Information Sciences 606 (2022) 877–891

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Low-rank tensor approximation with local structure for multi-view intrinsic subspace clustering

Lele Fu [a], Jinghua Yang [b], Chuan Chen [c,*], Chuanfu Zhang [a]

[a] School of Systems Science and Engineering, Sun Yat-sen University, Guangzhou, China
[b] Faculty of Information Technology, Macau University of Science and Technology, Macau, China
[c] School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

ARTICLE INFO

Article history:
Received 12 January 2022
Received in revised form 29 March 2022
Accepted 26 May 2022
Available online 31 May 2022

Keyword:
Multi-view clustering
Low-rank tensor
Subspace representation
Local manifold

ABSTRACT

Among various multi-view clustering approaches, tensor-based multi-view subspace clustering methods aim to explore the high-order correlations across varying views and have achieved encouraging effects. Nevertheless, there are still some demerits in them: (1) View-specific information hinders the mining of global consensus. (2) The local structure inside individual view lacks consideration. (3) Clustering results are not utilized to reversely guide the low-rank tensor optimization. In order to tackle these drawbacks, we propose a unified model termed as Low-rank Tensor Approximation with Local Structure for Multi-view Intrinsic Subspace Clustering. Specifically, the proposed model learns multiple intrinsic subspace representations via the rank preserving decomposition, which is to mitigate the impact of view-specific information on enhancing the global consistency. Then, these intrinsic subspace representations are assembled into a 3-order target tensor with tensor nuclear norm constraint. To preserve the consistent locality, we adopt the manifold regularization to constrain each view when mapping into the intrinsic subspace. Furthermore, since the learned label indicator matrix implicitly characterizes the cluster structure, which is used to guide the optimization of the low-rank tensor representation. Finally, abundant experiments on six real-word datasets demonstrate that the proposed method is superior over other state-of-the-art clustering methods.

© 2022 Elsevier Inc. All rights reserved.

1. Introduction

Clustering is a traditional but important problem with a wide range of applications in pattern recognition, computer vision, natural language processing, etc. With the advancement of multimedia and feature extraction technologies, multi-view data emerge in large quantities. Compared to single-view data, multi-view data come from multiple feature levels, which contain diverse and complementary information in different views. In order to efficiently count and analyze multi-view data, multi-view clustering methods [1–4] have received increasing attention.

Subspace clustering [5,6] aims at grouping samples into multiple underlying subspaces. As an outstanding clustering technology, it has gained much attention and been introduced into multi-view learning filed, thereby spawning plenty of multi-view subspace clustering works. For example, Li et al. [7] flexibly constructed the underlying representation in a

* Corresponding author.
E-mail addresses: lawrencefzu@gmail.com (L. Fu), yangjinghua110@126.com (J. Yang), chenchuan@mail.sysu.edu.cn (C. Chen), zhangchf9@mail.sysu.edu.cn (C. Zhang).

https://doi.org/10.1016/j.ins.2022.05.091
0020-0255/© 2022 Elsevier Inc. All rights reserved.

subspace, which was enforced to be close to multiple views. Zhang et al. [8] explored a latent subspace, wherein the underlying representation of multi-view data was reconstructed. Kang et al. [9] proposed an efficient anchor graph fusion for dealing with large-scale datasets. Chen et al. [10] used the least squares regression to capture the grouping effect. The above works learn discriminative subspace representations from the matrix level. As for multi-view representation learning, exploring the correlations among views at the tensor level is a better option, this is mainly because it is conducive to pursuing the cross-view consistency by imposing low-rank constraint on the representation tensor.

In order to capture the high-order correlations among multiple views, tensor-based multi-view subspace clustering methods have been further developed. Zhang et al. [11] aggregated multiple subspace representations into a 3-order tensor, which was imposed with low-rank constraint by combining the nuclear norms of all matrices unfolded along each mode. Xie et al. [12] leveraged tensor singular value decomposition (t-SVD) based tensor nuclear norm (TNN) to restore a low-rank self-representation tensor. Gao et al. [13] proposed a weighted tensor nuclear norm based on t-SVD to preserve the low-rank properties of representation tensor. Chen et al. [14] proposed a generalized nonconvex low-rank tensor approximation to cope with multi-view subspace clustering problem. Although these tensor-based approaches have achieved promising effects, they still suffer from issues that to be resolved. Firstly, different views represent different statistical properties [15] and even contradict each other, the view-specific information is not conducive to reinforcing the global consistency, which is usually ignored by these methods. Secondly, the global consistency sought by many tensor-based methods is certainly essential, however, the local structure [16] inside a single view has equally positive significance for the final clustering results, while which has not been given enough attention. Finally, most tensor-based methods [11–14,17,18] first solve the low-rank tensor representation, and then obtain the label indicator matrix via spectral clustering algorithm, which does not utilize the correlation between the two items. Nevertheless, the fact is that the label indicator matrix implicitly characterizes the cluster structure, which can be used as auxiliary supervised information to guide the learning of the low-rank tensor representation. Also, the label indicator matrix tends to be optimal in the dynamic optimization.

In light of the above points, we propose a joint framework that can handle the three concerns at the same time. Concretely, we perform the rank preserving decomposition on the initial self-representation matrices to factor out the intrinsic subspace representations, which are assembled into a 3-order tensor to be optimized. Then, the t-SVD based TNN is used to constrain the target tensor for enhancing the cross-view consensus and exploiting the high-order correlations of views. In order to protect the local structure of data, the manifold regularization is imposed on each view, thereby allowing data points that are closed in the original space to maintain this property in the intrinsic subspace. Moreover, for using clustering results to reversely boost the learning of the low-rank tensor representation, the latter is optimized uniformly with the label indicator matrix in the form of spectral embedding. In general, the contributions of this paper can be outlined as follows.

- Considering that different views have their own specific statistic properties that are not conducive to the enhancement of global consistency, we adopt the rank preserving decomposition to learn intrinsic subspace representations from the initial subspace representations, thereby removing view-specific information and facilitating the pursuit of global consistency.
- The refinement of local structure of data is beneficial to obtain encouraging clustering results. Therefore, we protect the local structure of each view via the manifold regularization. Meanwhile, we use the clustering results to guide the learning of low-rank tensor representation, so as to obtain the optimal results of both in the optimization process.
- For effectively solving the model LTALS, we propose an iterative algorithm based on alternating direction method of multipliers algorithm (ADMM). Moreover, numerous experimental results on six real-world datasets verify the superiority of the proposed method over compared algorithms.

The remainder of this paper is arranged as follows. Section 2 briefly reviews several mainstream multi-view clustering methods. In Section 3, we elaborate the proposed LTALS and the optimization algorithm. Experimental results are analysed in Section 4. Section 5 summarizes the paper.

2. Related Works

A variety of multi-view clustering methods are developed in recent years. Based on the focus of thesis and ease of our narrative, they can be roughly divided into three categories: graph-based methods, subspace-based methods, and tensor-based methods. A brief review is provided as follows.

Graph-based methods aim at refining a consistent similarity matrix of multiple graphs constructed from multi-source features. For instance, Qiang et al. [19] proposed the fast multi-view discrete clustering based on anchor graphs, which solved the spectral clustering problem in close to linear time. Yu et al. [20] proposed a fine-grained similarity fusion method to integrate multiple affinity matrices. For addressing high-dimensional data and achieving effective multi-graph fusion, the work [21] proposed a multi-view projected clustering approach, which learned the structured graph and adaptive weight parameters. Given that the equality of affinity graph is critical for clustering task, the work [22] constructed a robust affinity graph via hypergraph embedding and sparse regression.

Subspace-based methods are dedicated to exploiting a discriminative subspace representation, which also reflects the relationship between data points. Self-representation technology is an important technology in subspace learning, based

on which a large number of representative works have been generated. Works [23–26] aimed at leaning refined self-representations through objective functions designed in sophisticated ways. In order to learn a robust affinity matrix, Rong et al. [27] fused multiple subspaces into a common subspace on the Grassmann manifold. Lv et al. [28] proposed to achieve information integration in a partition space to enhance the model robustness. Li et al. [29] attempted to tackle high-dimensional and noisy data via energy preserving embedding in a latent subspace. Zhang et al. [30] proposed a one-step framework to preserve the nonnegative property and reinforce the discrimination of affinity matrix.

For capturing high-order correlations between diverse feature representations, tensor-based methods have attracted the attention of researchers. The works [11,12] are the earliest typical multi-view clustering works by introducing low-rank tensor learning. Based on this theoretical framework, many interesting works were subsequently proposed. Faced with the challenge of various kinds of errors, Wang et al. [31] simultaneously adopted group l_1-norm and $l_{2,1}$-norm to constrain the error tensor. Wu et al. [32] integrated graph learning and low-rank tensor learning into a unified process for obtaining the reliable representation. For handling nonlinear structure of multiple features and distinguishing the contributions of diverse views, Chen et al. [33] employed the kernel trick to perform feature space transformation and learned a adaptive weight for each view. Zhang et al. [34] utilized a well-designed constraint matrix to guide the learning of subspace representation for fusing prior information of data. Fu et al. [35] constructed a unified framework based on low-rank tensor optimization for multi-view learning.

3. The Proposed Method

In this section, we introduce the Low-rank Tensor Approximation with Local Structure for Multi-view Intrinsic Subspace Clustering and its iterative optimization process. The meanings of some relevant notations are explained in Table 1.

3.1. Problem Formulation

In order to pursue the global consistency and mine the high-order correlations between views, the objective functions of the original tensor-based multi-view clustering models [11,12] are formulated as

$$\begin{aligned}
&\min_{\mathscr{Z},\mathbf{E}} \|\mathscr{Z}\|_* + \lambda\|\mathbf{E}\|_{2,1} \\
&\text{s.t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)}\mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, v = 1, 2, \ldots, m, \\
&\mathscr{Z} = \Phi\left(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \ldots, \mathbf{Z}^{(m)}\right), \\
&\mathbf{E} = \left[\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; \ldots; \mathbf{E}^{(m)}\right],
\end{aligned} \tag{1}$$

where $\lambda > 0$ is a penalty hyperparameter, the operator $\Phi(\cdot)$ denotes the aggregation of multiple matrices into a 3-order tensor. The model [11] selected the sum of nuclear norms (SNN) to capture the low-rank property while the model [12] chose the t-SVD based TNN. The work [36] has demonstrated that the t-SVD based TNN is the tightest convex relaxation for l_1-norm of the tensor multirank. In view of this, we also adopt this tensor nuclear norm to encode the low-rank components of target tensor in LTALS. To be specific, the t-SVD based TNN of a tensor $\mathscr{T} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is defined as

$$\begin{aligned}
\|\mathscr{T}\|_* &= \|bdiag(\mathscr{T}_f)\|_* = \|bdiag(\mathscr{D}_f)\|_* \\
&= \sum_{i=1}^{\min\{n_1,n_2\}} \sum_{j=1}^{n_3} |\mathscr{D}_f^{(j)}(i,i)|,
\end{aligned} \tag{2}$$

where $\mathscr{D}_f^{(j)}$ is obtained by $\mathscr{T}_f^{(j)} = \mathscr{U}_f^{(j)} * \mathscr{D}_f^{(j)} * \mathscr{V}_f^{(j)^T}$. For more details, please refer to the appendix.

