Generalized Nonconvex Low-Rank Tensor Approximation for Multi-View Subspace Clustering

Yongyong Chen, Shuqin Wang, Chong Peng, Zhongyun Hua, *Member, IEEE*, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*— The low-rank tensor representation (LRTR) has become an emerging research direction to boost the multi-view clustering performance. This is because LRTR utilizes not only the pairwise relation between data points, but also the view relation of multiple views. However, there is one significant challenge: LRTR uses the tensor nuclear norm as the convex approximation but provides a biased estimation of the tensor rank function. To address this limitation, we propose the generalized nonconvex low-rank tensor approximation (GNLTA) for multi-view subspace clustering. Instead of the pairwise correlation, GNLTA adopts the low-rank tensor approximation to capture the high-order correlation among multiple views and proposes the generalized nonconvex low-rank tensor norm to well consider the physical meanings of different singular values. We develop a unified solver to solve the GNLTA model and prove that under mild conditions, any accumulation point is a stationary point of GNLTA. Extensive experiments on seven commonly used benchmark databases have demonstrated that the proposed GNLTA achieves better clustering performance over state-of-the-art methods.**

***Index Terms*— Multi-view clustering, nonconvex low-rank tensor approximation, spectral clustering, subspace clustering.**

I. Introduction

MULTI-VIEW learning has achieved the ideal performance in many applications such as face clustering [1], outlier detection [2], human action recognition [3], [4]. Different from the single-view learning based on one specific feature, different heterogeneous and complementary features are explored to investigate the correlations among data points in multi-view learning. One popular example is that an image can be represented by various types of features, *e.g.*, textures, color and edges. However, it is hard to obtain the label information in real applications. Thus, multi-view clustering has attracted great interests in recent years [1], [5].

Multi-view clustering (MVC) refers to the process of clustering a large number of unlabeled data points into several groups by using multiple features, so that similar samples are classified into the same group while dissimilar ones into different groups. Considering that high-dimensional data can be well represented by several low-dimensional subspaces, multi-view subspace clustering (MVSC) methods have been proposed and gathered a great deal of attention. Most of existing MVSC methods generally follow the similar pipeline: (1) construct the affinity matrix of the data points from the raw features or the learnt representation matrices; (2) perform the spectral clustering algorithm [6] on the affinity matrix and yield the clustering results. A core of these methods is how to exactly learn a proper affinity matrix at the first step.

A. Related Work

The first way to learn the affinity matrix is using the raw multi-view features. For example, Nie *et al*. [7], [8] used the Euclidean distance of raw features and the rank constraint to learn an affinity matrix with exact clusters to measure the similarity of pairwise data points. To eliminate noise, Zhan *et al*. [9], [10] developed to learn a consensus graph with minimizing disagreement between different views to improve the quality of the affinity matrix. Instead of fixed graph, Wang *et al*. [11] proposed to fuse data graph matrices into a unified graph. Similar idea was adopted by [9].

The second way is also the most popular one which constructs the affinity matrix by using the learnt representation matrices. Under the assumption that observed data usually lie in/near some low-dimensional subspaces, sparse subspace clustering (SSC) and low-rank representation (LRR) pursue the sparse and low-rank coefficients of the data points over the data itself by the l_1-norm and the nuclear norm, respectively. Reference [12] adopted the low-rank and sparse decomposition to learn a shared representation matrix. Cao *et al*. [13] utilized the Hilbert Schmidt Independence Criterion to pursue the diverse representations. Zhang *et al*. [14] combined the latent

Manuscript received January 6, 2021; revised March 3, 2021; accepted March 19, 2021. Date of publication March 30, 2021; date of current version April 5, 2021. This work was funded in part by the National Natural Science Foundation of China under Grant 62071142, in part by the Science and Technology Development Fund, Macau SAR (File no. 189/2017/A3), and by University of Macau (File no. MYRG2018-00136-FST). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lucio Marcenaro. *(Corresponding author: Zhongyun Hua.)*

Yongyong Chen is with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, also with the Bio-Computing Research Center, Harbin Institute of Technology, Shenzhen 518055, China, and also with the Shenzhen Key Laboratory of Visual Object Detection and Recognition, Shenzhen 518055, China (e-mail: YongyongChen.cn@gmail.com).

Shuqin Wang is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China (e-mail: ShuqinWang.cn@hotmaiil.com).

Chong Peng is with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China (e-mail: pchong1991@163.com).

Zhongyun Hua is with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China (e-mail: huazyum@gmail.com).

