf{H}\mathbf{Z} - \mathbf{E}_s\|_{\infty} < \epsilon, \|\mathbf{D} - \mathbf{Z}\|_{\infty} < \epsilon.$$
 - 6: **end while**
-

Algorithm 2 MVSC-CSLFS

Input:

$$\{\mathbf{X}^v\}_{v=1}^V, K_s, K_c, \lambda_1, \lambda_2, \lambda_3, \mu, \mu_{max}, \rho, \epsilon.$$

Output:

\mathbf{Z} and \mathbf{H} .

- 1: **Initialize:** Set the initial values of $\{\mathbf{P}_s^v\}_{v=1}^V, \{\mathbf{P}_c^v\}_{v=1}^V, \{\mathbf{H}_s^v\}_{v=1}^V, \mathbf{H}_c, \{\mathbf{Z}^v\}_{v=1}^V, \mathbf{Z}_c, \{\mathbf{D}^v\}_{v=1}^V, \mathbf{D}_c, \{\mathbf{E}_r^v\}_{v=1}^V, \{\mathbf{E}_s^c\}_{v=1}^V, \mathbf{E}_s^c$ equal to 0, let $\{\pi_1^v = \frac{1}{V}\}_{v=1}^V$ and $\{\pi_2^v = \frac{1}{V+1}\}_{v=1}^{V+1}$.
 - 2: **while** not converged **do**
 - 3: fix others and update $\{\mathbf{P}_s^v\}_{v=1}^V, \{\mathbf{P}_c^v\}_{v=1}^V, \{\mathbf{H}_s^v\}_{v=1}^V, \mathbf{H}_c, \{\mathbf{Z}^v\}_{v=1}^V, \mathbf{Z}_c, \{\mathbf{D}^v\}_{v=1}^V, \mathbf{D}_c, \{\mathbf{E}_r^v\}_{v=1}^V, \{\mathbf{E}_s^c\}_{v=1}^V, \{\mathbf{E}_s\}_{v=1}^V, \pi_1, \pi_2$, and lagrangian multipliers by solving (11), (12), (16), (20), (25), (26), (30), (31), (36), (38), (39), (41), (42), and (44) in turn;
 - 4: update $\mu = \min(\rho\mu, \mu_{max})$
 - 5: check convergence condition:

$$\frac{1}{V} \sum_{v=1}^V \|\mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v\|_{\infty} < \epsilon,$$

$$\frac{1}{V} \sum_{v=1}^V \|\mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z}^v - \mathbf{E}_s^v\|_{\infty} < \epsilon,$$

$$\|\mathbf{H}_c - \mathbf{H}_c \mathbf{Z}_c - \mathbf{E}_s^c\|_{\infty} < \epsilon, \frac{1}{V} \sum_{v=1}^V \|\mathbf{D}^v - \mathbf{Z}^v\|_{\infty} < \epsilon,$$

$$\|\mathbf{D}_c - \mathbf{Z}_c\|_{\infty} < \epsilon.$$
 - 6: **end while**
-

3.3.8. Lagrangian multipliers sub-problems

The lagrangian multipliers can be updated by the rules:

$$\begin{cases} \Lambda_1^v = \Lambda_1^v + \mu(\mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v) \\ \Lambda_2 = \Lambda_2 + \mu(\mathbf{H} - \mathbf{H}\mathbf{Z} - \mathbf{E}_s) \\ \Lambda_3 = \Lambda_3 + \mu(\mathbf{D} - \mathbf{Z}) \end{cases} \quad (43)$$

$$\begin{cases} \Lambda_1^v = \Lambda_1^v + \mu(\mathbf{X}^v - \mathbf{P}_s^v \mathbf{H}_s^v - \mathbf{P}_c^v \mathbf{H}_c - \mathbf{E}_r^v) \\ \Lambda_2^v = \Lambda_2^v + \mu(\mathbf{H}_s^v - \mathbf{H}_s^v \mathbf{Z}^v - \mathbf{E}_s^v) \\ \Lambda_3 = \Lambda_3 + \mu(\mathbf{H}_c - \mathbf{H}_c \mathbf{Z}_c - \mathbf{E}_s^c) \\ \Lambda_4^v = \Lambda_4^v + \mu(\mathbf{D}^v - \mathbf{Z}^v) \\ \Lambda_5 = \Lambda_5 + \mu(\mathbf{D}_c - \mathbf{Z}_c) \end{cases} \quad (44)$$

The complete algorithms w.r.t MVSC-CSLF and MVSC-CSLFS are summarized in Algorithms 1 and 2, respectively.

3.4. Complexity and convergence analysis

Computational complexity of the algorithm mainly comes from some complex computations, such as matrix inversion and svd decomposition, and therefore we mainly consider these operations in this subsection.

3.4.1. MVSC-CSLF

The computational complexity mainly comes from five sources. For $\{\mathbf{P}_s^v\}_{v=1}^V$ and $\{\mathbf{P}_c^v\}_{v=1}^V$, the complexities are $O(\sum_{v=1}^V (M^v)^3)$; for $\{\mathbf{H}_s^v\}_{v=1}^V$ and \mathbf{H}_c , the complexities are $O(VK_s^3)$ and $O(K_c^3)$; for \mathbf{Z} , the matrix inversion's complexity is $O(N^3)$. Therefore, the total computational cost of this problem is $O(\sum_{v=1}^V (M^v)^3 + VK_s^3 + K_c^3 + N^3)$.

3.4.2. MVSC-CSLFS

The computational complexity mainly comes from seven sources. For $\{\mathbf{P}_s^v\}_{v=1}^V$ and $\{\mathbf{P}_c^v\}_{v=1}^V$, the complexities are $O(\sum_{v=1}^V (M^v)^3)$; for $\{\mathbf{H}_s^v\}_{v=1}^V$, \mathbf{H}_c , the complexities are $O(VK_s^3)$ and $O(K_c^3)$; for $\{\mathbf{Z}^v\}_{v=1}^V$, \mathbf{Z}_c , the complexities are $O(VN^3)$ and $O(N^3)$; for \mathbf{D}_c , the complexity is $O(N^3)$. Therefore, the total computational cost of this problem is $O(\sum_{v=1}^V (M^v)^3 + VK_s^3 + K_c^3 + (V+2)N^3)$.

Mostly, K_s and K_c should smaller than M^v , i.e., $K_s, K_c \ll M^v$, therefore the computational cost of our proposed two algorithms can be rewritten as $O(\sum_{v=1}^V (M^v)^3 + N^3)$ and $O(\sum_{v=1}^V (M^v)^3 + (V+2)N^3)$. And for large scale data, i.e. $M^v \ll N$, the computational cost are $O(N^3)$ and $O((V+2)N^3)$, respectively.

3.4.3. Convergence analysis

MVSC-CSLF and MVSC-CSLFS both involve multiple sub-problems, which makes it hard to give a solid proof of their convergence. However, we can give an empirical demonstration by extensive experiments on multiple real-world datasets referred to [12, 14, 13]. The convergence curves of MVSC-CSLF and MVSC-CSLFS on six datasets are shown in Fig. 2 and Fig. 3, respectively. The promising performance indicates that the proposed methods converge rate fast and stable.