Table 1
Explanation of varying notations.

Notations	Descriptions
$\mathbf{a}, \mathbf{A}, \mathscr{A}$	a vector, a matrix, a tensor
$d^{(v)}, n$	the feature dimension of the v-th view, the number of samples
$\|\|\mathbf{M}\|\|_{2,1}$	$\|\|\mathbf{M}\|\|_{2,1} = \sum_j \|\|\mathbf{M}(:,j)\|\|_2$
$\|\|\mathbf{M}\|\|_\infty$	$\|\|\mathbf{M}\|\|_\infty = max_i \sum_j \|\mathbf{M}_{ij}\|$
$Tr(\cdot)$	the trace operator
$\|\|\mathscr{A}\|\|_F$	$\|\|\mathscr{A}\|\|_F = \sqrt{\sum_{ijk} \|\boldsymbol{A}_{ijk}\|^2}$
$\mathscr{A}^{(i)}$	the i-th frontal slice of $\mathscr{A}$
$\mathscr{A}_f$	fast Fourier transformation (FFT) on $\mathscr{A}$ along the third dimension, i.e., $\mathscr{A}_f = fft(\mathscr{A}, [\,], 3)$. Moreover, $\mathscr{A}$ is rederived via the inverse FFT, i.e., $\mathscr{A}_f = ifft(\mathscr{A}, [\,], 3)$

It is obvious that Eq. (1) directly imposes the tensor nuclear norm on the target tensor consisting of the original self-representation matrices. However, different feature representations have different statistic properties [15], which may contradict each other and increase the difficulty of pursing the global consistency. Therefore, the view-specific information of each view should be partitioned out without damaging the intrinsic cluster structure before constructing the target tensor. For this goal, we adopt the rank preserving decomposition to factorize the initial self-representation matrix $\mathbf{Z}^{(v)}$ into the following form:

$$\begin{aligned} &\mathbf{Z}^{(v)} = \mathbf{P}^{(v)}\mathbf{S}^{(v)} \\ &\text{s.t. } \mathbf{P}^{(v)^T}\mathbf{P}^{(v)} = \mathbf{I}, \end{aligned} \tag{3}$$

where $\mathbf{P}^{(v)} \in \mathbb{R}^{n\times n}$ denotes a projection matrix and $\mathbf{S}^{(v)} \in \mathbb{R}^{n\times n}$ is the learned intrinsic subspace representation. Since $\mathbf{P}^{(v)}$ is an orthogonal matrix, the ranks of $\mathbf{Z}^{(v)}$ and $\mathbf{S}^{(v)}$ are equal, thereby retaining the clustering properties, while $\mathbf{P}^{(v)}$ contains view-specific information. Then, we use $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$ instead of $\left\{\mathbf{Z}^{(v)}\right\}_{v=1}^{m}$ to construct the target tensor $\mathscr{S}$.

According to [37], if two data points are close to each other in the original space, they still maintain this property in the new low-dimensional space. Following this assumption, we think that if two original samples $\mathbf{x}_i^{(v)}$ and $\mathbf{x}_j^{(v)}$ are close, then the corresponding $\mathbf{s}_i^{(v)}$ and $\mathbf{s}_j^{(v)}$ in the intrinsic subspace should be similar. Thus, we have the following manifold regularization:

$$\sum_i\sum_j ||\mathbf{s}_i^{(v)} - \mathbf{s}_j^{(v)}||_2^2\mathbf{W}_{ij}^{(v)} = \frac{1}{2}Tr\left(\mathbf{S}^{(v)}\mathbf{L}^{(v)}\mathbf{S}^{(v)^T}\right), \tag{4}$$

where $\mathbf{W}_{ij}^{(v)}$ represents the similarity between $\mathbf{x}_i^{(v)}$ and $\mathbf{x}_j^{(v)}$ and is computed via the Gaussian kernel function with k-nearest neighbors. $\mathbf{L}^{(v)}$ is the Laplacian matrix of the v-th view, which is calculated by $\mathbf{L}^{(v)} = \mathbf{D}^{(v)} - \mathbf{W}^{(v)}$. Specifically, $\mathbf{D}^{(v)}$ is a diagonal matrix and defined as $\mathbf{D}_{ii}^{(v)} = \sum_j \mathbf{W}_{ij}^{(v)}$.

After obtaining the low-rank tensor $\mathscr{Z}$, most tensor-based methods solve the final affinity matrix via $\frac{1}{m}\sum_{v=1}^{m}\left(\left|\mathbf{Z}^{(v)}\right| + |\mathbf{Z}^{(v)^T}|\right)/2$, which is used as the input of spectral clustering algorithm to get the label indicator matrix. It can be known that the learning of the low-rank tensor representation and label indicator matrix is two independent processes, so that the cluster structure information contained in the latter is not employed in the optimization process. Therefore, we want to make up for this shortcoming and make the both tend to optimality in a unified optimization process. To tackle this issue, we propose the following formula:

$$\begin{aligned} &\min_{\mathbf{F}}\sum_{v=1}^{m}\sum_{i=1}^{n}\sum_{j=1}^{n}||\mathbf{f}_i - \mathbf{f}_j||_2^2\overline{\mathbf{S}}_{ij}^{(v)} = \min_{\mathbf{F}}\sum_{v=1}^{m}Tr\left(\mathbf{F}^T\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}\mathbf{F}\right) \\ &\text{s.t. } \mathbf{F}^T\mathbf{F} = \mathbf{I}, \end{aligned} \tag{5}$$

where $\mathbf{f}_i$ and $\mathbf{f}_j$ represent the i-th and j-th row of $\mathbf{F}$, respectively. The affinity $\overline{\mathbf{S}}_{ij}^{(v)}$ between the i-th and j-th sample is obtained by $\bar{\mathbf{S}}_{ij}^{(v)} = (\mathbf{S}_{ij} + \mathbf{S}_{ji})/2$. Laplacian matrix $\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}$ is defined as $\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)} = \mathbf{D}^{(v)} - \overline{\mathbf{S}}^{(v)}$, and $\mathbf{D}_{ii}^{(v)} = \sum_j \overline{\mathbf{S}}^{(v)}$. $\mathbf{F} \in \mathbb{R}^{n\times c}$ is the learned label indicator matrix. The motivation behind Eq. (5) is that when $\mathbf{F}$ is obtained by performing eigenvector decomposition on $\sum_{v=1}^{m}\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}$, we want to use $\mathbf{F}$ to promote the optimization of $\left\{\overline{\mathbf{S}}^{(v)}\right\}_{v=1}^{m}$ (i.e., the tensor $\mathscr{S}$) in reverse. According to graph embedding theory [37], if the distance between the label indicator vector $\mathbf{f}_i$ and $\mathbf{f}_j$ is too large, then the values of the corresponding $\left\{\overline{\mathbf{S}}_{ij}^{(v)}\right\}_{v=1}^{m}$ should be smaller, otherwise tend to be larger. Finally, in light of the above discussions, we formulate the final objective function as the following form:

$$\begin{aligned} &\min_{\mathscr{S},\mathbf{Z}^{(v)},\mathbf{E},\mathbf{F}} \underbrace{\|\mathscr{S}\|_* + \lambda\|\mathbf{E}\|_{2,1}}_{\textbf{Low-rank tensor learning}} + \beta\sum_{v=1}^{m}\underbrace{Tr\left(\mathbf{S}^{(v)}\mathbf{L}^{(v)}\mathbf{S}^{(v)^T}\right)}_{\textbf{Local manifold}} + \alpha\sum_{v=1}^{m}\underbrace{Tr\left(\mathbf{F}^T\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}\mathbf{F}\right)}_{\textbf{Spectral embedding}} \\ &\text{s.t. } \mathbf{X}^{(v)} = \mathbf{X}^{(v)}\mathbf{Z}^{(v)} + \mathbf{E}^{(v)}, \mathbf{Z}^{(v)} = \mathbf{P}^{(v)}\mathbf{S}^{(v)}, \\ &\mathbf{P}^{(v)^T}\mathbf{P}^{(v)} = \mathbf{I}, \mathbf{F}^T\mathbf{F} = \mathbf{I}, \\ &\mathscr{S} = \Phi\left(\mathbf{S}^{(1)}, \mathbf{S}^{(2)}, \ldots, \mathbf{S}^{(m)}\right), \\ &\mathbf{E} = \left[\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; \ldots; \mathbf{E}^{(m)}\right], \end{aligned} \tag{6}$$

where $\lambda > 0, \beta > 0$, and $\alpha > 0$ are three penalty hyperparameters. When the iterative optimization process is complete, the global unique $\mathbf{F}$ can be gained, on which the kmeans algorithm is used to get the clustering results. Fig. 1 presents the overall idea of the proposed LTALS.

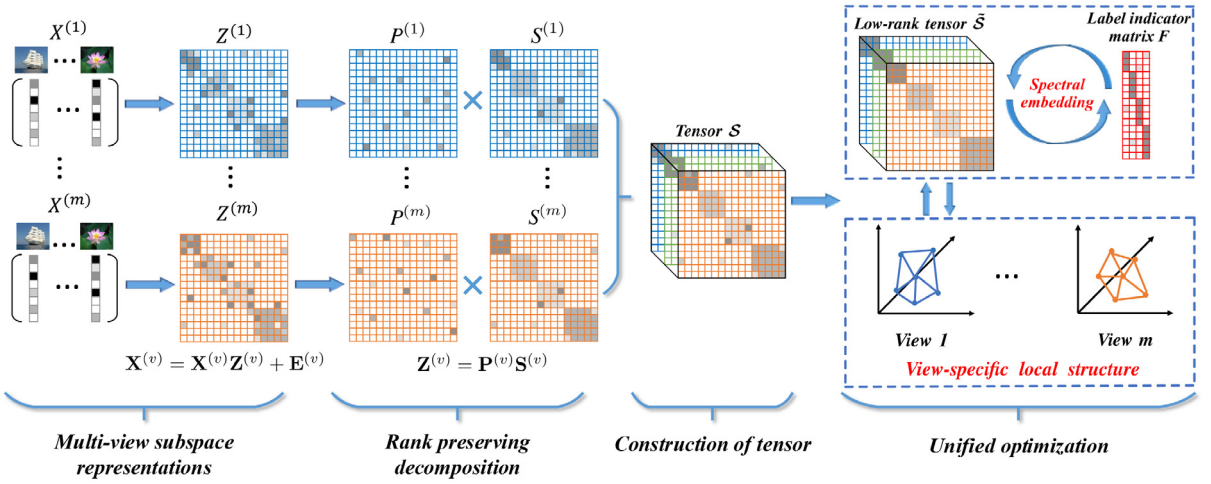

Fig. 1. The overall framework of the proposed LTALS. The rank preserving decomposition is performed on the original self-representation matrices $\left\{\mathbf{Z}^{(v)}\right\}_{v=1}^{m}$ to obtain the intrinsic subspace representations $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$, which consist of the target tensor $\mathcal{S}$. The local structure of each view is preserved while optimizing the low-rank tensor $\tilde{\mathcal{S}}$. Furthermore, the label indicator matrix **F** as the auxiliary supervised information is used to guide the optimization of low-rank tensor $\tilde{\mathcal{S}}$, which are in a process of mutual promoting optimization.