Yicong Zhou is with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: yicongzhou@um.edu.mo).

Digital Object Identifier 10.1109/TIP.2021.3068646

representation learning and the low-rank representation into a unified model. Following the studies in [13] and [14], the work [15] projected the high-dimensional multiple features onto a latent space and learned each view-specific representation. Instead of the popular nuclear norm, structured matrix factorization [16], deep matrix factorization [17], and least squares regression [18] with graph regularizer were used to pursue the ideal representations besides local manifold structures. In summary, these methods belong to the matrix optimization and explored only the pairwise correlation within views. Thus, they may not fully make use of the correlation among the view dimension [5].

Recently, by storing all representation matrices into a third-order tensor (called *representation tensor*), several methods for MVSC have achieved promising performance based on the low-rank tensor representation [1], [19], [20]. This is because the tensor learning is capable of extracting high-order information from a multidimensional perspective. For example, Zhang *et al*. [5] firstly proposed to construct the third-order representation tensor and extended the single-view LRR into MVSC. Along this line, the study in [1] explored a higher version of tensor nuclear norm instead of the unfolding-based tensor nuclear norm. The tensor nuclear norm was integrated with hyper-laplacian regularizer [21] and the kernel trick [22], [23] for MVSC. To simultaneously exploit multi-view features, the work in [24] organized the multi-view features into the third-order tensorial data which were recovered by the sparse and low-rank tensor decomposition. To explore the correlation between the representation tensor and the affinity matrix, Chen *et al*. [25] proposed to learn them jointly. All of them have the cubic complexity, which severely limits their scalability to big data. The reason is that they handle the MVSC task in a self-representation way and perform the matrix inversion inevitably. To reduce the computational complexity, Wu *et al*. [26] proposed an essential tensor learning method to avoid the matrix inversion. Following [26], the work in [27] proposed the adaptive transition probability matrix learning method. However, one common challenge is that all of them use the matrix nuclear norm and tensor nuclear norm to investigate the low-rank characteristic of the representations. For example, the studies in [12], [14] exploited the convex matrix nuclear norm while that of [1], [5], [26] used the convex tensor nuclear norm. Essentially, the nuclear norm is the l_1 norm of all singular values. To solve the nuclear norm-regularized subproblem, the singular value thresholding (SVT) operator was adopted to shrink all singular values with the same threshold. This means that both of the matrix nuclear norm and tensor nuclear norm treat each singular value equally and over-penalize top larger singular values. In fact, we expect to shrink less on the larger singular values since they usually represent the major information of the underneath data [28]. Meanwhile, the small singular values usually come from noise and should be more likely to be vanished. Unfortunately, the equal treatment to all singular values by the nuclear norm and the corresponding SVT algorithm would not provide the shrinkage effect as we expect and thus result in information loss. Moreover, the straight summation of all singular values cannot approximate the true rank function well if there are large singular values in a matrix and thus would result in a biased estimation of the rank function [28]–[30]. Several recent studies [31]–[33] have been proposed based on the benefit of nonconvex penalty functions for single-view subspace clustering. Some efficient single-view subspace clustering methods were also developed. For example, Peng *et al*. [34] proposed the scalable SSC in the sampling, clustering, coding, and classifying manner. Following this line, the study [35] proposed the selective sampling-based scalable SSC to reduce the complexity of the representation learning and spectral clustering steps. The work in [36] proposed the online low-rank subspace clustering by basis dictionary pursuit. When handling the multi-view features, they may lead to a compromised clustering performance, since they cannot fully exploit the correlation among multiple views. Therefore, there is an urgent need to design the nonconvex low-rank regularized multi-view clustering method to strengthen the ability of recovering latent patterns from multi-view data. We attempt to overcome the above limitation by integrating the nonconvex optimization into the low-rank tensor learning and propose the generalized nonconvex low-rank tensor approximation method (GNLTA). Unlike these nonconvex single-view subspace clustering methods [31]–[33], our GNLTA not only utilizes the pairwise relation between data points, but also the view dimension by the tensor optimization technique. In addition, most existing nonconvex subspace clustering methods were solved case-by-case. We also provide a general solver with theoretical guarantee.

Deep multi-view clustering approaches have recently flourished. For example, Ji *et al*. [37] proposed the deep subspace clustering network by introducing a novel self-expressive layer to mimic the self-representation property. Following this line, the deep multimodal subspace clustering networks was proposed in [38] which consists of multimodal encoder, self-expressive layer, and multimodal decoder. The work in [39] developed the reciprocal multi-layer subspace learning for MVC under the auto-encode framework. Xie *et al*. [40] adopted the joint learning strategy to utilize the multi-view features and complementary information. Although deep clustering methods have shown impressive performance, they often rely heavily on a large number of training samples [41]. This means that it is still intractable to well train a network for MVC task.