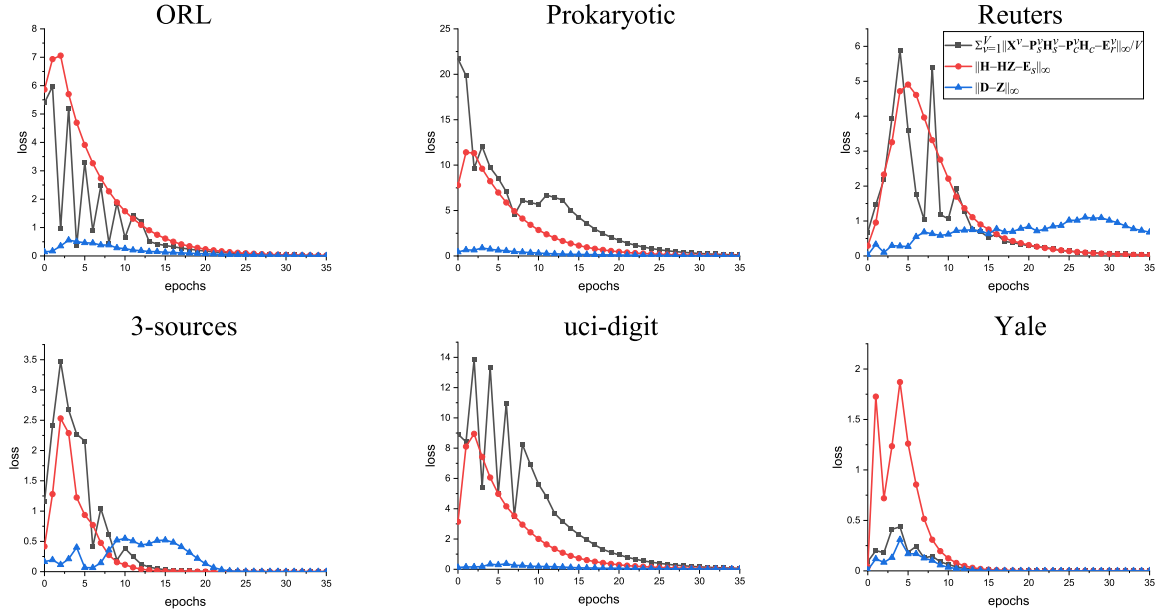


Figure 2: Convergence of MVSC-CSLFS. We display the curves about three stop criteria in red, gray, and blue versus the iteration numbers on six datasets.

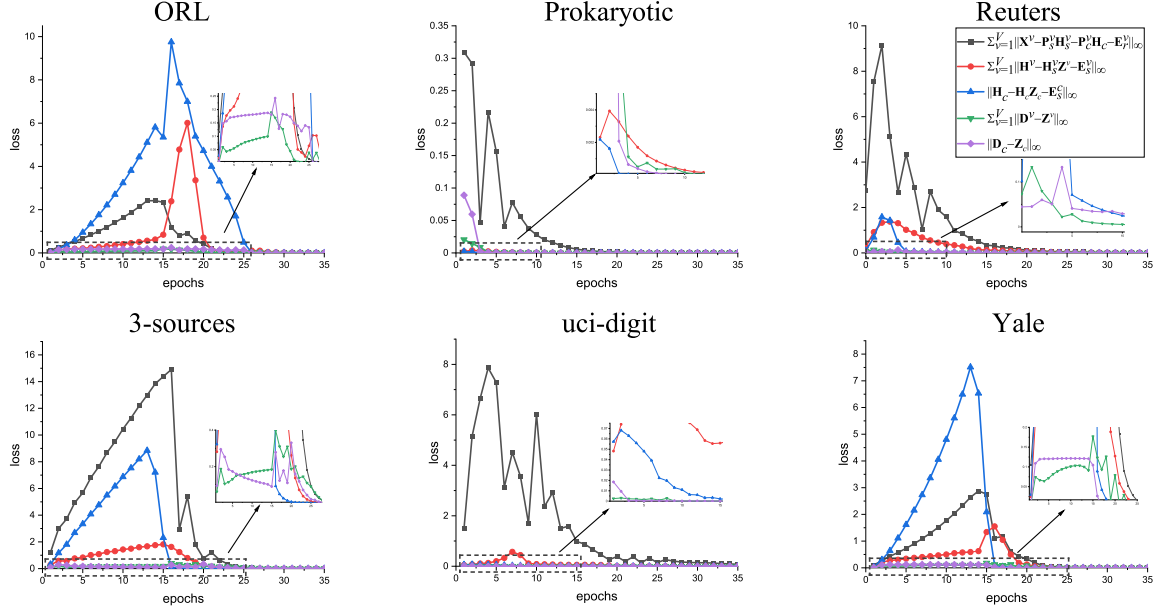


Figure 3: Convergence of MVSC-CSLFS. We display the curves about five stop criteria in gray, red, blue, green, and purple versus the iteration numbers on six datasets.

4. Experiment

4.1. datasets

Six real-world datasets are selected to test and validate the performance of proposed algorithms. Among them, **ORL**⁴ and **Yale**⁴ are human face data, and we select intensity, LBP, and Gabor features as three views; **Prokaryotic**[44] is a species data set described from textual data, proteome composition, and genomic representations, each of them is viewed as a view; **Reuters**[45] is document data that the contents are written by 6 languages, each language is viewed as a view, and they are all in the bag-of-words representations; **3-sources**⁵ is news data set with its contents come from three sources (BBC, Reuters, and the Guardian), and they are all in the bag-of-words representations; **uci-digit**⁶ concludes six types of features of handwritten number, and 3 of them are utilized to construct three views. Table 2 summarizes the above information.

Table 2: Description of datasets

Data	View number (V)	Clusters (C)	Instance number (N)	Dimension of each view (M^v)
ORL	3	40	400	[4096,3304,6750]
Reuters	5	6	600	[21526,24892,34121,15487,11539]
3-sources	3	6	169	[3560,3631,3068]
Yale	3	15	165	[4096,3304,6750]
uci-digit	3	10	2000	[216,76,64]
Prokaryotic	3	4	551	[438,3,393]

⁴<http://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html>

⁵<http://mlg.ucd.ie/datasets/3sources.html>

⁶<http://archive.ics.uci.edu/ml/datasets/Multiple+Features>

4.2. Compared Methods

In comprising experiment, 9 algorithms are selected as baseline methods:

- (1) **LRSC** [27]: the method impose a low-rank constraint on subspace clustering. The algorithm has four variants depending on whether the data is corrupted and whether the constraint is relaxed or exact. Since LRSC is a single-view method, we perform each variant on each view’s data and select the best result.
- (2) **LMSC** [12]: the method combines representation learning with subspace clustering. It learn a common latent representation first and then perform self-expressive reconstruction on it with a low-rank constraint.
- (3) **CSMSC** [10]: the method assumes that each view’s subspace representation includes two parts: consistent and view-specific ones. It separates them from each view, and constraints them with nuclear and Frobenius norm, respectively.
- (4) **MLRSSC** [9]: the method performs self-expressive reconstruction on each view, but impose sparse and low-rank constraints on each subspace representation, simultaneously.
- (5) **MLRSSC-C** [9]: the method is based on MLRSSC, but propose a centroid alignment strategy.
- (6) **KMLRSSC** [9]: the method is based on MLRSSC, and add kernel to enhance the processing ability of nonlinear data.
- (7) **KMLRSSC-C** [9]: the method is based on MLRSSC-C, and add kernel to enhance the processing ability of non-linear data.
- (8) **CSI** [13]: the method is differ from most existing algorithms when integrating multiple subspace representations. CSI propose a strategy that can obtain a common subspace representation by keeping each view’s local manifold preserving.
- (9) **FCMSC** [14]: the method concatenates all views and performs self-expressive reconstruction on it with low-rank and local manifold preserving regularization.