3.2. Optimization Process

Eq. (6) can be solved by ADMM, whose augmented Lagrangian function is written as

$$\begin{aligned}\mathcal{F}\left(\left\{\mathbf{Z}^{(v)}\right\}_{v=1}^{m};\left\{\mathbf{E}^{(v)}\right\}_{v=1}^{m};\mathcal{H};\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m};\left\{\mathbf{P}^{(v)}\right\}_{v=1}^{m};\mathbf{F}\right) &= \|\mathcal{H}\|_{*} + \lambda\|\mathbf{E}\|_{2,1} + \beta\sum_{v=1}^{m} Tr\left(\mathbf{S}^{(v)}\mathbf{L}^{(v)}\mathbf{S}^{(v)^{T}}\right) + \alpha\sum_{v=1}^{m} Tr\left(\mathbf{F}^{T}\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}\mathbf{F}\right) \\ &+ \sum_{v=1}^{m}\left(\frac{\mu}{2}\|\mathbf{X}^{(v)} - \mathbf{X}^{(v)}\mathbf{Z}^{(v)} - \mathbf{E}^{(v)} + \frac{\mathbf{J}^{(v)}}{\mu}\|_{F}^{2}\right) + \sum_{v=1}^{m}\left(\frac{\rho}{2}\|\mathbf{Z}^{(v)} - \mathbf{P}^{(v)}\mathbf{S}^{(v)} + \frac{\mathbf{M}^{(v)}}{\rho}\|_{F}^{2}\right) + \frac{\rho}{2}\|\mathcal{S} - \mathcal{H} + \frac{\mathcal{K}}{\rho}\|_{F}^{2},\end{aligned} \tag{7}$$

where $\mathcal{H}$ is an auxiliary variable and we have $\mathcal{S} = \mathcal{H}$, $\left\{\mathbf{J}^{(v)}\right\}_{v=1}^{m}$, $\left\{\mathbf{M}^{(v)}\right\}_{v=1}^{m}$, and $\mathcal{K}$ are Lagrange multipliers, μ and ρ are penalty parameters. The details of updating rules for each variable based on ADMM are described below.

(1) **Update $\mathbf{Z}^{(v)}$**: Since each $\mathbf{Z}^{(v)}$ is solved in the same way, it is sufficient to focus on the solution of only one of them. When fixing other unrelated items, the $\mathbf{Z}^{(v)}$ subproblem becomes

$$\min_{\mathbf{Z}^{(v)}} \frac{\mu}{2}\|\mathbf{X}^{(v)} - \mathbf{X}^{(v)}\mathbf{Z}^{(v)} - \mathbf{E}^{(v)} + \frac{\mathbf{J}^{(v)}}{\mu}\|_{F}^{2} + \frac{\rho}{2}\|\mathbf{Z}^{(v)} - \mathbf{P}^{(v)}\mathbf{S}^{(v)} + \frac{\mathbf{M}^{(v)}}{\rho}\|_{F}^{2}. \tag{8}$$

By taking the partial derivative of $\mathbf{Z}^{(v)}$ and setting the value to zero, we get

$$\mathbf{Z}^{(v)} = \left(\mu\mathbf{X}^{(v)^{T}}\mathbf{X}^{(v)} + \rho\mathbf{I}\right)^{-1}\left(\mu\mathbf{X}^{(v)^{T}}\mathbf{X}^{(v)} - \mu\mathbf{X}^{(v)^{T}}\mathbf{E}^{(v)} + \mathbf{X}^{(v)^{T}}\mathbf{J}^{(v)} + \rho\mathbf{P}^{(v)}\mathbf{S}^{(v)} - \mathbf{M}^{(v)}\right). \tag{9}$$

(2) **Update $\mathbf{S}^{(v)}$**: For optimizing $\mathbf{S}^{(v)}$, tensors $\mathcal{S}$, $\mathcal{H}$, and $\mathcal{K}$ are decomposed into matrix forms, i.e., $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$, $\left\{\mathbf{H}^{(v)}\right\}_{v=1}^{m}$, and $\left\{\mathbf{K}^{(v)}\right\}_{v=1}^{m}$. Focusing only on the solution of one of $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$ and retaining the terms associated with $\mathbf{S}^{(v)}$, $\mathbf{S}^{(v)}$ is solved by

$$\min_{\mathbf{S}^{(v)}} \beta Tr\left(\mathbf{S}^{(v)}\mathbf{L}^{(v)}\mathbf{S}^{(v)^{T}}\right) + \alpha Tr\left(\mathbf{F}^{T}\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}\mathbf{F}\right) + \frac{\rho}{2}\|\mathbf{Z}^{(v)} - \mathbf{P}^{(v)}\mathbf{S}^{(v)} + \frac{\mathbf{M}^{(v)}}{\rho}\|_{F}^{2} + \frac{\rho}{2}\|\mathbf{S}^{(v)} - \mathbf{H}^{(v)} + \frac{\mathbf{K}^{(v)}}{\rho}\|_{F}^{2}, \tag{10}$$

by setting the partial derivatives of Eq. (10) with respect to $\mathbf{S}^{(v)}$ to zero, the following formula can be gained

$$\left(\rho\mathbf{P}^{(v)^{T}}\mathbf{P}^{(v)} + \rho\mathbf{I}\right)\mathbf{S}^{(v)} + \mathbf{S}^{(v)}\left(\beta\mathbf{L}^{(v)^{T}} + \beta\mathbf{L}^{(v)}\right) = \rho\mathbf{P}^{(v)^{T}}\mathbf{Z}^{(v)} + \mathbf{P}^{(v)^{T}}\mathbf{M}^{(v)} + \rho\mathbf{H}^{(v)} - \mathbf{K}^{(v)} + \alpha\mathbf{F}\mathbf{F}^{T}, \tag{11}$$

which satisfies the form of Sylvester equation $\mathbf{A}\mathbf{S}^{(v)} + \mathbf{S}^{(v)}\mathbf{B} = \mathbf{C}$. Thus, we can update $\mathbf{S}^{(v)}$ by the off-the-shelf solver [38].

(3) **Update $\mathbf{P}^{(v)}$**: Fixing $\left\{\mathbf{Z}^{(v)}\right\}_{v=1}^{m}$, $\left\{\mathbf{E}^{(v)}\right\}_{v=1}^{m}$, $\mathcal{H}$, $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$, and **F**, we have

$$\begin{aligned}&\min_{\mathbf{P}^{(v)}}\tfrac{\rho}{2}||\mathbf{Z}^{(v)}-\mathbf{P}^{(v)}\mathbf{S}^{(v)}+\tfrac{\mathbf{M}^{(v)}}{\rho}||_F^2=\min_{\mathbf{P}^{(v)}}\tfrac{\rho}{2}||\left(\mathbf{Z}^{(v)}+\tfrac{\mathbf{M}^{(v)}}{\rho}\right)-\mathbf{P}^{(v)}\mathbf{S}^{(v)}||_F^2\\&\text{s.t.}\quad \mathbf{P}^{(v)^T}\mathbf{P}^{(v)}=\mathbf{I}.\end{aligned}\tag{12}$$

According to [39], the updating formula of $\mathbf{P}^{(v)}$ can be written as

$$\mathbf{P}^{(v)}=\mathbf{U}^{(v)}\mathbf{V}^{(v)^T},\tag{13}$$

where the solutions of $\mathbf{U}^{(v)}$ and $\mathbf{V}^{(v)}$ can be obtained by SVD implemented on the matrix $\left(\mathbf{Z}^{(v)}+\frac{1}{\rho}\mathbf{M}^{(v)}\right)\mathbf{S}^{(v)^T}$:

$$\left(\mathbf{Z}^{(v)}+\frac{1}{\rho}\mathbf{M}^{(v)}\right)\mathbf{S}^{(v)^T}=\mathbf{U}^{(v)}\Sigma^{(v)}\mathbf{V}^{(v)^T}.\tag{14}$$

(4) **Update E**: Keeping the relevant terms about **E**, the problem is transformed into the following form

$$\begin{aligned}&\min_{\mathbf{E}}\lambda\|\mathbf{E}\|_{2,1}+\sum_{v=1}^{m}\left(\tfrac{\mu}{2}||\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}-\mathbf{E}^{(v)}+\tfrac{\mathbf{J}^{(v)}}{\mu}||_F^2\right)\\&=\min_{\mathbf{E}}\lambda||\mathbf{E}||_{2,1}+\tfrac{\mu}{2}\sum_{v=1}^{m}||\mathbf{E}^{(v)}-\left(\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}+\tfrac{\mathbf{J}^{(v)}}{\mu}\right)||_F^2\\&=\min_{\mathbf{E}}\lambda||\mathbf{E}||_{2,1}+\tfrac{\mu}{2}||\mathbf{E}-\mathbf{B}||_F^2,\end{aligned}\tag{15}$$

where $\mathbf{B}=\left[\mathbf{B}^{(1)};\mathbf{B}^{(2)};\cdots;\mathbf{B}^{(m)}\right]$ and $\mathbf{B}^{(v)}=\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}+\frac{1}{\mu}\mathbf{J}^{(v)}$. According to [6], **E** is computed by

$$\mathbf{E}^{*}_{:,j}=\begin{cases}\frac{||\mathbf{B}_{:,j}||_2-\frac{\lambda}{\mu}}{||\mathbf{B}_{:,j}||_2}\mathbf{B}_{:,j}, & \left\|\mathbf{B}_{:,j}\right\|_2>\frac{\lambda}{\mu}\\ \mathbf{0}, & \text{otherwise.}\end{cases}\tag{16}$$

(5) **Update** $\mathscr{H}$: Retaining the items related of $\mathscr{H}$, we obtain the following problem

$$\begin{aligned}&=\min_{\mathscr{H}}||\mathscr{H}||_*+\tfrac{\rho}{2}||\mathscr{S}-\mathscr{H}+\tfrac{\mathscr{K}}{\rho}||_F^2\\&=\min_{\mathscr{H}}||\mathscr{H}||_*+\tfrac{\rho}{2}||\mathscr{H}-\left(\mathscr{S}+\tfrac{\mathscr{K}}{\rho}\right)||_F^2.\end{aligned}\tag{17}$$

For effectively optimizing the subproblem, we rotate the size of intrinsic subspace tensor representation $\mathscr{H}$ from $n\times n\times m$ to $n\times m\times n$ following [12], Fig. 2 presents how tensor rotation is performed. We can see that the low-rank properties are preserved across views after the rotation operation, which helps to strengthen the global consistency. Moreover, the computational cost can be greatly decreased by the rotation operation, which has been verified in [12]. Following [40], we get the updating approach of $\mathscr{H}$ by the tensor tubal-shrinkage operator

$$\mathscr{H}^{*}=\mathscr{U}*\mathscr{G}_{m/\rho}(\mathscr{D})*\mathscr{V}^{T},\tag{18}$$

where $\mathscr{G}_{m/\rho}(\mathscr{D})=\mathscr{D}*\mathscr{Q}$. $\mathscr{Q}$ is an f-diagonal tensor and its diagonal element of k-th frontal slice is calculated by $\widehat{\mathscr{Q}}(i,i,k)=max\left\{1-\frac{m/\rho}{\widehat{\mathscr{D}}(i,i,k)},0\right\}$.