B. Our Contributions

Inspired by the huge success of the nonconvex low-rank matrix approximation in image denoising [28], [29] and the power of low-rank tensor learning, we investigate to extend the nonconvex low-rank matrix approximation into the tensor to deal with the MVSC task. To this end, we propose a Generalized Nonconvex Low-rank Tensor Approximation (GNLTA) for multi-view subspace clustering. Different from the existing tensor-based MVC methods based on the convex matrix and tensor nuclear norm, GNLTA proposes one general nonconvex low-rank tensor approximation for better capture of the high correlation among multiple views. In addition, GNLTA learns the representation from the similarity matrices

TABLE I
BASIC NOTATIONS AND THEIR DESCRIPTIONS

Notation	Meaning
$\mathcal{X}$, X, x	tensor, matrix, vector
$\mathcal{X}^{(k)}$	the k-th frontal slice of tensor $\mathcal{X}$
$\hat{\mathcal{X}} = \mathbf{fft}(\mathcal{X}, [], 3)$	fast Fourier transformation along tube fiber
n, M	the number of samples, views
d_v	feature dimension of the v-th view
$X^{(v)} \in \mathbb{R}^{d_v \times n}$	feature matrix of the v-th view
$\mathcal{Z} \in \mathbb{R}^{n \times n \times M}$	the representation tensor
$P \in \mathbb{R}^{n \times n \times M}$	the similarity tensor
$E^{(v)} \in \mathbb{R}^{n \times n}$	the sample-specific corruptions
$\|\cdot\|_{2,1}$, $\|\cdot\|_F$	$l_{2,1}$-norm, Frobenius norm
$\|\cdot\|_{\circledast}$, $\|\cdot\|_\infty$	t-SVD-nuclear norm, infinity norm

and avoids the matrix inversion perfectly. Thus, the proposed GNLTA costs less running time and has the better tensor learning for multi-view subspace clustering.

The contributions and novelty of this paper are summarized as follows:

- To address the biased estimation of the convex nuclear norm, we propose a new GNLTA model for multi-view subspace clustering. Instead of treating each singular value equally in existing methods, GNLTA assigns different weights to different singular values using the nonconvex functions such that larger singular values still dominate more important information of the representation tensor. Thus, the main idea of GNLTA is totally different from existing ones.
- Based on the alternating direction method of multipliers and the difference of convex method, we provide a general solver for the GNLTA model, which is significantly different from most existing methods that solve them case-by-case.
- From the theoretical aspect, we prove that under mild conditions, any accumulation point is a stationary point of GNLTA. Extensive experiments demonstrate that the GNLTA model has lower computational time and better clustering performance than twenty state-of-the-art methods on seven popular multi-view datasets.

C. Organization of This Paper

The rest of this paper is structured as follows. Some notations and preliminaries are introduced in Section II. Section III presents our generalized nonconvex low-rank tensor approximation (GNLTA) model and designs an iterative algorithm under the alternating direction method of multipliers framework and the difference of convex method. We provide the experimental performance of the proposed GNLTA on seven real-world multi-view databases in Section IV and finally conclude the whole paper in Section V.

II. NOTATIONS AND PRELIMINARIES

We use calligraphy letters, capital letters, and lowercase letters (*e.g.*, $\mathcal{X}$, X, x) to denote the tensors, matrices, and vectors, respectively. The frequently used notations are summarized in Table I.

The rest of this section aims to describe the tensor nuclear norm [42] which has widely adopted to depict the low-rank property for multi-view clustering. For a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its block circular matrix $\mathbf{bcirc}(\mathcal{X})$ and block diagonal matrix $\mathbf{bdiag}(\mathcal{X})$ are defined as

$$\mathbf{bcirc}(\mathcal{X}) = \begin{bmatrix} \mathcal{X}^{(1)} & \mathcal{X}^{(n_3)} & \cdots & \mathcal{X}^{(2)} \\ \mathcal{X}^{(2)} & \mathcal{X}^{(1)} & \cdots & \mathcal{X}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}^{(n_3)} & \mathcal{X}^{(n_3-1)} & \cdots & \mathcal{X}^{(1)} \end{bmatrix},$$

$$\mathbf{bdiag}(\mathcal{X}) = \begin{bmatrix} \mathcal{X}^{(1)} & & & \\ & \mathcal{X}^{(2)} & & \\ & & \ddots & \\ & & & \mathcal{X}^{(n_3)} \end{bmatrix}.$$