4.3. Evaluation Metrics and Adjacency Matrix

We use ACC (accuracy), NMI (normalized mutual information), ARI (Adjusted Rand Index), Precision, Recall, F-score as the evaluation metrics. Among them, except for NMI and ARI, other metrics are not clustering metrics. Therefore the predicted labels should be first mapped into the true categories by applying Kuhn-Munkres algorithm. Note that a higher value indicates better performance for all the above metrics.

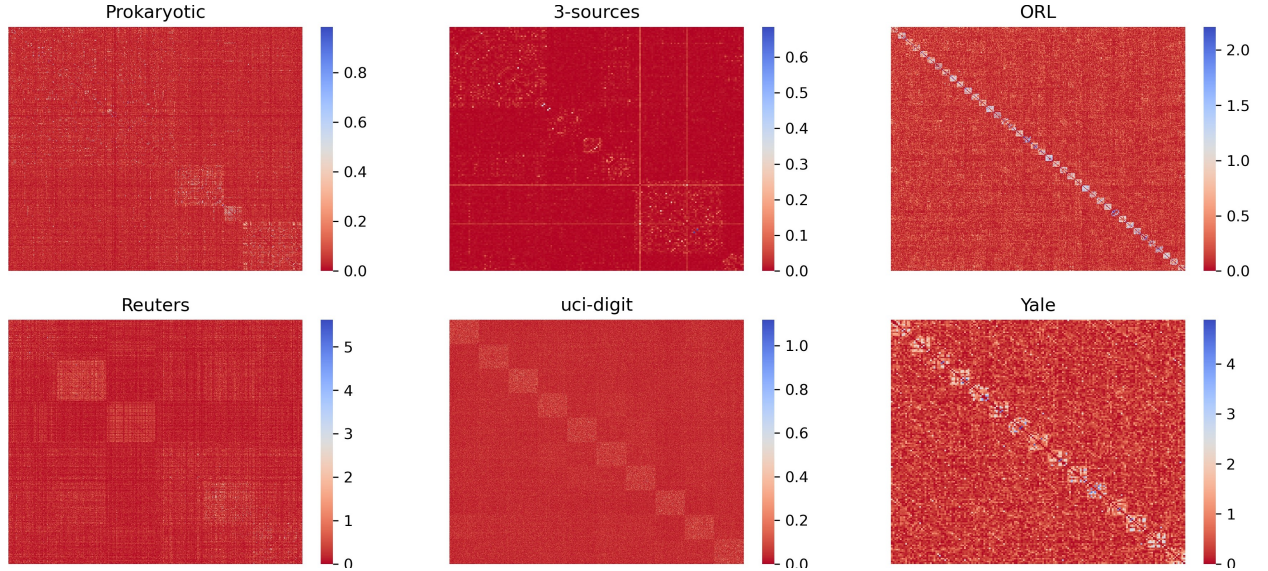


Figure 4: Visualization for MVSC-CSLF’s adjacency matrices on six datasets.

Once model training is completed, the learned subspace representation can be utilized to construct adjacency matrix, which can be applied to the spectral clustering [46]. MVSC-CSLF constructs adjacency matrix according to

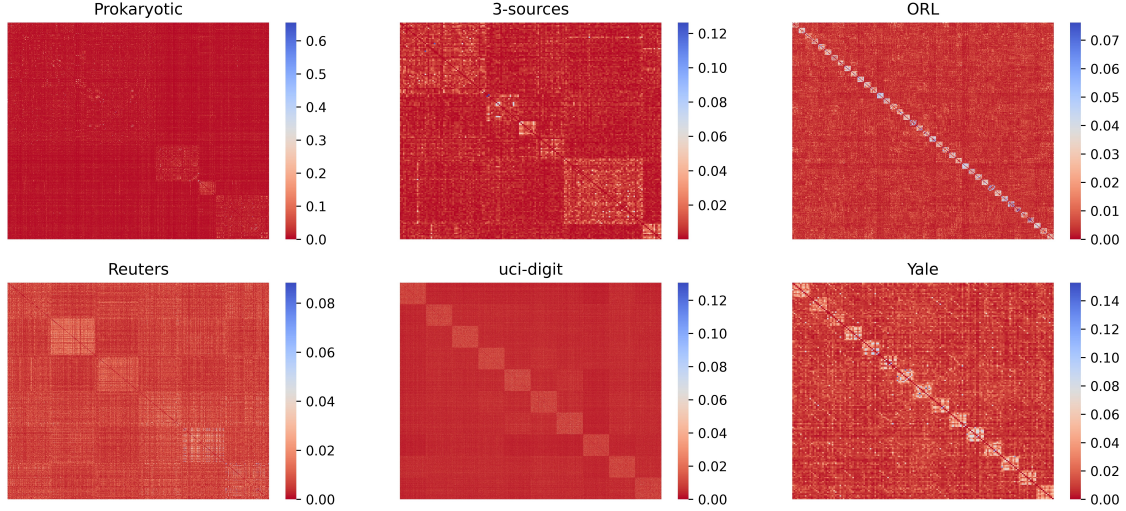


Figure 5: Visualization for MVSC-CSLFS's adjacency matrices on six datasets.

the following rule:

$$\frac{|\mathbf{Z}| + |\mathbf{Z}^T|}{2}. \quad (45)$$

MVSC-CSLFS has $V + 1$ coefficient matrix, and its construction rule is:

$$\frac{1}{2} \left(\frac{|\mathbf{Z}_c| + |\mathbf{Z}_c^T|}{2} + \frac{1}{V} \sum_{v=1}^V \frac{|\mathbf{Z}^v| + |(\mathbf{Z}^v)^T|}{2} \right), \quad (46)$$

where $|\cdot|$ denotes element-wise absolute operator.

Due to the higher correlation between samples belonging to the same category, the adjacency matrix should present block diagonalization when we sort samples by category. Therefore, we visualize the adjacency matrices of MVSC-CSLF and MVSC-CSLFS computed thorough equations (45) and (46) on six datasets, which are shown in Fig. 4 and Fig. 5, respectively. These figures show that the adjacency matrices present obvious block diagonalization according to category on all datasets. The property is more clearly expressed in MVSC-CSLFS than in MVSC-CSLF, which is one of the reasons that MVSC-CSLFS performs better in most cases.

4.4. Clustering Performance on Real-World datasets

In this subsection, we compare our proposed methods with 9 baseline approaches, and Table 3 tabulates the clustering results on six datasets in terms of six metrics. Note that we conduct each method 10 times, and the results are reported in form of the mean score, as well as the standard deviation.

The results in Table 3 show that multi-view algorithms perform better than single view algorithm on all datasets. As refereed in subsection 4.2, LRSC's four variants are conducted on all views of each datasets, and the best result is selected as comparison. Even though, LRSC performs worse than almost all multi-view algorithms, and the clustering results are difficult to find a trade-off between Precision and Recall on Reuters and Prokaryotic. Furthermore, it is not hard to see that our proposed MVSC-CSLF and MVSC-CSLFS have satisfied performance on all datasets. In particular, MVSC-CSLFS outperforms than others on ORL and Yale in terms of all metrics, and also has best score on uci-digit and Prokaryotic in terms of three and four metrics, respectively. On Reuters, MVSC-CSLF performs best in terms of ACC and F-score, and since LRSC cannot find a good trade-off between Recall and F-score on this data, MVSC-CSLF actually also has the best score in terms of Recall. As for 3-sources, FSCMC performs best, but expect for this method our proposed approaches have a bigger improvement that other compared algorithms.