(6) **Update F**: When $\left\{\mathbf{Z}^{(v)}\right\}_{v=1}^{m}$, $\left\{\mathbf{E}^{(v)}\right\}_{v=1}^{m}$, $\mathscr{H}$, $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$, and $\left\{\mathbf{P}^{(v)}\right\}_{v=1}^{m}$ are fixed, we have

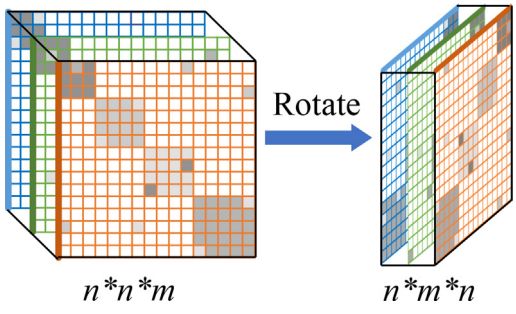

Fig. 2. The illustration of tensor rotation.

$$\begin{aligned}&\min_{\mathbf{F}}\sum_{v=1}^{m}Tr\left(\mathbf{F}^T\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}\mathbf{F}\right)=\min_{\mathbf{F}}Tr\left(\mathbf{F}^T\mathbf{L}\mathbf{F}\right)\\&\text{s.t. }\mathbf{F}^T\mathbf{F}=\mathbf{I},\end{aligned}\tag{19}$$

where $\mathbf{L}=\sum_{v=1}^{m}\mathbf{L}_{\mathbf{S}^{(v)}}^{(v)}$. The updated $\mathbf{F}^*$ is composed of the eigenvectors corresponding to the first c smallest eigenvalues of $\mathbf{L}$.

(7) **Update the Lagrangian multipliers $\mathbf{J}^{(v)},\mathbf{M}^{(v)},\mathscr{K}$, and penalty parameters μ,ρ**:

$$\begin{aligned}&\mathbf{J}^{(v)^*}=\mathbf{J}^{(v)}+\mu\left(\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}^{(v)}-\mathbf{E}^{(v)}\right);\\&\mathbf{M}^{(v)^*}=\mathbf{M}^{(v)}+\rho\left(\mathbf{Z}^{(v)}-\mathbf{P}^{(v)}\mathbf{S}^{(v)}\right);\\&\mathscr{K}^*=\mathscr{K}+\rho(\mathscr{Z}-\mathscr{K});\\&\mu^*=min(\omega*\mu,\mu_{max});\\&\rho^*=min(\omega*\rho,\rho_{max}).\end{aligned}\tag{20}$$

We summarize the main steps of the proposed LTALS in Algorithm 1. For Algorithm 1, the computational complexity lies mainly in the updating of $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m},\left\{\mathbf{P}^{(v)}\right\}_{v=1}^{m},\mathbf{E},\mathscr{H}$, and $\mathbf{F}$. Specifically, updating $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$ via solving Sylvester equation needs $\mathcal{O}(mn^3)$. SVD operation is adopted when updating $\left\{\mathbf{P}^{(v)}\right\}_{v=1}^{m}$, so it takes $\mathcal{O}(mn^3)$. Solving the subproblem of $\mathbf{E}$ costs $\mathcal{O}(mn^2)$. Due to using FFT, inverse FFT, and SVD operations when optimizing $\mathscr{H}$, it takes $\mathcal{O}(m^2n^2+mn^2log(n))$. For solving $\mathbf{F},\mathcal{O}(cn^2)$ is costed. In general, the computational complexity of Algorithm 1 is $\mathcal{O}(t(mn^3+mn^2log(n)+(m+m^2+c)n^2))$, where t is the number of iterations.

Algorithm 1: Low-rank Tensor Approximation with Local Structure for Multi-view Intrinsic Subspace Clustering

Input: Multiple features $\left\{\mathbf{X}^{(m)}\right\}_{v=1}^{m},\mathbf{X}^{(v)}\in\mathbb{R}^{d^{(v)}\times n},\lambda,\beta,\alpha$, and number of clusters c.
Output: Label indicator matrix $\mathbf{F}$.
1: Initialize $\mathscr{S}_0=\mathscr{H}_0=\mathscr{K}_0=\mathbf{0},\mathbf{E}_0=\mathbf{0},\mathbf{J}_0^{(v)}=\mathbf{0},\mathbf{M}_0^{(v)}=\mathbf{0},\omega=2,\varepsilon=10^{-6},\mu_0,\rho_0,\mu_{max}=\rho_{max}=10^{10},k=0$.
2: **while**not convergent**do**
3: **for** $v=1:m$ **do**
4: Update $\mathbf{Z}_{k+1}^{(v)}$ by Eq. (9);
5: Update $\mathbf{S}_{k+1}^{(v)}$ by Eq. (11);
6: Update $\mathbf{P}_{k+1}^{(v)}$ by Eq. (13);
7: **end for**
8: Update $\mathbf{E}_{k+1}$ by Eq. (16);
9: Update $\mathscr{H}_{k+1}$ by Eq. (18);
10: Update $\mathbf{F}_{k+1}$ by Eq. (19);
11: Update $\mathbf{J}_{k+1}^{(v)},\mathbf{M}_{k+1}^{(v)},\mathscr{K}_{k+1},\mu_{k+1}$, and ρ_{k+1} by Eq. (20);
12: Check the convergence conditions:
$||\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{Z}_{k+1}^{(v)}-\mathbf{E}_{k+1}^{(v)}||_\infty\leqslant\varepsilon$,
$||\mathbf{S}_{k+1}^{(v)}-\mathbf{H}_{k+1}^{(v)}||_\infty\leqslant\varepsilon$.
13: $k=k+1$;
14: **end while** 15: **return** Matrix $\mathbf{F}$;
16: Obtain the data labels by performing the kmeans algorithm on $\mathbf{F}$.

4. Experiments

4.1. Datasets

Six challenging real-world datasets are selected to test the performance of the proposed LTALS, whose information is described below. Moreover, Table 2 also presents detailed statistics of them.

BBCnews [1] contains 685 news documents consisted of 5 themes, including entertainmentn, sport, business, politics, and technology. Every sample has four kinds of feature representations.

[1] http://mlg.ucd.ie/datasets/bbc.html

Table 2
Statistics of six datasets.

Feature type	BBCnews	Caltech	NUS-WIDE	UCI	WikipediaArticles	Youtube
1	View1 (4569)	Gabor (48)	CH (65)	PIX (240)	View1 (128)	CH (2000)
2	View2 (4633)	WM (40)	CM (226)	FOU (76)	View2 (10)	HME (1024)
3	View3 (4665)	CENT (254)	CORR (145)	MOR (6)	-	HOG (64)
4	View4 (4684)	HOG (1984)	ED (74)	-	-	MFCC (512)
5	-	GIST (512)	WT (129)	-	-	VS (64)
6	-	LBP (928)	-	-	-	SS (647)
Samples	685	2386	2000	2000	693	2000
Clusters	5	20	31	10	10	10

Caltech [2] is a collection of object images. A total of 2386 images of 20 classes are selected to form the experimental dataset, and each image has six features: Gabor feature, WM feature, CENT feature, HOG feature, GIST feature, and LBP feature.

NUS-WIDE [3] contains 2000 web images that are collected from 31 classes, each image is represented by five kinds of features: CH feature, CM feature, CORR feature, ED feature, and WT feature.

UCI [41] is comprised of 2000 digit images ranging in [0, 9]. There are three feature representations in this dataset: PIX feature, FOU feature, and MOR feature.

WikipediaArticles [4] contains 693 documents divided into 10 categories. Each document is represented by two feature representations.

Youtube [5] is a video dataset with 2000 samples of 10 categories, and each sample is represented by six features, including CH feature, HME feature, HOG feature, MFCC feature, VS feature, and SS feature.

4.2. Compared Methods

In experiments, a classical clustering algorithm and nine multi-view clustering methods are compared with the proposed method LTALS, we briefly introduce them below.

SPC_{best} implemented standard spectral clustering algorithm on the most informative view to obtain the optimal results.

MCGC [42] learned a consistent graph via maximizing agreement among views, whose Laplacian matrix was imposed with rank constraint.

LMVSC [9] proposed an efficient anchor graph fusion method for dealing with large-scale datasets.

MCLES [43] learned the global structure and the label indicator matrix in a latent embedding space.

LTMSC [11] adopted the sum of nuclear norms (SNN) to constrain the self-representation tensor.

t-SVD-MSC [12] leveraged t-SVD based TNN to capture the high-order correlations between different views.

ETLMSC [17] learned a purity transition probability tensor, then the spectral clustering by Markov chain was used to obtain the clustering results.

TISRL [18] imposed t-SVD based TNN on the intrinsic representations learned from subspace representations of all views.

HLR-M^2VS [44] considered the nonlinear subspace representation problem and imposed the hyper-Laplacian regularization on the target tensor.

CGL [45] integrated the low-rank tensor optimization and spectral embedding into a joint learning process to explore a discriminative label indicator matrix.

4.3. Evaluation Metrics

For quantifying the performance of varying clustering algorithms, the following mainstream evaluation metrics are employed: clustering accuracy (ACC), normalized mutual information (NMI), purity, adjusted rank index (ARI), F-score, and precision. For ARI, its value varies in [-1, 1]. For the other metrics, their values range in [0, 1]. The lager values indicate the better performance. Moreover, for observing the statistic significance of comparison clustering results, we perform t-test to compute the p-value using the proposed LTALS's experimental results of ten runs against that of its closest competitor for each metric. The smaller the p-value, the more it indicates the superiority of the proposed method.

4.4. Experimental Results

Considering the effect of algorithm stability, each experiment is run ten times on Matlab 2018b installed on a PC with Intel Core i5-9500 CPU @ 3.00 GHz and 24G RAM, we record the means and variances. All clustering results on six datasets

[2] http://www.vision.caltech.edu/Image_ Datasets/Caltech101
[3] http://lms.comp.nus.edu.sg/research/NUS-WIDE.html
[4] http://lig-membres.imag.fr/grimal/data.html
[5] http://archive.ics.uci.edu/ml/datasets

are presented in Tables 3–8. The numerical results are bolded and the suboptimal results are underlined. We have some significant observations from these statistics.