The block vectorization is defined as $\mathbf{bvec}(\mathcal{X}) = [\mathcal{X}^{(1)}; \cdots; \mathcal{X}^{(n_3)}]$. The inverse operations of **bvec** and **bdiag** are defined as $\mathbf{bvfold}(\mathbf{bvec}(\mathcal{X})) = \mathcal{X}$ and $\mathbf{bdfold}(\mathbf{bdiag}(\mathcal{X})) = \mathcal{X}$, respectively. Let $\mathcal{Y} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$. The **t-product** $\mathcal{X} * \mathcal{Y}$ is an $n_1 \times n_4 \times n_3$ tensor, $\mathcal{X} * \mathcal{Y} = \mathbf{bvfold}(\mathbf{bcirc}(\mathcal{X}) * \mathbf{bvec}(\mathcal{Y}))$. The **transpose** of $\mathcal{X}$ is $\mathcal{X}^T \in \mathbb{R}^{n_2 \times n_1 \times n_3}$ by transposing each of the frontal slices and then reversing the order of transposed frontal slices 2 through n_3. The **identity tensor** $\mathcal{I} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ is a tensor whose first frontal slice is an $n_1 \times n_1$ identity matrix and the rest frontal slices are zero. A tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ is **orthogonal** if it satisfies $\mathcal{X}^T * \mathcal{X} = \mathcal{X} * \mathcal{X}^T = \mathcal{I}$.

Definition 1 (t-SVD): Given $\mathcal{X}$, its t-SVD is defined as

$$\mathcal{X} = \mathcal{U} * \mathcal{G} * \mathcal{V}^T,$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal tensors, $\mathcal{G} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is an f-diagonal tensor. Each of its frontal slices is a diagonal matrix.

Based on the t-SVD, the tensor multirank and its convex surrogate *i.e.*, tensor nuclear norm can be defined as follows:

Definition 2 (Tensor Multirank): The multirank of a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a vector whose i-th element is the rank of the i-th frontal slice of $\hat{\mathcal{X}}$.

Definition 3 (Tensor Nuclear Norm): The tensor nuclear norm of a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, denoted as $\|\mathcal{X}\|_{\circledast}$, is defined as the sum of singular values of all the frontal slices of $\hat{\mathcal{X}}$, i.e.,

$$\|\mathcal{X}\|_{\circledast} = \sum_{i=1}^{\min\{n_1, n_2\}} \sum_{k=1}^{n_3} |\hat{\mathcal{G}}(i, i, k)|. \tag{1}$$

III. GENERALIZED NONCONVEX LOW-RANK TENSOR APPROXIMATION FOR MVSC

In this section, we will propose a Generalized Nonconex Low-rank Tensor Approximation (GNLTA) model for multi-view subspace clustering. The main idea of the proposed GNLTA is that we search a general approximation for the tensor rank instead of the convex tensor nuclear norm. We will provide a general solver to overcome the nonconvexity of the proposed generalized tensor norm and the unified theoretical guarantee.

A. The Proposed GNLTA Model

Inherited from the low-rank tensor representation, Xie *et al.* [1] proposed a unifying multi-view self-representation model:

$$\begin{aligned}&\min_{\mathcal{Z},E}\|\mathcal{Z}\|_{\circledast}+\alpha\sum_{v=1}^{M}\|E^{(v)}\|_{2,1}\\&s.t.\ X^{(v)}=X^{(v)}Z^{(v)}+E^{(v)},\quad v=1,\cdots,M,\\&\quad\mathcal{Z}=\Phi(Z^{(1)},Z^{(2)},\cdots,Z^{(V)}),\end{aligned}\tag{2}$$

where $X^{(v)}\in\mathbb{R}^{d_v\times n}$ denotes the v-th feature, d_v is the dimension of feature, n is the number of data points. $E^{(v)}$ is noise. The last equation aims to construct the representation tensor $\mathcal{Z}$. The low-rank constraint is imposed on $\mathcal{Z}$. This is because the number of clusters is much smaller than the sample number [1], [26]. A non-negligible shortcoming of Eq. (2) is the high computation cost. This is because it needs to perform the matrix inversion on $X^{(v)^T}X^{(v)}+\mu I$, where μ is a parameter and I represents the identity matrix. To solve this, Wu *et al.* [26] developed an essential tensor learning model for multi-view clustering:

$$\begin{aligned}&\min_{\mathcal{Z},\mathcal{E}}\|\mathcal{Z}\|_{\circledast}+\alpha\|\mathcal{E}\|_{2,1}\\&s.t.\ \mathcal{P}=\mathcal{Z}+\mathcal{E},\end{aligned}\tag{3}$$

where $\mathcal{P}\in\mathbb{R}^{n\times n\times M}$ is constructed by M similarity matrices $P^{(v)}$ which is computed by the Markov chain [43]. The details are as follows: the Markov chain first calculates the similarity matrix of the v-th view $A^{(v)}\in\mathbb{R}^{n\times n}$ by $A^{(v)}_{i,j}=\exp(-\frac{\|x_i^{(v)}-x_j^{(v)}\|_2^2}{\sigma^2})$, where σ is the standard deviation. Accordingly, $P^{(v)}$ is constructed by $D^{(v)^{-1}}A^{(v)}$, where $D^{(v)}$ is a diagonal matrix with $D^{(v)}_{ii}=\Sigma_j A^{(v)}_{ij}$. However, both of them used the tensor nuclear norm which may be a biased surrogate of the tensor rank. The reason is that the tensor nuclear norm simply sums all nuclear norms of all frontal slices but different singular values represent different energy, resulting in over-penalizing top singular values. For the matrix case, many works [28], [29] have proven that the convex nuclear norm is a biased matrix rank estimator.

Different from the tensor nuclear norm which simply sums all nuclear norms, we pursue a general surrogate to approximate the original tensor rank, *i.e.*,

$$\|\mathcal{Z}\|_{\theta}=\frac{1}{n}\sum_{j}\sum_{i}\phi\Big(\sigma_i(\hat{\mathcal{Z}}^{(j)}),\theta\Big),\tag{4}$$

where θ is a nonnegative parameter. $\phi(\cdot)$ represents the nonconvex functions which are shown in the following Eqs. (5)-(7). Inspired by the work in [30], we adopt three nonconvex functions to approximate the tensor rank:

$$\|\mathcal{Z}\|_{l,\theta}=\frac{1}{n}\sum_{j}\sum_{i}(1-e^{\frac{-\sigma_i(\hat{\mathcal{Z}}^{(j)})}{\theta}});\tag{5}$$

$$\|\mathcal{Z}\|_{g,\theta}=\frac{1}{n}\sum_{i}\frac{(1+\theta)*\sigma_i(\hat{\mathcal{Z}}^{(j)})}{\theta+\sigma_i(\hat{\mathcal{Z}}^{(j)})};\tag{6}$$

$$\|\mathcal{Z}\|^{\theta}_{\omega,\theta}=\frac{1}{n}\sum_{i}\omega_i*\sigma_i^{\theta}(\hat{\mathcal{Z}}^{(j)});\tag{7}$$

where those three nonconvex functions are named as Laplace, Geman, and weighted Schatten functions, respectively. By replacing the tensor nuclear norm with Eq. (4), our model is formulated as:

$$\begin{aligned}&\min_{\mathcal{Z},\mathcal{E}}\|\mathcal{Z}\|_{\theta}+\alpha\|\mathcal{E}\|_{2,1}\\&s.t.\ \mathcal{P}=\mathcal{Z}+\mathcal{E}.\end{aligned}\tag{8}$$

Due to the nonconvexity of three functions, it is intractable to directly solve the proposed GNLTA model by the off-the-shelf tools. To overcome this, we propose an easy-to-implement algorithm based on the alternation direction method of multipliers (ADMM) and the difference of convex method.

B. Optimization of GNLTA

The main idea of ADMM is to iteratively update one variable by fixing other variables. While the difference of convex method is to overcome the nonconvexity by linearizing the nonconvex functions. We first transform the constrained optimization problem into the following unconstrained one:

$$\begin{aligned}\mathcal{L}_{\mu}(\mathcal{Z},\mathcal{E};\Gamma)=&\|\mathcal{Z}\|_{\theta}+\alpha\|\mathcal{E}\|_{2,1}\\&+<\mathcal{P}-\mathcal{Z}-\mathcal{E},\Gamma>+\frac{\mu}{2}\|\mathcal{P}-\mathcal{Z}-\mathcal{E}\|_F^2,\end{aligned}\tag{9}$$

where $\mu>0$ is the penalty parameter. Symbol Γ denotes the Lagrange multiplier of size $n\times n\times M$ with respect to the linear constraint $\mathcal{P}=\mathcal{Z}+\mathcal{E}$. Eq. (9) is the augmented Lagrangian function of Eq. (8). Then, each variable is updated by fixing the other variables at their latest values. Specifically, we update them as follows:

1) $\mathcal{Z}$-Subproblem: By fixing variables $\mathcal{E}$ and Γ, variable $\mathcal{Z}$ is updated by minimizing Eq. (9):

$$\begin{aligned}\mathcal{Z}_{k+1}&=\arg\min_{\mathcal{Z}}\mathcal{L}_{\mu_k}(\mathcal{Z},\mathcal{E}_k,\Gamma_k),\\&=\arg\min_{\mathcal{Z}}\|\mathcal{Z}\|_{\theta}+\frac{\mu_k}{2}\|\mathcal{Z}-(\mathcal{P}-\mathcal{E}_k+\frac{\Gamma_k}{\mu_k})\|_F^2.\end{aligned}\tag{10}$$

Unfortunately, it is difficult to directly yield the closed-form solution of Eq. (10) due to the nonconvexity of the generalized tensor rank approximation. To solve this problem, we use the difference of convex method to linearize three functions. Inherited from Eq. (4), Eq. (10) can be equivalently transformed into the frequency domain and separated n problems. The j-th problem is

$$\mathcal{Z}^{(j)}_{k+1}=\arg\min_{\hat{\mathcal{Z}}^{(j)}}\frac{1}{\mu_k}\sum_{i=1}^{M}\phi(\sigma_i(\hat{\mathcal{Z}}^{(j)}),\theta)+\frac{1}{2}\|\hat{\mathcal{Z}}^{(j)}-\hat{\mathcal{T}}_k^{(j)}\|_F^2,\tag{11}$$

where $\hat{\mathcal{Z}}=\mathbf{fft}(\mathcal{Z},[],3)$ and $\hat{\mathcal{T}}_k=\mathbf{fft}(\mathcal{P}-\mathcal{E}_k+\frac{\Gamma_k}{\mu_k},[],3)$. $\hat{\mathcal{Z}}^{(j)}$ is the j-th frontal slice of $\hat{\mathcal{Z}}$. Considering the nonascending order of singular values and the antimonotone property of gradient of these three nonconvex functions, we have

$$0\le\nabla\phi(\sigma_1^k,\theta)\le\nabla\phi(\sigma_2^k,\theta)\le\cdots\le\nabla\phi(\sigma_M^k,\theta),\tag{12}$$

$$\begin{aligned}&\phi(\sigma_i(\hat{\mathcal{Z}}^{(j)}),\theta)\\&\quad\le\phi(\sigma_i^k,\theta)+\nabla\phi(\sigma_i^k,\theta)(\sigma_i(\hat{\mathcal{Z}}^{(j)})-\sigma_i^k),\end{aligned}\tag{13}$$

where σ_i^k denotes the i-th singular value of $\hat{\mathcal{Z}}_k^{(j)}$. $\nabla\phi(\sigma_i^k,\theta)$ is the gradient of $\phi(\sigma_i(\hat{\mathcal{Z}}^{(j)}),\theta)$ at σ_i^k. Eq. (13) is derived by the supergradient definition of the concave function. Based on the above three equations and the difference of convex method, Eq. (11) is relaxed into

$$\begin{aligned}\mathcal{Z}_{k+1}^{(j)} &= \arg\min_{\hat{\mathcal{Z}}^{(j)}} \frac{1}{\mu_k}\sum_{i=1}^{M}\phi(\sigma_i^k,\theta)\\ &\quad+\nabla\phi(\sigma_i^k,\theta)(\sigma_i(\hat{\mathcal{Z}}^{(j)})-\sigma_i^k)+\frac{1}{2}\|\hat{\mathcal{Z}}^{(j)}-\hat{T}_k^{(j)}\|_F^2,\\ &= \arg\min_{\hat{\mathcal{Z}}^{(j)}} \frac{1}{\mu_k}\sum_{i=1}^{M}\nabla\phi(\sigma_i^k,\theta)\sigma_i(\hat{\mathcal{Z}}^{(j)})+\frac{1}{2}\|\hat{\mathcal{Z}}^{(j)}-\hat{T}_k^{(j)}\|_F^2.\end{aligned} \tag{14}$$

The optimal solution of Eq. (14) can be yielded by the generalized weighted singular value thresholding (WSVT) [44], [45].