Table 3: Clustering results

Dataset	Method	ACC	NMI	AR	Precision	Recall	F-score
ORL	LRSC	0.769(0.015)	0.883(0.015)	0.688(0.037)	0.668(0.037)	0.726(0.036)	0.695(0.036)
	LMSC	0.822(0.037)	0.927(0.013)	0.774(0.043)	0.813(0.038)	0.822(0.037)	0.802(0.041)
	CSMSC	0.868(0.0012)	0.942(0.005)	0.827(0.002)	0.860(0.002)	0.804(0.003)	0.831(0.001)
	MLRSSC	0.637(0.034)	0.813(0.017)	0.524(0.037)	0.640(0.033)	0.637(0.034)	0.623(0.034)
	MLRSSC-Centroid	0.780(0.027)	0.917(0.009)	0.729(0.029)	0.785(0.033)	0.780(0.027)	0.761(0.032)
	KMLRSSC	0.786(0.041)	0.903(0.016)	0.721(0.042)	0.793(0.040)	0.786(0.041)	0.774(0.042)
	KMLRSSC-Centroid	0.783(0.031)	0.907(0.008)	0.721(0.026)	0.793(0.039)	0.783(0.031)	0.772(0.035)
	CSI	0.863(0.152)	0.930(0.095)	0.805(0.158)	0.782(0.170)	0.840(0.141)	0.810(0.158)
	FCSMC	0.847(0.019)	0.928(0.008)	0.782(0.021)	0.746(0.025)	0.833(0.020)	0.787(0.020)
	MVSC-CSLF	0.849(0.009)	0.931(0.006)	0.800(0.012)	0.854(0.009)	0.849(0.009)	0.837(0.010)
	MVSC-CSLFS	0.885(0.007)	0.947(0.005)	0.842(0.011)	0.889(0.008)	0.885(0.007)	0.874(0.007)
Reuters	LRSC	0.209(0.007)	0.131(0.007)	0.015(0.001)	0.172(0.000)	0.835(0.013)	0.285(0.001)
	LMSC	0.495(0.001)	0.355(0.001)	0.223(0.010)	0.312(0.009)	0.497(0.008)	0.383(0.006)
	CSMSC	0.513(0.001)	0.358(0.001)	0.247(0.001)	0.343(0.001)	0.444(0.001)	0.387(0.001)
	MLRSSC	0.530(0.033)	0.382(0.015)	0.282(0.026)	0.501(0.055)	0.530(0.033)	0.483(0.051)
	MLRSSC-Centroid	0.514(0.035)	0.370(0.011)	0.268(0.026)	0.484(0.040)	0.514(0.035)	0.459(0.046)
	KMLRSSC	0.571(0.023)	0.374(0.019)	0.304(0.020)	0.618(0.037)	0.571(0.023)	0.567(0.030)
	KMLRSSC-Centroid	0.551(0.024)	0.357(0.016)	0.294(0.016)	0.591(0.046)	0.551(0.024)	0.540(0.027)
	CSI	0.527(0.009)	0.394(0.009)	0.288(0.007)	0.364(0.006)	0.511(0.008)	0.425(0.007)
	FCSMC	0.518(0.001)	0.369(0.001)	0.268(0.000)	0.357(0.000)	0.470(0.001)	0.405(0.000)
	MVSC-CSLF	0.582(0.010)	0.352(0.004)	0.290(0.012)	0.604(0.005)	0.582(0.005)	0.581(0.012)
	MVSC-CSLFS	0.571(0.001)	0.373(0.001)	0.297(0.01)	0.419(0.002)	0.397(0.001)	0.444(0.002)
Yale	LRSC	0.675(0.027)	0.706(0.019)	0.491(0.032)	0.501(0.030)	0.547(0.030)	0.523(0.030)
	LMSC	0.789(0.018)	0.789(0.019)	0.628(0.030)	0.632(0.029)	0.672(0.027)	0.651(0.028)
	CSMSC	0.752(0.001)	0.784(0.001)	0.615(0.005)	0.673(0.002)	0.610(0.006)	0.794(0.029)
	MLRSSC	0.660(0.050)	0.697(0.033)	0.477(0.054)	0.482(0.056)	0.544(0.045)	0.511(0.050)
	MLRSSC-Centroid	0.627(0.034)	0.702(0.021)	0.458(0.024)	0.711(0.050)	0.627(0.034)	0.648(0.041)
	KMLRSSC	0.649(0.053)	0.689(0.032)	0.485(0.046)	0.679(0.070)	0.649(0.053)	0.653(0.059)
	KMLRSSC-Centroid	0.639(0.043)	0.680(0.032)	0.480(0.044)	0.661(0.046)	0.639(0.043)	0.640(0.045)
	CSI	0.734(0.000)	0.777(0.000)	0.606(0.000)	0.606(0.000)	0.659(0.000)	0.739(0.000)
	FCSMC	0.775(0.025)	0.796(0.019)	0.589(0.041)	0.569(0.046)	0.674(0.028)	0.617(0.037)
	MVSC-CSLF	0.804(0.008)	0.813(0.013)	0.629(0.022)	0.874(0.010)	0.804(0.008)	0.828(0.009)
	MVSC-CSLFS	0.859(0.017)	0.861(0.016)	0.727(0.024)	0.902(0.018)	0.859(0.017)	0.853(0.006)
uci-digit	LRSC	0.763(0.003)	0.697(0.001)	0.626(0.003)	0.650(0.003)	0.678(0.002)	0.664(0.003)
	LMSC	0.905(0.004)	0.833(0.006)	0.812(0.008)	0.835(0.007)	0.840(0.007)	0.838(0.007)
	CSMSC	0.903(0.000)	0.835(0.001)	0.810(0.000)	0.823(0.000)	0.834(0.004)	0.828(0.000)
	MLRSSC	0.881(0.061)	0.852(0.022)	0.814(0.050)	0.877(0.070)	0.881(0.061)	0.873(0.070)
	MLRSSC-Centroid	0.893(0.055)	0.853(0.023)	0.818(0.050)	0.891(0.063)	0.893(0.055)	0.887(0.063)
	KMLRSSC	0.894(0.049)	0.861(0.018)	0.826(0.044)	0.888(0.060)	0.894(0.049)	0.886(0.058)
	KMLRSSC-Centroid	0.910(0.046)	0.865(0.018)	0.840(0.039)	0.910(0.051)	0.910(0.046)	0.906(0.053)
	CSI	0.863(0.000)	0.915(0.000)	0.846(0.000)	0.805(0.000)	0.929(0.000)	0.862(0.000)
	FCSMC	0.916(0.001)	0.848(0.001)	0.827(0.001)	0.840(0.001)	0.849(0.001)	0.844(0.001)
	MVSC-CSLF	0.913(0.000)	0.834(0.000)	0.818(0.001)	0.917(0.000)	0.913(0.000)	0.913(0.000)
	MVSC-CSLFS	0.917(0.002)	0.852(0.001)	0.828(0.004)	0.921(0.000)	0.917(0.000)	0.917(0.000)
Prokaryotic	LRSC	0.582(0.005)	0.079(0.011)	0.038(0.029)	0.407(0.011)	0.964(0.047)	0.571(0.001)
	LMSC	0.709(0.032)	0.418(0.034)	0.394(0.052)	0.669(0.067)	0.585(0.068)	0.618(0.026)
	CSMSC	0.658(0.006)	0.352(0.004)	0.367(0.005)	0.665(0.005)	0.527(0.007)	0.588(0.004)
	MLRSSC	0.646(0.038)	0.309(0.018)	0.324(0.037)	0.529(0.011)	0.531(0.006)	0.498(0.005)
	MLRSSC-Centroid	0.620(0.009)	0.196(0.008)	0.260(0.005)	0.444(0.033)	0.399(0.018)	0.363(0.024)
	KMLRSSC	0.638(0.045)	0.420(0.041)	0.355(0.071)	0.609(0.032)	0.641(0.053)	0.589(0.041)
	KMLRSSC-Centroid	0.655(0.045)	0.405(0.030)	0.350(0.072)	0.606(0.031)	0.657(0.046)	0.597(0.042)
	CSI	0.619(0.000)	0.413(0.000)	0.291(0.000)	0.634(0.000)	0.440(0.000)	0.520(0.000)
	FCSMC	0.674(0.001)	0.443(0.000)	0.394(0.001)	0.688(0.001)	0.537(0.001)	0.603(0.001)
	MVSC-CSLF	0.763(0.001)	0.449(0.004)	0.521(0.004)	0.614(0.001)	0.596(0.007)	0.600(0.002)
	MVSC-CSLFS	0.809(0.001)	0.489(0.002)	0.568(0.001)	0.617(0.001)	0.635(0.001)	0.624(0.001)
3-sources	LRSC	0.596(0.026)	0.462(0.009)	0.351(0.020)	0.471(0.037)	0.585(0.054)	0.518(0.004)
	LMSC	0.734(0.020)	0.676(0.018)	0.579(0.017)	0.729(0.024)	0.621(0.047)	0.669(0.018)
	CSMSC	0.798(0.002)	0.747(0.005)	0.673(0.004)	0.730(0.005)	0.773(0.001)	0.751(0.003)
	MLRSSC	0.677(0.060)	0.593(0.026)	0.550(0.067)	0.535(0.048)	0.596(0.060)	0.544(0.059)
	MLRSSC-Centroid	0.666(0.051)	0.590(0.019)	0.534(0.053)	0.512(0.045)	0.551(0.068)	0.513(0.056)
	KMLRSSC	0.609(0.038)	0.526(0.027)	0.423(0.035)	0.536(0.59)	0.564(0.079)	0.517(0.060)
	KMLRSSC-Centroid	0.621(0.025)	0.531(0.020)	0.457(0.031)	0.544(0.057)	0.562(0.051)	0.519(0.046)
	CSI	0.763(0.000)	0.747(0.000)	0.622(0.000)	0.717(0.000)	0.702(0.000)	0.709(0.000)
	FCSMC	0.908(0.003)	0.795(0.006)	0.793(0.008)	0.887(0.004)	0.795(0.007)	0.839(0.006)
	MVSC-CSLF	0.843(0.016)	0.696(0.014)	0.659(0.034)	0.810(0.013)	0.873(0.007)	0.833(0.012)
	MVSC-CSLFS	0.894(0.000)	0.754(0.000)	0.793(0.000)	0.840(0.000)	0.860(0.000)	0.821(0.000)