First of all, most of multi-view approaches outperform the single-view method SPC_{best}. The main reason is that multi-view approaches can exploit and utilize the hidden complementary information in multi-view data, thus improving the clustering effects. Although SPC_{best} uses the data with the most discriminative features from multiple views, it ultimately fails to achieve the best results because of the information monotonicity of a single view. Secondly, the tensor-based approaches generally perform well, which aggregate the affinity matrices of all views into a 3-order tensor and attempt to recover the low-rank part. The final low-rank representations take into account the consistency among multiple views and mine the complementary information in them, thus obtaining discriminative data information and achieving good clustering performance. Moreover, it can be seen that the results of LTMSC are less favorable than other tensor-based methods, mainly because the former uses SNN to preserve low-rank properties, while the latter adopts the t-SVD based TNN. Thirdly, the proposed LTALS achieves priority on all datasets. Particularly, LTALS gains excellent results on BBCnews dataset with approximately 0.990 for each metric, near perfect performance are also achieved on UCI dataset. As for the rest of datasets,

Table 3
Comparison of experimental results on BBCnews dataset.

Methods	ACC	NMI	Purity	ARI	F-score	Precision
SPC_{best}	0.438 ± 0.002	0.295 ± 0.001	0.548 ± 0.001	0.204 ± 0.001	0.399 ± 0.000	0.382 ± 0.002
MCGC	0.350 ± 0.000	0.039 ± 0.000	0.366 ± 0.000	0.001 ± 0.000	0.373 ± 0.000	0.235 ± 0.000
LMVSC	0.588 ± 0.000	0.503 ± 0.000	0.737 ± 0.000	0.417 ± 0.000	0.546 ± 0.000	0.579 ± 0.000
MCLES	0.706 ± 0.012	0.482 ± 0.017	0.706 ± 0.012	0.474 ± 0.032	0.626 ± 0.021	0.508 ± 0.024
LTMSC	0.579 ± 0.001	0.424 ± 0.006	0.632 ± 0.003	0.401 ± 0.003	0.574 ± 0.003	0.524 ± 0.002
t-SVD-MSC	0.958 ± 0.000	0.866 ± 0.000	0.958 ± 0.000	0.900 ± 0.000	0.923 ± 0.000	0.925 ± 0.000
ETLMSC	0.953 ± 0.005	0.889 ± 0.006	0.953 ± 0.005	0.893 ± 0.010	0.917 ± 0.008	0.949 ± 0.005
TISRL	0.902 ± 0.000	0.785 ± 0.000	0.902 ± 0.000	0.805 ± 0.000	0.849 ± 0.000	0.879 ± 0.000
HLR-M^2VS	0.959 ± 0.000	0.875 ± 0.000	0.959 ± 0.000	0.903 ± 0.000	0.926 ± 0.000	0.922 ± 0.000
CGL	0.924 ± 0.000	0.842 ± 0.000	0.924 ± 0.000	0.861 ± 0.000	0.893 ± 0.000	0.907 ± 0.000
LTALS	**0.997 ± 0.000**	**0.988 ± 0.000**	**0.997 ± 0.000**	**0.994 ± 0.000**	**0.995 ± 0.000**	**0.996 ± 0.000**
p-value	1.811e-42	1.601e-20	5.223e-42	1.097e-42	2.251e-42	9.961e-16

Table 4
Comparison of experimental results on Caltech dataset.

Methods	ACC	NMI	Purity	ARI	F-score	Precision
SPC_{best}	0.424 ± 0.010	0.540 ± 0.006	0.762 ± 0.008	0.310 ± 0.007	0.367 ± 0.007	0.719 ± 0.011
MCGC	0.537 ± 0.000	0.586 ± 0.000	0.721 ± 0.000	0.392 ± 0.000	0.480 ± 0.000	0.541 ± 0.000
LMVSC	0.505 ± 0.000	0.579 ± 0.000	0.774 ± 0.000	0.390 ± 0.000	0.458 ± 0.000	0.672 ± 0.000
MCLES	0.452 ± 0.015	0.595 ± 0.029	0.731 ± 0.000	0.226 ± 0.031	0.333 ± 0.023	0.392 ± 0.017
LTMSC	0.529 ± 0.047	0.598 ± 0.021	0.789 ± 0.019	0.419 ± 0.050	0.476 ± 0.049	0.788 ± 0.036
t-SVD-MSC	0.613 ± 0.029	0.722 ± 0.010	0.702 ± 0.021	0.486 ± 0.032	0.537 ± 0.031	0.385 ± 0.029
ETLMSC	0.476 ± 0.037	0.671 ± 0.022	0.841 ± 0.020	0.362 ± 0.030	0.415 ± 0.027	0.814 ± 0.042
TISRL	0.553 ± 0.035	0.719 ± 0.017	0.829 ± 0.013	0.426 ± 0.034	0.476 ± 0.033	0.803 ± 0.027
HLR-M^2VS	0.564 ± 0.032	0.653 ± 0.012	0.826 ± 0.010	0.458 ± 0.034	0.512 ± 0.033	0.834 ± 0.020
CGL	0.608 ± 0.020	0.709 ± 0.007	0.861 ± 0.006	0.509 ± 0.015	0.560 ± 0.014	0.869 ± 0.017
LTALS	**0.633 ± 0.023**	**0.758 ± 0.020**	**0.899 ± 0.012**	**0.548 ± 0.019**	**0.598 ± 0.014**	**0.883 ± 0.015**
p-value	8.173e-3	1.323e-7	3.870e-10	7.767e-3	1.230e-2	1.401e-5

Table 5
Comparison of experimental results on NUS-WIDE dataset.

Methods	ACC	NMI	Purity	ARI	F-score	Precision
SPC_{best}	0.120 ± 0.006	0.142 ± 0.004	0.206 ± 0.006	0.023 ± 0.003	0.066 ± 0.003	0.090 ± 0.003
MCGC	0.151 ± 0.000	0.159 ± 0.000	0.216 ± 0.000	0.033 ± 0.000	0.092 ± 0.000	0.089 ± 0.000
LMVSC	0.143 ± 0.000	0.159 ± 0.000	0.224 ± 0.000	0.036 ± 0.000	0.077 ± 0.000	0.107 ± 0.000
MCLES	0.169 ± 0.005	0.152 ± 0.003	0.233 ± 0.005	0.039 ± 0.002	0.095 ± 0.002	0.107 ± 0.005
LTMSC	0.147 ± 0.001	0.160 ± 0.003	0.228 ± 0.002	0.035 ± 0.001	0.078 ± 0.002	0.104 ± 0.001
t-SVD-MSC	0.139 ± 0.004	0.163 ± 0.003	0.231 ± 0.003	0.034 ± 0.003	0.077 ± 0.003	0.101 ± 0.003
ETLMSC	0.131 ± 0.005	0.150 ± 0.006	0.234 ± 0.006	0.032 ± 0.004	0.073 ± 0.004	0.102 ± 0.005
TISRL	0.139 ± 0.000	0.165 ± 0.000	0.234 ± 0.000	0.033 ± 0.000	0.076 ± 0.000	0.101 ± 0.000
HLR-M^2VS	0.142 ± 0.000	0.164 ± 0.000	0.235 ± 0.000	0.036 ± 0.000	0.078 ± 0.000	0.105 ± 0.000
CGL	0.144 ± 0.004	**0.175 ± 0.003**	**0.239 ± 0.003**	0.038 ± 0.002	0.080 ± 0.002	**0.109 ± 0.003**
LTALS	**0.202 ± 0.004**	0.084 ± 0.003	0.210 ± 0.004	**0.042 ± 0.004**	**0.143 ± 0.002**	0.080 ± 0.003
p-value	4.360e-14	1.349e-7	3.842e-7	3.020e-1	3.386e-11	3.338e-1

Table 6
Comparison of experimental results on UCI dataset.

Methods	ACC	NMI	Purity	ARI	F-score	Precision
SPC_{best}	0.714 ± 0.000	0.649 ± 0.000	0.714 ± 0.000	0.552 ± 0.000	0.596 ± 0.000	0.595 ± 0.000
MCGC	0.836 ± 0.000	0.791 ± 0.000	0.836 ± 0.000	0.732 ± 0.000	0.759 ± 0.000	0.750 ± 0.000
LMVSC	0.825 ± 0.000	0.731 ± 0.000	0.825 ± 0.000	0.662 ± 0.000	0.696 ± 0.000	0.690 ± 0.000
MCLES	0.931 ± 0.012	0.868 ± 0.013	0.931 ± 0.017	0.856 ± 0.011	0.871 ± 0.015	0.869 ± 0.012
LTMSC	0.798 ± 0.009	0.767 ± 0.010	0.815 ± 0.008	0.718 ± 0.014	0.746 ± 0.013	0.733 ± 0.012
t-SVD-MSC	0.955 ± 0.000	0.932 ± 0.000	0.955 ± 0.000	0.924 ± 0.000	0.932 ± 0.000	0.930 ± 0.000
ETLMSC	0.926 ± 0.036	0.923 ± 0.030	0.936 ± 0.030	0.900 ± 0.039	0.911 ± 0.032	0.889 ± 0.040
TISRL	0.953 ± 0.000	0.931 ± 0.000	0.953 ± 0.000	0.912 ± 0.000	0.921 ± 0.000	0.920 ± 0.000
HLR-M^2VS	0.856 ± 0.001	0.870 ± 0.001	0.862 ± 0.001	0.819 ± 0.001	0.837 ± 0.001	0.820 ± 0.004
CGL	0.849 ± 0.017	0.931 ± 0.016	0.891 ± 0.018	0.856 ± 0.028	0.871 ± 0.025	0.813 ± 0.028
LTALS	**0.995 ± 0.000**	**0.987 ± 0.000**	**0.995 ± 0.000**	**0.989 ± 0.000**	**0.990 ± 0.000**	**0.990 ± 0.000**
p-value	9.298e-39	2.588e-37	9.298e-39	2.743e-39	2.710e-39	6.965e-39

Table 7
Comparison of experimental results on WikipediaArticles dataset.

Methods	ACC	NMI	Purity	ARI	F-score	Precision
SPC_{best}	0.552 ± 0.001	0.519 ± 0.004	0.600 ± 0.001	0.410 ± 0.000	0.473 ± 0.000	0.485 ± 0.000
MCGC	0.502 ± 0.000	0.418 ± 0.000	0.528 ± 0.000	0.265 ± 0.000	0.362 ± 0.000	0.299 ± 0.000
LMVSC	0.556 ± 0.000	0.475 ± 0.000	0.570 ± 0.000	0.331 ± 0.000	0.410 ± 0.000	0.380 ± 0.000
MCLES	0.543 ± 0.003	0.474 ± 0.004	0.563 ± 0.003	0.359 ± 0.005	0.430 ± 0.004	0.421 ± 0.005
LTMSC	0.531 ± 0.003	0.495 ± 0.005	0.575 ± 0.003	0.407 ± 0.002	0.471 ± 0.002	0.479 ± 0.002
t-SVD-MSC	0.556 ± 0.001	0.484 ± 0.002	0.580 ± 0.002	0.408 ± 0.001	0.471 ± 0.001	0.480 ± 0.001
ETLMSC	0.547 ± 0.032	0.495 ± 0.025	0.581 ± 0.036	0.400 ± 0.031	0.464 ± 0.026	0.473 ± 0.037
TISRL	0.547 ± 0.000	0.490 ± 0.000	0.587 ± 0.000	0.394 ± 0.000	0.458 ± 0.000	0.475 ± 0.000
HLR-M^2VS	0.546 ± 0.000	0.482 ± 0.000	0.574 ± 0.000	0.408 ± 0.000	0.471 ± 0.000	0.478 ± 0.000
CGL	0.542 ± 0.001	0.498 ± 0.000	0.594 ± 0.001	0.371 ± 0.001	0.441 ± 0.001	0.432 ± 0.001
LTALS	**0.564 ± 0.001**	**0.521 ± 0.002**	**0.601 ± 0.002**	**0.420 ± 0.002**	**0.483 ± 0.002**	**0.488 ± 0.003**
p-value	5.956e-10	3.731e-4	1.412e-4	2.324e-9	2.078e-10	9.094e-2

Table 8
Comparison of experimental results on Youtube dataset.