Lemma 1: For any $\frac{1}{\mu_k}>0$ *and the given data* $\hat{T}_k=\mathbf{fft}\{\mathcal{P}-\mathcal{E}_k+\frac{\Gamma_k}{\mu_k},[],3\}$*, the optimal solution* $\mathcal{Z}_{k+1}^{(j)}$ *to the problem (14) is given by the WSVT [44]*

$$\hat{\mathcal{Z}}_{k+1}^{(j)}=U\mathcal{S}_{\frac{\nabla\phi}{\mu_k}}(\Sigma)V^T, \tag{15}$$

where $\hat{T}_k^j=U\Sigma V^T$ *is the SVD of* $\hat{T}_k^j$ *and* $\mathcal{S}_{\frac{\nabla\phi}{\mu_k}}(\Sigma)=\mathbf{diag}\{\max(\Sigma_{ii}-\frac{\nabla\phi(\sigma_i^k)}{\mu_k},0)\}$.

2) $\mathcal{E}$-Subproblem: Similarly, fixing variables $\mathcal{Z}$ and Γ, variable $\mathcal{E}$ is updated by minimizing Eq. (9) as follows:

$$\begin{aligned}\mathcal{E}_{k+1} &= \arg\min_{\mathcal{E}}\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E};\Gamma_k),\\ &= \arg\min_{\mathcal{E}}\alpha\|\mathcal{E}\|_{2,1}+\frac{\mu_k}{2}\|\mathcal{E}-\mathcal{C}_k\|_F^2,\end{aligned} \tag{16}$$

where $\mathcal{C}_k=\mathcal{P}-\mathcal{Z}_{k+1}+\frac{\Gamma_k}{\mu_k}$. After vertically concatenating together along the column of $\mathcal{E}$ and $\mathcal{C}_k$ into matrices $E=[\mathcal{E}^{(1)};\mathcal{E}^{(2)};\cdots;\mathcal{E}^{(M)}]$ and $C_k=[\mathcal{C}_k^{(1)};\mathcal{C}_k^{(2)};\cdots;\mathcal{C}_k^{(M)}]$, Eq. (16) is rewritten as

$$\arg\min_{E}\frac{\alpha}{\mu_k}\|E\|_{2,1}+\frac{1}{2}\|E-C_k\|_F^2, \tag{17}$$

and the j-th column of E_{k+1} is

$$\begin{cases}\dfrac{\|C_k(:,j)\|_2-\frac{\alpha}{\mu_k}}{\|C_k(:,j)\|_2}C_k(:,j), & \text{if } \dfrac{\alpha}{\mu_k}<\|C_k(:,j)\|_2;\\ 0, & \text{otherwise.}\end{cases} \tag{18}$$

3) Γ and μ-Subproblems: Multiplier Γ and penalty parameter μ are updated by

$$\Gamma_{k+1}=\Gamma_k+\mu_k(\mathcal{P}-\mathcal{Z}_{k+1}-\mathcal{E}_{k+1}), \tag{19}$$

$$\mu_{k+1}=\min\{\beta*\mu_k,\mu_{max}\}, \tag{20}$$

where β is set to 2 to accelerate the convergence. μ_{max} denotes the max value of penalty parameter [46].

A summary of the proposed GNLTA algorithm for multi-view clustering is given in Algorithm 1, where the convergence condition is defined as follows:

$$\max\left\{\begin{array}{l}\|\mathcal{Z}_{k+1}-\mathcal{Z}_k\|_\infty,\\ \|\mathcal{E}_{k+1}-\mathcal{E}_k\|_\infty,\\ \|\mathcal{P}-\mathcal{Z}_{k+1}-\mathcal{E}_{k+1}\|_\infty\end{array}\right\}\le tol, \tag{21}$$

Algorithm 1 GNLTA for Multi-View Subspace Clustering

Input: Similarity matrices $\{P^{(v)}\}_{v=1}^M$ by Markov chain, trade-off parameter α, parameter θ;

Initialize: $\mathcal{Z}_1,\ \mathcal{E}_1,\ \Gamma_1$ initialized to $\mathbf{0}$; $\beta=2$, $\mu_1=10^{-3}$, $\mu_{max}=10^8$;

```
Store all similarity matrices {P^(v)}_{v=1}^M into a third-order tensor P;
while not converged do
    for j = 1 to n do
        Update Ẑ_{k+1}^(j) by Eq. (15);
    end for
    Update E_{k+1} by Eq. (18) and transform E_{k+1} into tensor E_{k+1};
    Update Γ_{k+1} and μ_{k+1} by Eqs. (19) and (20);
    Check the convergence condition in Eq. (21);
end while
Construct affinity matrix
```

$$S=\frac{1}{M}\sum_{v=1}^{M}(|\mathcal{Z}^{(v)}|+|\mathcal{Z}^{(v)^T}|);$$

Output: Clustering result by performing the spectral clustering on S.

where $tol>0$ is a pre-defined tolerance. Once obtaining the optimal $\mathcal{Z}$, the final affinity matrix is constructed by averaging all frontal slices of $\mathcal{Z}$.