It is worth noting that MVSC-CSLF, which only has a pretty simple constraint strategy on coefficient matrix, performs better than many of compared methods on most datasets. The promising performance indicates that well distributed representation is significant for subspace clustering, and when obtained suitable latent representation, even a simple form can derive satisfactory clustering results. However, such strategy cannot make full use of the obtained representations, therefore the proposed MVSC-CSLFS which has advanced trick can further make a significant improvement on clustering performance.

4.5. Ablation Study: Is the latent representation good?

The core assumption in this study is that the proposed MVSC-CSLF and MVSC-CSLFS perform well because they have good representations, and thus we design an ablation experiment to demonstrate this. First of all, MVSC-CSLFS is employed on ORL and Yale to calculate the latent representations: \mathbf{H}_s^1 , \mathbf{H}_s^2 , \mathbf{H}_s^3 , \mathbf{H}_c , and \mathbf{H} . Among them, the first three ones are view-specific representations, \mathbf{H}_c is consistent representation of all views, and \mathbf{H} is the joint latent representation concatenated by the first four ones. Then LRSC [27] (a single view subspace clustering algorithm, and we use its first variant: uncorrupted data and relaxed constraints) is employed to test the performance on both raw data and latent representations in terms of three major clustering metrics: ACC, NMI, and ARI. As shown in Table 4, LRSC is conducted on 3 views' raw data, 3 views' specific, consistent, and joint representations, respectively. And the results obviously indicate that the latent representations have made a significant improvement compared with raw data. Note that joint representation may not bring a better improvement, and this is why we further propose an advanced strategy to improve on the basis of MVSC-CSLF. However, the ability of joint representation is also satisfied and MVSC-CSLF only has one coefficient matrix that can reduce computing and storage cost. Therefore, how to choose the two proposed algorithms should be based on the actual situation and requirement.

Table 4: Comparison of the clustering performance on raw data and latent representation

Data	Method	ACC	NMI	ARI
ORL	view1	0.691(0.021)	0.834(0.009)	0.567(0.022)
	view2	0.769(0.015)	0.883(0.015)	0.688(0.037)
	view3	0.560(0.018)	0.747(0.010)	0.418(0.021)
	view1-specific	0.855(0.015)	0.933(0.005)	0.809(0.013)
	view2-specific	0.847(0.020)	0.928(0.005)	0.787(0.013)
	view3-specific	0.815(0.016)	0.910(0.009)	0.744(0.021)
	consistent	0.847(0.017)	0.928(0.007)	0.788(0.017)
	joint	0.836(0.026)	0.923(0.010)	0.775(0.033)
Yale	view1	0.675(0.027)	0.706(0.019)	0.491(0.032)
	view2	0.587(0.015)	0.597(0.009)	0.371(0.013)
	view3	0.508(0.020)	0.571(0.020)	0.310(0.026)
	view1-specific	0.726(0.011)	0.736(0.009)	0.525(0.016)
	view2-specific	0.715(0.018)	0.734(0.010)	0.544(0.013)
	view3-specific	0.733(0.005)	0.741(0.003)	0.544(0.007)
	consistent	0.718(0.009)	0.713(0.011)	0.516(0.016)
	joint	0.718(0.018)	0.733(0.013)	0.537(0.019)

4.6. Hyper-parameters Tuning Strategy

Both MVSC-CSLF and MVSC-CSLFS have five hyperparameters, where λ_1 , λ_2 , and λ_3 are trade-off parameters of different loss and penalty terms, K_s and K_c are view-specific and consistent latent representations' dimensions, respectively. We design a greedy grid search strategy in terms of the metric of ACC to find the best hyperparameters' combination. Firstly, we fix λ_1 , λ_2 , λ_3 , and search the best combination of K_s and K_c range from 10 to 350, the results on different datasets w.r.t MVSC-CSLF and MVSC-CSLFS are shown in Fig. 6 and Fig. 7, respectively. Secondly, fix λ_3 , K_s and K_c , we search the best combination of λ_1 and λ_2 , and the search space is $\{0.01, 0.1, 1, 10, 100\}$. The results on different datasets w.r.t MVSC-CSLF and MVSC-CSLFS are shown in Fig. 8 and Fig. 9, respectively. Finally, we fix others and tune λ_3 from $\{0.1, 0.5, 1, 5, 10\}$, and the results on different datasets w.r.t MVSC-CSLF and MVSC-CSLFS are shown in Fig. 10.