Methods	ACC	NMI	Purity	ARI	F-score	Precision
SPC_{best}	0.263 ± 0.000	0.171 ± 0.000	0.290 ± 0.000	0.096 ± 0.000	0.188 ± 0.000	0.183 ± 0.000
MCGC	0.284 ± 0.000	0.137 ± 0.000	0.297 ± 0.000	0.083 ± 0.000	0.176 ± 0.000	0.173 ± 0.000
LMVSC	0.272 ± 0.000	0.148 ± 0.000	0.292 ± 0.000	0.079 ± 0.000	0.173 ± 0.000	0.169 ± 0.000
MCLES	-	-	-	-	-	-
LTMSC	0.300 ± 0.001	0.184 ± 0.000	0.321 ± 0.001	0.112 ± 0.000	0.203 ± 0.000	0.198 ± 0.000
t-SVD-MSC	0.263 ± 0.017	0.162 ± 0.009	0.334 ± 0.011	0.098 ± 0.005	0.193 ± 0.005	0.180 ± 0.005
ETLMSC	0.266 ± 0.015	0.147 ± 0.006	0.294 ± 0.012	0.084 ± 0.006	0.180 ± 0.005	0.171 ± 0.007
TISRL	0.305 ± 0.001	0.178 ± 0.002	0.341 ± 0.001	0.111 ± 0.001	0.201 ± 0.001	0.197 ± 0.001
HLR-M^2VS	0.257 ± 0.002	0.143 ± 0.001	0.282 ± 0.001	0.081 ± 0.000	0.177 ± 0.000	0.168 ± 0.000
CGL	-	-	-	-	-	-
LTALS	**0.323 ± 0.002**	**0.199 ± 0.001**	**0.355 ± 0.002**	**0.135 ± 0.001**	**0.230 ± 0.001**	**0.208 ± 0.001**
p-value	5.510e-15	3.037e-9	6.499e-14	1.206e-13	1.959e-16	8.768e-9

Table 9
Running time (in seconds) of diverse multi-view approaches on six datasets.

Method	MCGC	LMVSC	MCLES	LTMSC	t-SVD-MSC	ETLMSC	TISRL	HLR-M^2VS	CGL	LTALS
BBCnews	3.68	2.22	4205.90	175.75	190.12	10.34	95.16	50.71	23.98	93.59
Caltech	133.78	8.07	93754.21	1968.00	843.05	146.12	1727.90	1344.50	595.05	2342.20
NUS-WIDE	162.89	2.65	14811.21	971.25	311.69	154.34	740.43	740.16	382.82	919.30
UCI	53.58	2.59	46567.20	1143.20	310.20	80.97	449.44	441.40	197.69	587.04
WikipediaArticles	0.26	0.47	141.50	8.63	3.02	0.87	5.51	2.78	13.62	4.24
Youtube	88.56	3.40	-	2029.10	407.60	309.33	1060.30	962.50	-	959.62

significant improvements compared with other methods can be seen from the clustering results. Fourthly, it can be seen that the *p*-value corresponding to each dataset is very small, which means that the proposed LTALS has positive significance in improving the clustering results compared to its closest competitor. Fifthly, Table 9 shows the running time of each multi-

view algorithm on six datasets. Obviously, the running efficiency of MCGC, LMVSC, and ETLMSC is much ahead of other methods. However, on the basis of excellent clustering results, the running efficiency of LTALS is also acceptable. Finally, compared with t-SVD-MSC, TISRL, and HLR-M^{2}VS, which are the most correlated with our method, LTALS also performs well, the leading situation may benefit from the unified integration of the rank preserving decomposition, the local structure protection, and the joint optimization of the label indicator matrix and low-rank tensor representation.

In addition, for observing the algorithm performance intuitively, we visualize the clustering results of eight selected methods (i.e., SPC$_{\text{best}}$, LMVSC, LTMSC, t-SVD-MSC, TISRL, ETLMSC, HLR-M^{2}VS, and LTALS) on UCI dataset in Fig. 3. Specifically, we splice the three feature matrices in UCI together, then the dimension reduction method t-SNE is utilized to map the high-dimensional features onto a two-dimensional space. In Fig. 3, different colors represent different clusters. It can be seen that LTALS achieves the clearest segmentation, while SPC$_{\text{best}}$ are the worst, which corresponds to the values of their evaluation metrics.

4.5. Parameter Sensitivity Analysis

In the proposed LTALS, there are three hyperparameters: λ, β, and α. To find the appropriate parameter values, we perform a two-level grid search strategy, i.e., coarse-grained and fine-grained strategies. In coarse-grained strategy, we randomly select the BBCnews dataset as the test dataset and empirically vary λ, β, and α in {0.0001, 0.001, 0.01, 0.1, 1, 10} to find approximate parameter intervals, respectively. We observe that the proposed LTALS achieves acceptable numerical results when λ, β, and α are tuned in {0.001, 0.01, 0.1, 1}, {0.001, 0.01, 0.1, 1}, and {0.0001, 0.001, 0.01}, respectively. Since the characteristics of different datasets are not same, in fine-grained strategy, we perform a fine-grained grid search strategy on each dataset by varying the three papameters according to the intervals obtained in coarse-grained strategy, thus obtaining more specific and suitable parameter interval for each dataset. In general, λ is set in [0.05, 0.2] with a step of 0.01, β is ranged in [0.01, 0.1] with a step of 0.01, α is tuned in [0.0001,0.005] with a step of 0.0001, then the proposed LTALS can yield the promising experimental results. In addition, the number of k-nearest neighbors is fixed to 6.

For observing the impact of diverse parameters on the model LTALS, we display the clustering performance with respect to different parameter settings in Figs. 4, 5. As shown in Fig. 4, λ, β, and α are tuned in {0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.4, 0.8, 1}, {0.001, 0.01, 0.1, 1}, and {0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.15, 0.2}, respectively. Generally, when α is set large, the performance of LTALS is unsatisfactory. The main reason may be that the label indicator matrix is not accurate enough in the early stage of optimization, and too large value of α causes the deviation of optimization of both low-rank tensor and label indicator matrix, thus resulting in inferior performance. For λ, which is mainly used to adjust the impact of noise, its value is not suggested to be set too large from Fig. 4. Fig. 5 exhibits the effect of the numbers of k-nearest neighbors on the performance of LTALS, it can be observed that the performance of LTALS is stable as the number of k-nearest neighbors continues to increase.

4.6. Ablation Experiments

To illustrate the significance of the rank preserving decomposition, the local structure protection, and the joint optimization of label indicator matrix and low-rank tensor representation, we design some ablation experiments to validate. Specif-

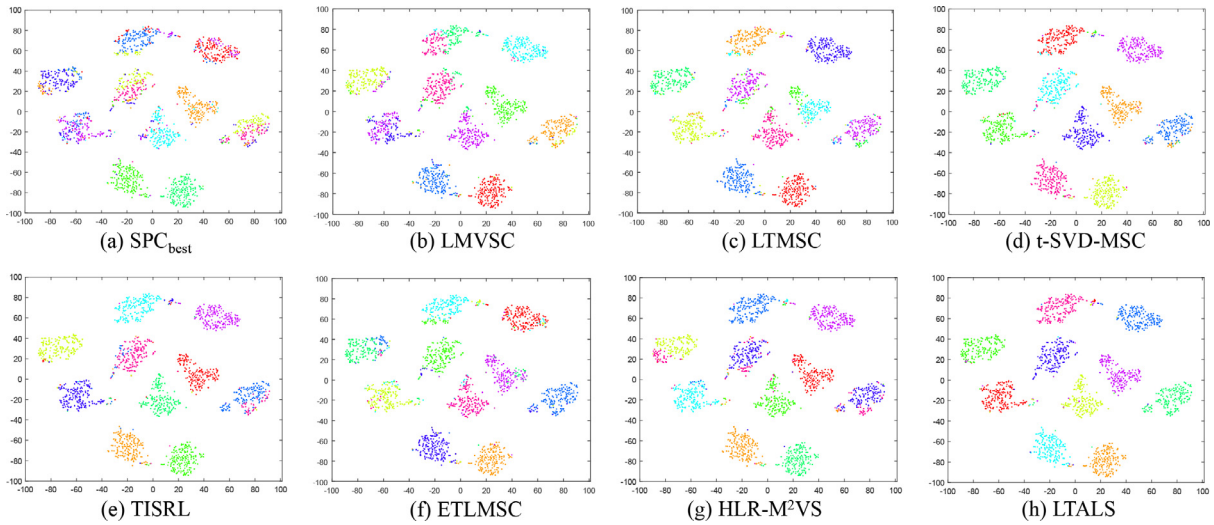

Fig. 3. Visualizations of eight clustering methods via t-SNE on UCI dataset.

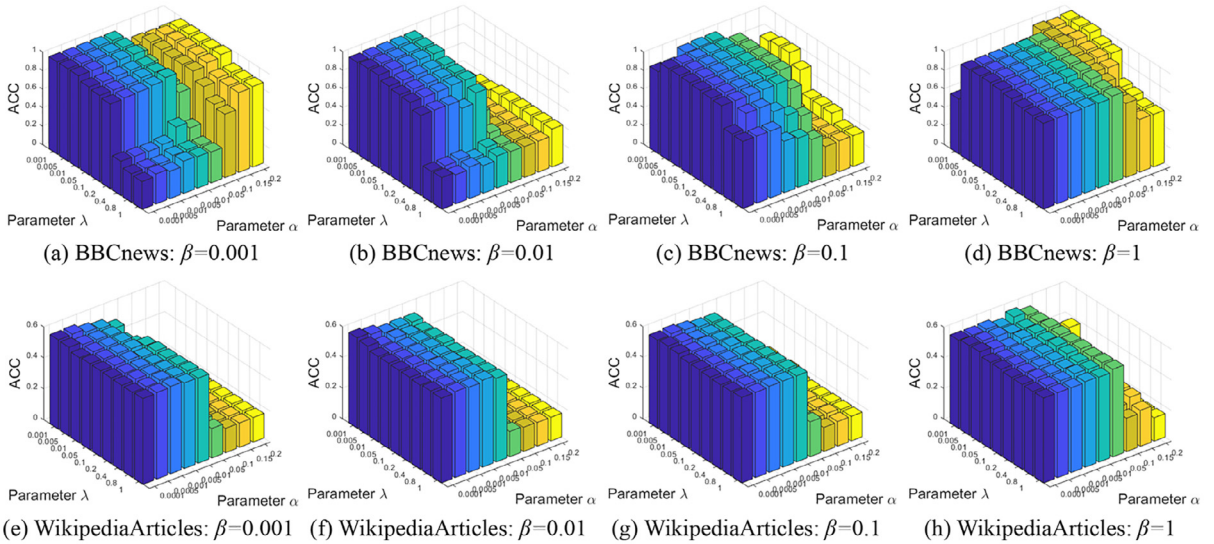

Fig. 4. ACC of the proposed method on BBCnews and WikipediaArticles datasets while λ ranges in {0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.4, 0.8, 1}, α ranges in {0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.15, 0.2}, and β varies in {0.001, 0.01, 0.1, 1}.