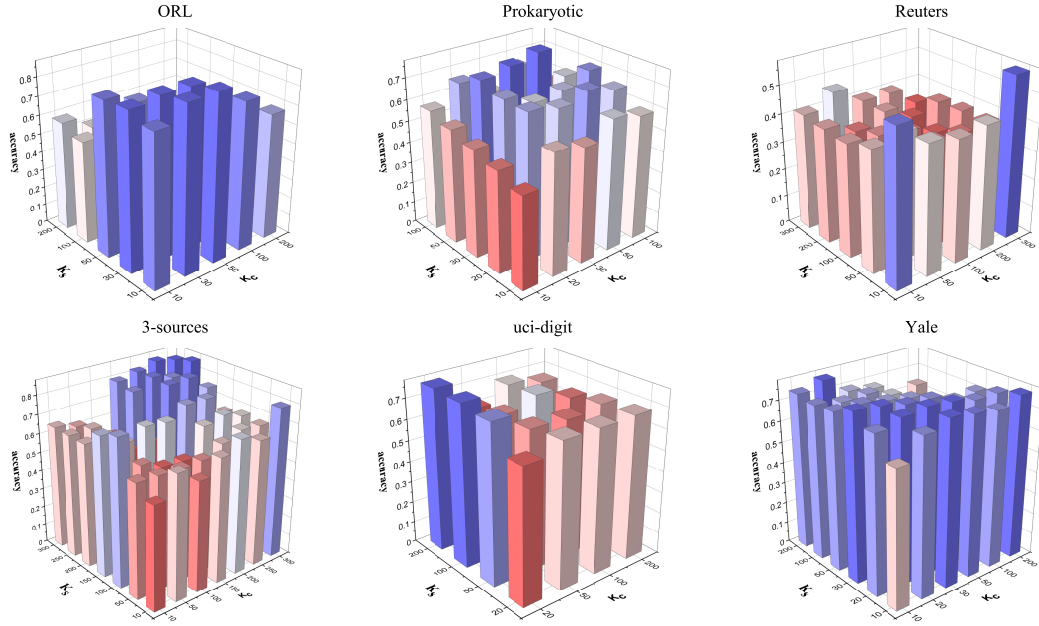


Figure 6: Visualization for the hyperparameter selection procedure: K_s and K_c . We conduct MVSC-CSLF 3 times with different hyperparameter combinations on each data set, and the three axes of the coordinate system represent K_s , K_c , and the mean value of clustering accuracy, respectively.

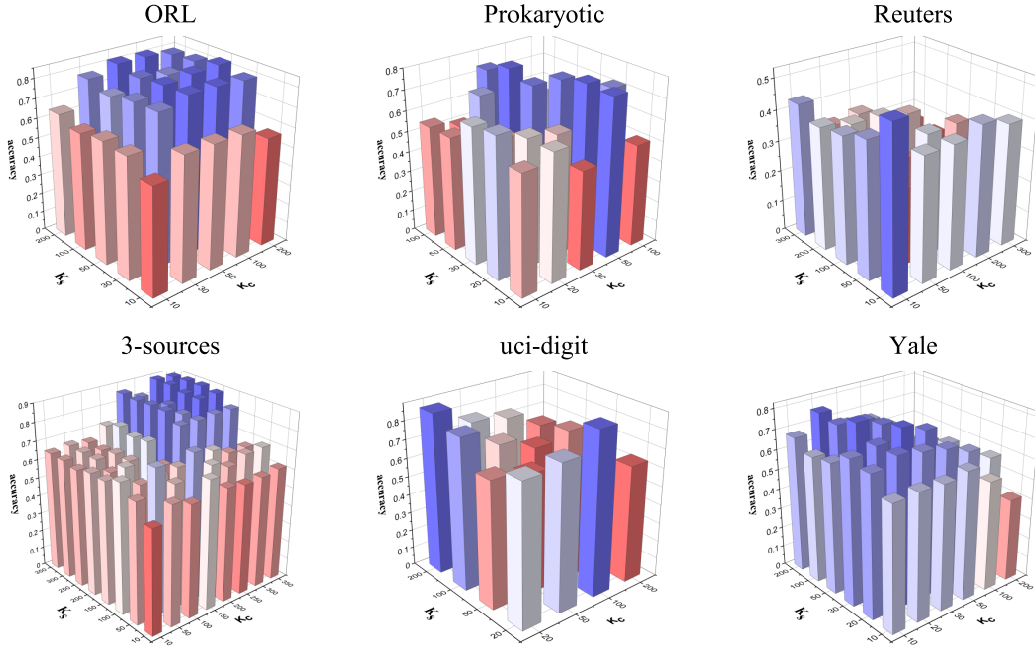


Figure 7: Visualization for the hyperparameter selection procedure: K_s and K_c . We conduct MVSC-CSLFS 3 times with different hyperparameter combinations on each data set, and the three axes of the coordinate system represent K_s , K_c , and the mean value of clustering accuracy, respectively.

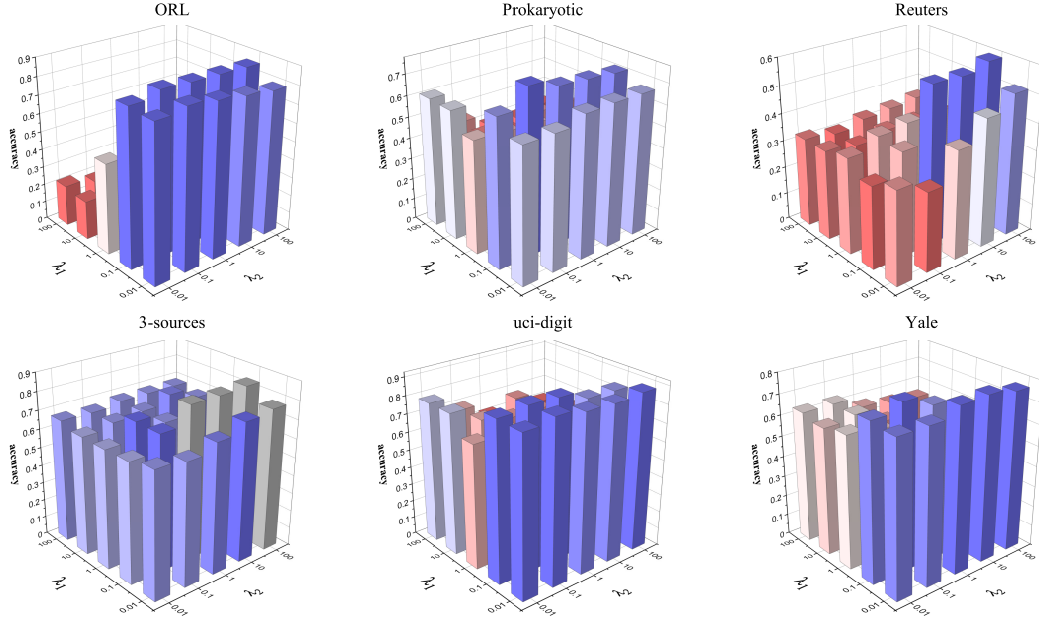


Figure 8: Visualization for the hyperparameter selection procedure: λ_1 and λ_2 . We conduct MVSC-CSLF 3 times with different hyperparameter combinations on each data set, and the three axes of the coordinate system represent λ_1 , λ_2 , and the mean value of clustering accuracy, respectively.