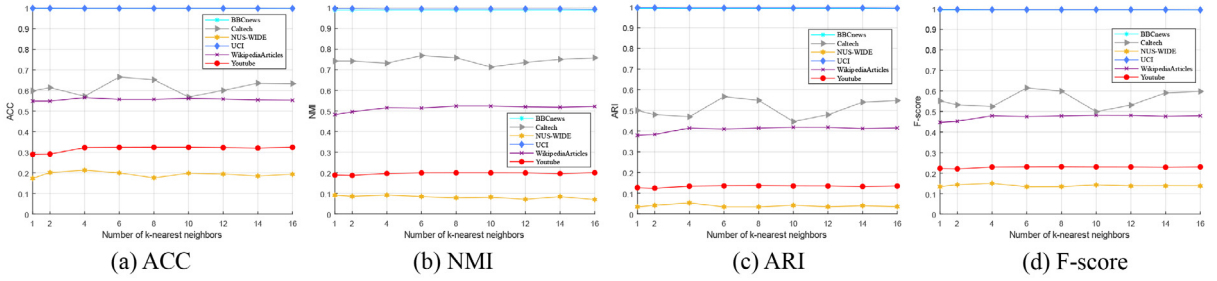

Fig. 5. Performance of the proposed LTALS with respect to different numbers of k-nearest neighbors.

ically, LTALS-OR represents the variant of LTALS that does not use the rank preserving decomposition to learn the intrinsic subspace representations $\left\{\mathbf{S}^{(v)}\right\}_{v=1}^{m}$, LTALS-OL represents the variant of LTALS without protecting the local structure, i.e., the third term in the objective function Eq. (6) is removed, and LTALS-OF denotes the variant of LTALS removing the joint optimization of label indicator matrix and low-rank tensor representation, that is, the fourth term in the objective function Eq. (6) is discarded. Furthermore, it is worth noting that if the rank preserving decomposition, the local structure protection, and the joint optimization of label indicator matrix and low-rank tensor representation are simultaneously cast off, the proposed LTALS degenerates to the t-SVD-MSC model [12]. Similarly, if the latter two are not considered, the proposed LTALS reduces to the TISRL model [18]. The clustering results of t-SVD-MSC and TISRL can be seen in Tables 3–8, which are not repeated in this subsection. As Table 10 shown, it is notable that the experimental results of LTALS are better than the three variants LTALS-OR, LTALS-OL, and LTALS-OF on all datasets. For example, LTALS improves the results by 7.5% and 2.9%, 10.9% and 13.4%, 10.8% and 7.1% on Caltech dataset compared with the three variants, respectively, which demonstrates that the three strategies are essential for boosting the clustering performance.

Table 10
Comparison of clustering results (ACC/NMI) of LTALS and its variants.

Datasets	BBCnews	Caltech	NUS-WIDE	UCI	WikipediaArticles	Youtube
LTALS-OR	0.327/0.005	0.549/0.701	0.123/0.015	0.885/0.814	0.175/0.036	0.102/0.005
LTALS-OL	0.977/0.937	0.515/0.596	0.166/0.069	0.721/0.684	0.549/0.478	0.293/0.195
LTALS-OF	0.945/0.892	0.516/0.659	0.136/**0.157**	0.866/0.878	0.544/0.496	0.290/0.164
LTALS	**0.997/0.988**	**0.633/0.758**	**0.202**/0.084	**0.995/0.987**	**0.564/0.521**	**0.323/0.199**

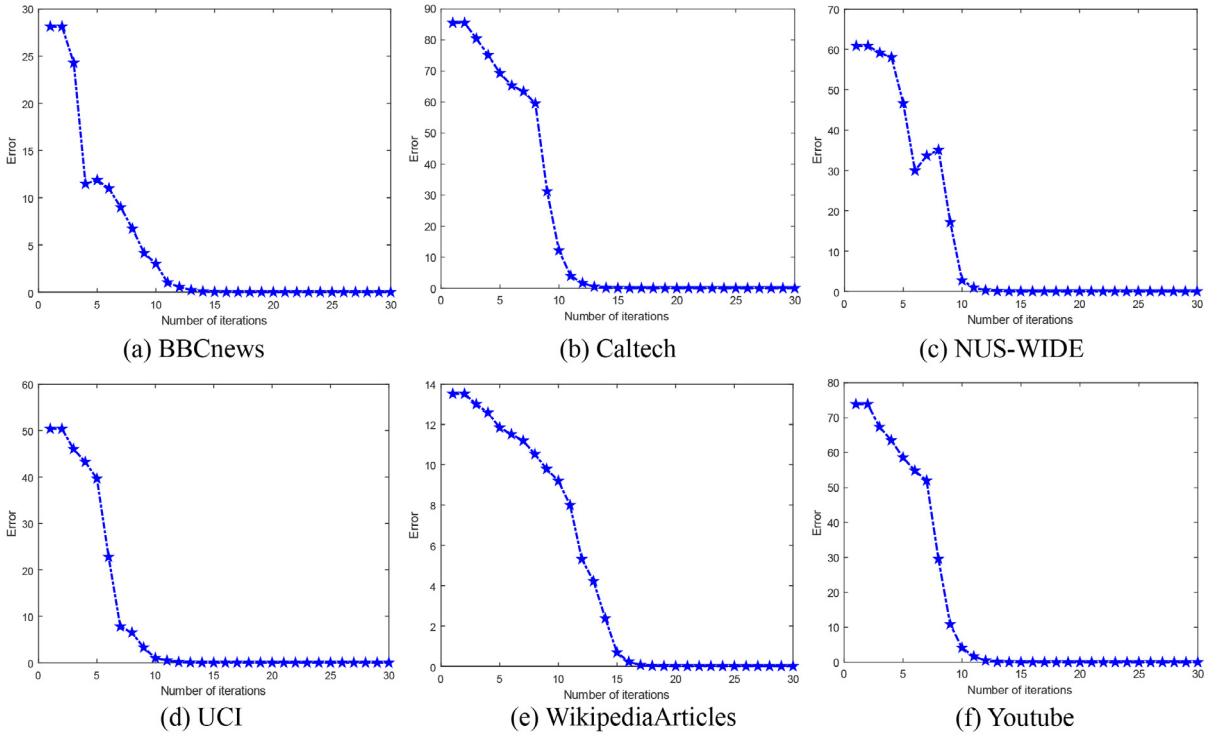

Fig. 6. Convergence curves on the test datasets.

4.7. Convergence Analysis

According to [46], since Algorithm 1 contains three or more blocks variables, it is not certain whether the proposed method LTALS will converge after iterations in theory. Nevertheless, the experiments can still demonstrate the good convergence property of LTALS. Fig. 6 shows the convergence curves on all datasets, where the x-axis represents the number of iterations, the y-axis error is defined as the maximum value of $\sum_{v=1}^{m}||\mathbf{X}^{(v)} - \mathbf{X}^{(v)}\mathbf{Z}_{k+1}^{(v)} - \mathbf{E}_{k+1}^{(v)}||_{\infty}$ and $\sum_{v=1}^{m}||\mathbf{S}_{k+1}^{(v)} - \mathbf{H}_{k+1}^{(v)}||_{\infty}$. As shown in Fig. 6, the values of error decrease very quickly. On the six datasets, twenty iterations are sufficient to bring the proposed LTALS to a state of convergence.

5. Conclusion

In this paper, we propose a low-rank tensor approximation with local structure (LTALS) for multi-view intrinsic subspace clustering. In the proposed LTALS, we perform rank preserving decomposition on the initial self-representation matrices to factorize out the intrinsic subspace representations, which are assembled into a 3-order target tensor. Then the t-SVD based TNN is imposed on the target tensor to strength the global consensus. In order to achieve the consistent locality during the projection of feature space, we adopt the manifold regularization to encode the local structure of data in the initial feature space. Moreover, we recognize that the learned label indicator matrix contains the cluster structure and use it to boost the optimization of the low-rank intrinsic subspace representation tensor, thus achieving simultaneous optimization of the two terms in a unified framework. For solving the proposed LTALS, an optimization algorithm based on ADMM is developed. Finally, a large number of experimental results on six real-world datasets show that the performance of LTALS is better than other state-of-the-art methods. In the future, as the size of the datasets continue to grow, we will further optimize the solution process of the proposed LTALS to accommodate large-scale datasets.

CRediT authorship contribution statement

Lele Fu: Conceptualization, Formal analysis, Methodology, Writing - original draft. **Jinghua Yang:** Conceptualization, Formal analysis, Methodology, Writing - review & editing. **Chuan Chen:** Supervision, Validation, Visualization. **Chuanfu Zhang:** Supervision, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (11801595), the Guang dong Basic and Applied Basic Research Foundation (2019A1515011043), the Natural Science Foundation of Guangdong (2018A030310076), and Macao Science and Technology Development Fund under Macao Funding Scheme for Key R & D Projects (0025/2019/AKP).

Appendix A. Preliminaries about Tensors

Definition 1. For a tensor $\mathscr{T} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its several basic operations $bdiag(\cdot), bcirc(\cdot), bvec(\cdot)$, and $bvfold(\cdot)$ are defined as

$$bdiag(\mathscr{T}) = \begin{bmatrix} \mathscr{T}^{(1)} & & & \\ & \mathscr{T}^{(2)} & & \\ & & \ddots & \\ & & & \mathscr{T}^{(n_3)} \end{bmatrix}, \tag{A.1}$$

$$bcirc(\mathscr{T}) = \begin{bmatrix} \mathscr{T}^{(1)} & \mathscr{T}^{(n_3)} & \cdots & \mathscr{T}^{(2)} \\ \mathscr{T}^{(2)} & \mathscr{T}^{(1)} & \cdots & \mathscr{T}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathscr{T}^{(n_3)} & \mathscr{T}^{(n_3-1)} & \cdots & \mathscr{T}^{(1)} \end{bmatrix}, \tag{A.2}$$

$$bvec(\mathscr{T}) = \left[\mathscr{T}^{(1)}; \mathscr{T}^{(2)}; \cdots; \mathscr{T}^{(n_3)}\right], \tag{A.3}$$

$$bvfold(bvec(\mathscr{T})) = \mathscr{T}. \tag{A.4}$$

Definition 2 (*t-product*). For tensors $\mathscr{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathscr{Y} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$, the t-product $\mathscr{A} \in \mathbb{R}^{n_1 \times n_4 \times n_3}$ is calculated via

$$\mathscr{A} = \mathscr{X} * \mathscr{Y} = bvfold(bcirc(\mathscr{X}) \cdot bvec(\mathscr{Y})), \tag{A.5}$$

Definition 3 (*Orthogonal tensor*). A tensor $\mathscr{T} \in \mathbb{R}^{n_1 \times n_1 \times n_2}$ is orthogonal if it satisfies

$$\mathscr{T}^T * \mathscr{T} = \mathscr{T} * \mathscr{T}^T = \mathscr{I}, \tag{A.6}$$

where tensor $\mathscr{I} \in \mathbb{R}^{n_1 \times n_1 \times n_2}$ is an identity tensor, whose first frontal slice is a $n_1 \times n_1$ identity matrix and the other frontal slices are all zeros.

Definition 4 (*t-SVD*). A tensor $\mathscr{C} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ can be decomposed as

$$\mathscr{C} = \mathscr{U} * \mathscr{D} * \mathscr{V}^T, \tag{A.7}$$

where $\mathscr{D} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a f-diagonal tensor, whose each frontal slice is a diagonal matrix. $\mathscr{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathscr{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal.

References

[1] Rongkai Xia, Yan Pan, Du. Lei, Jian Yin, Robust multi-view spectral clustering via low-rank and sparse decomposition, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2014, pp. 2149–2155.
[2] Fu. Lele, Pengfei Lin, Athanasios V Vasilakos, Shiping Wang, An overview of recent multi-view clustering, Neurocomputing 402 (2020) 148–161.
[3] Shudong Huang, Xu. Zenglin, Ivor W. Tsang, Zhao Kang, Auto-weighted multi-view co-clustering with bipartite graphs, Information Sciences 512 (2020) 18–30.
[4] Shiping Wang, Fu. Lele, Zhewen Wang, Xu. Haiping, William Zhu, Multigraph random walk for joint learning of multiview clustering and semisupervised classification, IEEE Transactions on Computational Social Systems (2021), https://doi.org/10.1109/TCSS.2021.3109151.

[5] Ehsan Elhamifar, Rene Vidal, Sparse subspace clustering: Algorithm, theory, and applications, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (11) (2013) 2765–2781.
[6] Guangcan Liu, Zhouchen Lin, Ju. Shuicheng Yan, Yong Yu Sun, Yi Ma, Robust recovery of subspace structures by low-rank representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (1) (2013) 171–184.
[7] Ruihuang Li, Changqing Zhang, Qinghua Hu, Pengfei Zhu, Zheng Wang, Flexible multi-view representation learning for subspace clustering, in: Proceedings of the International Joint Conference on Artificial Intelligence, 2019, pp. 2916–2922.
[8] Changqing Zhang, Fu. Huazhu, Hu. Qinghua, Xiaochun Cao, Yuan Xie, Dacheng Tao, Xu. Dong, Generalized latent multi-view subspace clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence 42 (1) (2020) 86–99.
[9] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. Large-scale multi-view subspace clustering in linear time. In Proceedings of the International Joint Conferences on Artificial Intelligence Organization, pages 4412–4419, 2020..
[10] Yongyong Chen, Shuqin Wang, Fangying Zheng, Yigang Cen, Graph-regularized least squares regression for multi-view subspace clustering, Knowledge-Based Systems 194 (2020) 105482.
[11] Changqing Zhang, Fu. Huazhu, Si Liu, Guangcan Liu, Xiaochun Cao, Low-rank tensor constrained multiview subspace clustering, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1582–1590.
[12] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, Qu. Yanyun, On unifying multi-view self-representations for clustering by tensor multi-rank minimization, International Journal of Computer Vision 126 (11) (2018) 1157–1179.
[13] Quanxue Gao, Wei Xia, Zhizhen Wan, De-Yan Xie, and Pu Zhang. Tensor-svd based graph learning for multi-view subspace clustering. In Proceedings of the AAAI Conference on Artificial Intelligence, pages 3930–3937, 2020..
[14] Yongyong Chen, Shuqin Wang, Chong Peng, Zhongyun Hua, Yicong Zhou, Generalized nonconvex low-rank tensor approximation for multi-view subspace clustering, IEEE Transactions on Image Processing 30 (2021) 4022–4035.
[15] Chang Xu, Dacheng Tao, and Chao Xu. A survey on multi-view learning. arXiv preprint arXiv:1304.5634, 2013..
[16] Ming Yin, Junbin Gao, Zhouchen Lin, Laplacian regularized low-rank representation and its applications, IEEE Transactions Pattern Analysis Machine Intelligence 38 (3) (2016) 504–517.
[17] Wu. Jianlong, Zhouchen Lin, Hongbin Zha, Essential tensor learning for multi-view spectral clustering, IEEE Transactions on Image Processing 28 (12) (2019) 5910–5922.
[18] Qinghai Zheng, Jihua Zhu, Zhongyu Li, Haoyu Tang, and Shuangxun Ma. Tensor-based intrinsic subspace representation learning for multi-view clustering. arXiv preprint arXiv:2010.09193, 2020..
[19] Qianyao Qiang, Bin Zhang, Fei Wang, Feiping Nie, Fast multi-view discrete clustering with anchor graphs, in: Proceedings of AAAI Conference on Artificial Intelligence, 2021, pp. 9360–9367.
[20] Yu. Xiao, Hui Liu, Wu. Yan, Caiming Zhang, Fine-grained similarity fusion for multi-view spectral clustering, Information Sciences 568 (2021) 350–368.
[21] Rong Wang, Feiping Nie, Zhen Wang, Hu. Haojie, Xuelong Li, Parameter-free weighted multi-view projected clustering with structured graph learning, IEEE Transactions on Knowledge and Data Engineering 32 (10) (2020) 2014–2025.
[22] Peiguang Jing, Su. Yuting, Zhengnan Li, Liqiang Nie, Learning robust affinity graph representation for multi-view clustering, Information Sciences 544 (2021) 155–167.
[23] Xiaochun Cao, Changqing Zhang, Fu. Huazhu, Si Liu, Hu.a. Zhang, Diversity-induced multi-view subspace clustering, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 586–594.
[24] Xiaobo Wang, Xiaojie Guo, Zhen Lei, Changqing Zhang, Stan Z Li, Exclusivity-consistency regularized multi-view subspace clustering, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 923–931.
[25] Shirui Luo, Changqing Zhang, Wei Zhang, Xiaochun Cao, Consistent and specific multi-view subspace clustering, in: Proceedings of the International Joint Conferences on Artificial Intelligence Organization, 2018, pp. 3730–3737.
[26] Xiaoqian Zhang, Zhenwen Ren, Huaijiang Sun, Keqiang Bai, Xinghua Feng, Zhigui Liu, Multiple kernel low-rank representation-based robust multi-view subspace clustering, Information Sciences 551 (2021) 324–340.
[27] Wentao Rong, Enhong Zhuo, Hong Peng, Jiazhou Chen, Haiyan Wang, Ch.u. Han, Hongmin Cai, Learning a consensus affinity matrix for multi-view clustering via subspaces merging on grassmann manifold, Information Sciences 547 (2021) 68–87.
[28] Juncheng Lv, Zhao Kang, Boyu Wang, Luping Ji, Xu. Zenglin, Multi-view subspace clustering via partition fusion, Information Sciences 560 (2021) 410–423.
[29] Haoran Li, Zhenwen Ren, Mithun Mukherjee, Yuqing Huang, Quansen Sun, Xingfeng Li, Liwan Chen, Robust energy preserving embedding for multi-view subspace clustering, Knowledge-Based Systems 210 (2020) 106489.
[30] GuangYu Zhang, YuRen Zhou, XiaoYu He, ChangDong Wang, Dong Huang, One-step kernel multi-view subspace clustering, Knowledge-Based Systems 189 (2020) 105126.
[31] Shuqin Wang, Yongyong Chen, Yi Jin, Yigang Cen, Yidong Li, Linna Zhang, Error-robust low-rank tensor approximation for multi-view clustering, Knowledge-Based Systems 215 (2021) 106745.
[32] Wu. Jianlong, Xingyu Xie, Liqiang Nie, Zhouchen Lin, Hongbin Zha, Unified graph and low-rank tensor learning for multi-view clustering, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pp. 6388–6395.
[33] Y. Chen, X. Xiao, Y. Zhou, Jointly learning kernel representation tensor and affinity matrix for multi-view clustering, IEEE Transactions on Multimedia 22 (8) (2020) 1985–1997.
[34] Changqing Zhang, Fu. Huazhu, Jing Wang, Wen Li, Xiaochun Cao, Hu. Qinghua, Tensorized multi-view subspace representation learning, International Journal of Computer Vision 128 (8) (2020) 2344–2361.
[35] Fu. Lele, Zhaoliang Chen, Sujia Huang, Sheng Huang, Shiping Wang, Multi-view learning via low-rank tensor optimization, in: 2021 IEEE International Conference on Multimedia and Expo, 2021, pp. 1–6.
[36] Zemin Zhang, Gregory Ely, Shuchin Aeron, Ning Hao, Misha Kilmer, Novel methods for multilinear data completion and de-noising based on tensor-svd, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 3842–3849.
[37] Shuicheng Yan, Xu. Dong, Benyu Zhang, Hong-Jiang Zhang, Qiang Yang, Stephen Lin, Graph embedding and extensions: A general framework for dimensionality reduction, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (1) (2006) 40–51.
[38] R.H. Bartels, G.W. Stewart, Solution of the matrix equation ax + xb = c [f4], Communications of the ACM 15 (9) (1972) 820–826.
[39] John C Gower, Garmt B Dijksterhuis, et al, Procrustes problems, volume 30, Oxford University Press on Demand, 2004.
[40] Hu. Wenrui, Dacheng Tao, Wensheng Zhang, Yuan Xie, Yehui Yang, The twist tensor nuclear norm for video completion, IEEE Transactions on Neural Networks and Learning Systems 28 (12) (2016) 2961–2973.
[41] Arthur Asuncion and David Newman. Uci machine learning repository, 2007..
[42] Kun Zhan, Feiping Nie, Jing Wang, Li Yang, Multiview consensus graph clustering, IEEE Transactions on Image Processing 28 (3) (2019) 1261–1270.
[43] Mansheng Chen, Ling Huang, Changdong Wang, Dong Huang, Multi-view clustering in latent embedding space, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pp. 3513–3520.
[44] Yuan Xie, Wensheng Zhang, Qu. Yanyun, Longquan Dai, Dacheng Tao, Hyper-laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning, IEEE Transactions Cybernetics 50 (2) (2020) 572–586.
[45] Zhenglai Li, Chang Tang, Xinwang Liu, Xiao Zheng, Guanghui Yue, Wei Zhang, En Zhu, Consensus graph learning for multi-view clustering, IEEE Transactions on Multimedia (2021), https://doi.org/10.1109/TMM.2021.3081930.
[46] Jonathan Eckstein, Dimitri P Bertsekas, On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators, Mathematical Programming 55 (1) (1992) 293–318.