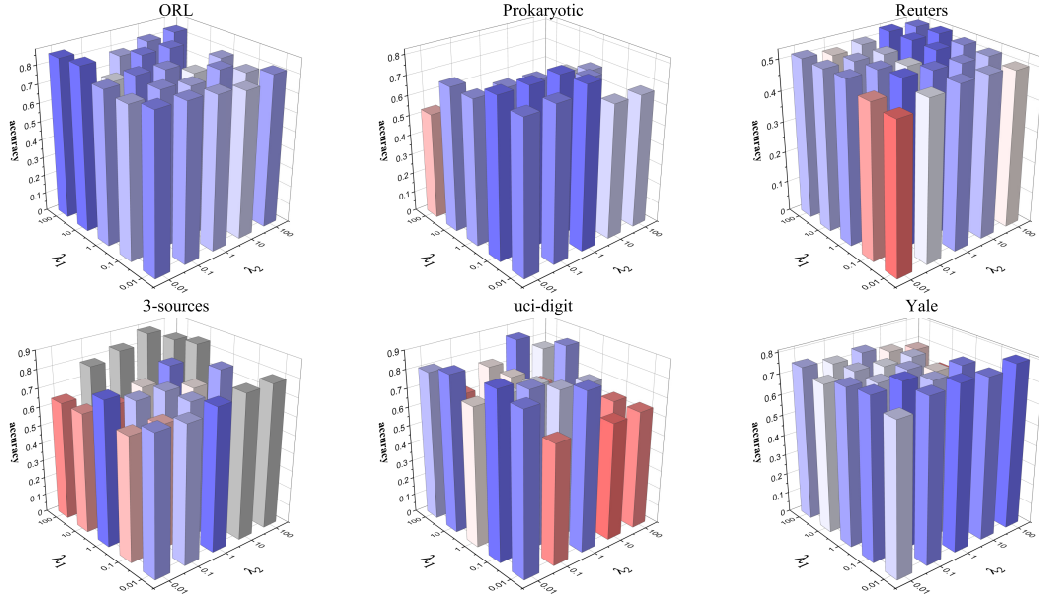


Figure 9: Visualization for the hyperparameter selection procedure: λ_1 and λ_2 . We conduct MVSC-CSLFS 3 times with different hyperparameter combinations on each data set, and the three axes of the coordinate system represent λ_1 , λ_2 , and the mean value of clustering accuracy, respectively.

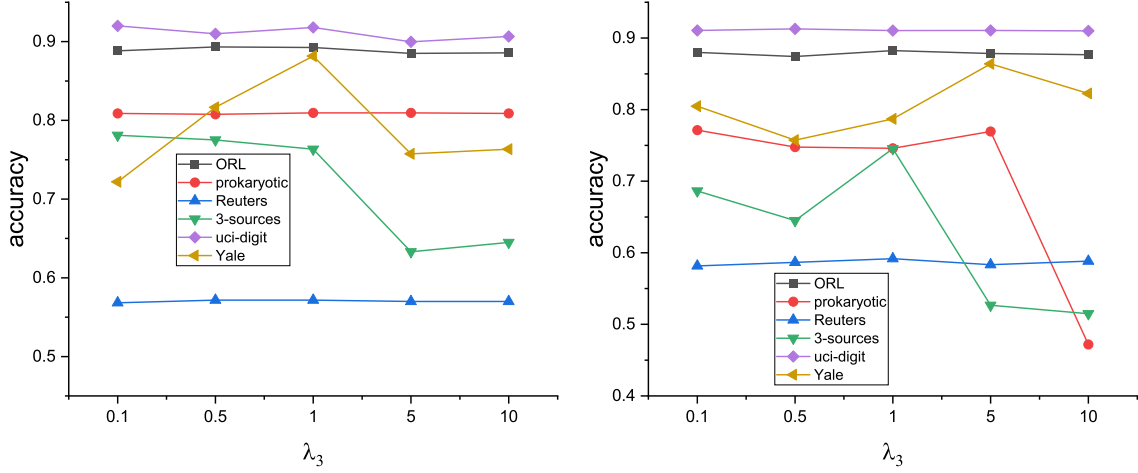


Figure 10: Visualization for the hyperparameter selection procedure w.r.t λ_3 of MVSC-CSLF and MVSC-CSLFS, respectively. We conduct MVSC-CSLF and MVSC-CSLFS 3 times with different λ_3 on each data set, therefore the two axes of the coordinate system represent λ_3 and the mean value of clustering accuracy, respectively.

5. Conclusion and Future Work

In this study, we emphasize the importance of a good representation performing in self-expressive subspace clustering, and introduce a matrix factorization method to explore and separate the consistent and view-specific information existing in multiple views. Based on the idea that representation learning process should be guided by specific learning tasks, we propose two multi-view subspace clustering algorithms MVSC-CSLF and MVSC-CSLFS. Among them, MVSC-CSLF use a feature-level fusion strategy to fuse consistent and complementary information into a joint latent representation, and then perform self-expressive reconstruction on it. MVSC-CSLFS use a subspace-level hierarchical strategy to impose the matching prior constraints on different subspace representations. Experimental results indicate that the quality of data representation is essential for subspace clustering. And our proposed MVSC-CSLFS can further improve clustering results by designing suitable self-expressive reconstruction strategies according to the properties of latent representations.

Although this study has achieved good results, there are still many challenges. One is to improve the capability of model to deal with non-linear and corrupted data. Another is to transfer this paradigm to deep neural network, which has more powerful representation ability. Finally, our proposed algorithms are time-consuming, which is a common problem of subspace learning. Therefore, some new techniques should be introduced to speed up the computation, such as binary representation [47].

References

- [1] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002) 971–987.
- [2] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume I, 2005, pp. 886–893.
- [3] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2004) 91–110.
- [4] S. Sun, L. Mao, Z. Dong, L. Wu, *Multiview Machine Learning*, Springer, 2019.
- [5] J. Liu, Y. Jiang, Z. Li, Z. H. Zhou, H. Lu, Partially shared latent factor learning with multiview data, *IEEE Trans. Neural Netw. Learn. Syst.* 26 (2015) 1233–1246.
- [6] C. Zhang, H. Fu, S. Liu, G. Liu, X. Cao, Low-rank tensor constrained multiview subspace clustering, in: *IEEE International Conference on Computer Vision*, 2015, pp. 1582–1590.
- [7] X. Cao, C. Zhang, H. Fu, Si Liu, Hua Zhang, Diversity-induced Multi-view Subspace Clustering, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 586–594.
- [8] C. Zhang, Q. Hu, H. Fu, P. Zhu, X. Cao, Latent Multi-view Subspace Clustering, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4333–4341.

- [9] M. Brbić, I. Kopriva, Multi-view low-rank sparse subspace clustering, *Pattern Recognit.* 73 (2018) 247–258.
- [10] S. Luo, C. Zhang, W. Zhang, X. Cao, Consistent and specific multi-view subspace clustering, in: *AAAI Conference on Artificial Intelligence*, 2018, pp. 3730–3737.
- [11] T. Zhou, C. Zhang, X. Peng, H. Bhaskar, J. Yang, Dual Shared-Specific Multiview Subspace Clustering, *IEEE T. Cybern.* (2019) 1–14.
- [12] C. Zhang, H. Fu, Q. Hu, X. Cao, Y. Xie, D. Tao, D. Xu, Generalized latent multi-view subspace clustering, *IEEE Trans. Pattern Anal. Mach. Intell.* 42 (2020) 86–99.
- [13] W. Weng, W. Zhou, J. Chen, H. Peng, H. Cai, Enhancing multi-view clustering through common subspace integration by considering both global similarities and local structures, *Neurocomputing* 378 (2020) 375–386.
- [14] Q. Zheng, J. Zhu, Z. Li, S. Pang, J. Wang, Y. Li, Feature concatenation multi-view subspace clustering, *Neurocomputing* 379 (2020) 89–102.
- [15] R. Vidal, Y. Ma, S. S. Sastry, *Generalized Principal Component Analysis*, Springer, 2016.
- [16] M. Chen, L. Huang, C.-D. Wang, D. Huang, Multi-View Clustering in Latent Embedding Space, in: *AAAI Conference on Artificial Intelligence*, 2020, pp. 3513–3520.
- [17] C.-Y. Lu, H. Min, Z.-Q. Zhao, L. Zhu, D.-S. Huang, S. Yan, Robust and Efficient Subspace Segmentation via Least Squares Regression, in: *European Conference on Computer Vision*, 2012, pp. 347–360.
- [18] J. Pan, M. Salzmann, H. Li, Efficient dense subspace clustering, in: *IEEE Winter Conference on Applications of Computer Vision*, 2014, pp. 461–468.
- [19] Z. Cui, X. Jing, P. Zhao, W. Zhang, J. Chen, A new subspace clustering strategy for ai-based data analysis in iot system, *IEEE Internet of Things Journal* 8 (2021) 12540–12549.
- [20] G. Xia, H. Sun, L. Feng, G. Zhang, Y. Liu, Human motion segmentation via robust kernel sparse subspace clustering, *IEEE Trans. Image Process.* 27 (2018) 135–150.
- [21] L. Wang, Z. Ding, Y. Fu, Learning transferable subspace for human motion segmentation, in: *AAAI Conference on Artificial Intelligence*, 2018, pp. 4195–4202.
- [22] W. Zheng, Y. Zong, X. Zhou, M. Xin, Cross-domain color facial expression recognition using transductive transfer subspace learning, *IEEE Trans. Affect. Comput.* 9 (2018) 21–37.
- [23] M. Liao, X. Gu, Face recognition based on dictionary learning and subspace learning, *Digit. Signal Prog.* 90 (2019) 110–124.
- [24] Y. Qiu, P. Hao, Self-supervised deep subspace clustering network for faces in videos, *The Visual Computer* 37 (2021) 2253–2261.
- [25] E. Elhamifar, R. Vidal, Sparse Subspace Clustering: Algorithm, Theory, and Applications, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 2765–2781.
- [26] V. M. Patel, H. Van Nguyen, R. Vidal, Latent Space Sparse Subspace Clustering, in: *IEEE International Conference on Computer Vision*, 2013, pp. 225–232.
- [27] R. Vidal, P. Favaro, Low rank subspace clustering (LRSC), *Pattern Recognit. Lett.* 43 (2014) 47–61.
- [28] V. M. Patel, R. Vidal, Kernel sparse subspace clustering, in: *IEEE International Conference on Image Processing*, 2014, pp. 2849–2853.
- [29] M. Yin, Y. Guo, J. Gao, Z. He, S. Xie, Kernel sparse subspace clustering on symmetric positive definite manifolds, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 5157–5164.
- [30] S. Xiao, M. Tan, D. Xu, Z. Y. Dong, Robust Kernel Low-Rank Representation, *IEEE Trans. Neural Netw. Learn. Syst.* 27 (2016) 2268–2281.
- [31] C. Lu, J. Feng, Z. Lin, T. Mei, S. Yan, Subspace Clustering by Block Diagonal Representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (2019) 487–501.
- [32] M. Abavisani, V. M. Patel, Multimodal sparse and low-rank subspace clustering, *Inf. Fusion* 39 (2018) 168–177.
- [33] D. Xie, W. Xia, Q. Wang, Q. Gao, S. Xiao, Multi-view clustering by joint manifold learning and tensor nuclear norm, *Neurocomputing* 380 (2020) 105–114.
- [34] X. Zhang, H. Sun, Z. Liu, Z. Ren, Q. Cui, Y. Li, Robust low-rank kernel multi-view subspace clustering based on the Schatten p -norm and correntropy, *Inf. Sci.* 477 (2019) 430–447.
- [35] P. Ji, T. Zhang, H. Li, M. Salzmann, I. D. Reid, Deep Subspace Clustering Networks, 2017. [arXiv:1709.02508](https://arxiv.org/abs/1709.02508).
- [36] Z. Xue, J. Du, D. Du, S. Lyu, Deep low-rank subspace ensemble for multi-view clustering, *Inf. Sci.* 482 (2019) 210–227.
- [37] P. Zhu, B. Hui, C. Zhang, D. Du, L. Wen, Q. Hu, Multi-view deep subspace clustering networks, 2019. [arXiv:1908.01978](https://arxiv.org/abs/1908.01978).
- [38] R.-k. Lu, J.-w. Liu, X. Zuo, Attentive multi-view deep subspace clustering net, *Neurocomputing* 435 (2021) 186–196.
- [39] G. Wahba, A Least Squares Estimate of Satellite Attitude, *SIAM Rev.* 7 (1965) 409–409.
- [40] J. L. Farrell, J. C. Stuelpnagel, R. H. Wessner, J. R. Velman, J. E. Brook, A Least Squares Estimate of Satellite Attitude (Grace Wahba), *SIAM Rev.* 8 (1966) 384–386.
- [41] R. H. Bartels, G. W. Stewart, Solution of the Matrix Equation $AX + XB = C$ [F4], *Commun. ACM* 15 (1972) 820–826.
- [42] J. Cai, E. Candès, Z. Shen, A Singular Value Thresholding Algorithm for Matrix Completion, *SIAM J. Optim.* 20 (2010) 1956–1982.
- [43] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, Y. Ma, Robust recovery of subspace structures by low-rank representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (2013) 171–184.
- [44] M. Brbić, M. Piškorec, V. Vidulin, A. Kriško, T. Šmuc, F. Supek, The landscape of microbial phenotypic traits and associated genes, *Nucleic Acids Res.* 44 (2016) 10074–10090.
- [45] D. D. Lewis, Y. Yang, T. G. Rose, F. Li, Rcv1: A new benchmark collection for text categorization research, *J. Mach. Learn. Res.* 5 (2004) 361–397.
- [46] A. Y. Ng, M. I. Jordan, Y. Weiss, On Spectral Clustering: Analysis and an algorithm, in: *Advances in Neural Information Processing Systems* 14, 2002, pp. 849–856.
- [47] Z. Zhang, L. Liu, F. Shen, H. T. Shen, L. Shao, Binary Multi-View Clustering, *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (2019) 1774–1782.