C. Connection With Several Existing Works

There are several related works including [1], [5], [26], [30]. The method in [30] extended the low-rank matrix approximation from the real domain into the quaternion domain with applications to color image processing. It cannot handle the multi-view clustering. Although the methods in [1], [5], [26] have achieved promising performance on multi-view clustering task, all of them face one common shortcoming: they used the existing tensor nuclear norm as the low-rank regularizers, which may lead to a compromised clustering performance. The reason is that the existing tensor nuclear norm treats each singular value equally and over-penalizes top singular values, resulting in a biased estimation of the rank function [29]. However, our GNLTA proposed a general surrogate to approximate the original tensor rank as shown in Eq. (4) which assigns different weights to different singular values. Thus, larger singular values of GNLTA still dominate more important information of the representation tensor. In addition, different from existing methods, we also provide a general solver with theoretical guarantee. In summary, the proposed GNLTA is significantly different from these related works.

D. Complexity Analysis

The computational complexity of the proposed Algorithm 1 is shown as follows: Variables $\mathcal{Z}$ and $\mathcal{E}$ are both of size $n\times n\times M$, where n and M denote the number of samples and views, respectively. Updating $\mathcal{Z}$ needs to perform the fast Fourier transformation (FFT), inverse FFT and singular value

decomposition with cost of $\mathcal{O}(Mn^2\log(n)+M^2n^2)$. Regarding the complexity of $\mathcal{E}$, it takes $\mathcal{O}(Mn^2)$ cost. Thus, the overall computational complexity of GNLTA is $\mathcal{O}(TMn^2\log(n))$ since $M \leq \log(n)$ and T represents the number of iterations.

E. Convergence Analysis

Before giving the theoretical convergence of GNLTA, three lemmas are firstly introduced.

Lemma 2: The augmented Lagrangian function in Eq. (9) $\mathcal{L}_\mu(\mathcal{Z},\mathcal{E};\Gamma)$ is monotonically decreasing with respect to variable $\mathcal{Z}$.

Proof: Since $\mathcal{Z}_{k+1}^{(j)}$ is the optimal solution of Eq. (14), we have

$$\frac{1}{\mu_k}\sum_{i=1}^{M}\nabla\phi(\sigma_i^k,\theta)\sigma_i^{k+1}+\frac{1}{2}\|\hat{\mathcal{Z}}_{k+1}^{(j)}-\hat{\mathcal{T}}_k^{(j)}\|_F^2, \leq \frac{1}{\mu_k}\sum_{i=1}^{M}\nabla\phi(\sigma_i^k,\theta)\sigma_i^{k}+\frac{1}{2}\|\hat{\mathcal{Z}}_{k}^{(j)}-\hat{\mathcal{T}}_k^{(j)}\|_F^2. \quad (22)$$

After simple operations, Eq. (22) is transformed into

$$\frac{1}{2}\|\hat{\mathcal{Z}}_{k+1}^{(j)}-\hat{\mathcal{T}}_k^{(j)}\|_F^2-\frac{1}{2}\|\hat{\mathcal{Z}}_{k}^{(j)}-\hat{\mathcal{T}}_k^{(j)}\|_F^2 \leq \frac{1}{\rho_k}\sum_{i=1}^{M}\nabla\phi(\sigma_i^k,\theta)(\sigma_i^k-\sigma_i^{k+1}). \quad (23)$$

Summing Eq. (13) from $i=1$ to M:

$$\sum_{i=1}^{M}(\phi(\sigma_i^{k+1},\theta)-\phi(\sigma_i^k,\theta)) \leq \sum_{i=1}^{M}\nabla\phi(\sigma_i^k,\theta)(\sigma_i^{k+1}-\sigma_i^k). \quad (24)$$

By adding Eq. (23) and Eq. (24) from $j=1$ to n, we yield

$$\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E}_k;\Gamma_k) \leq \mathcal{L}_{\mu_k}(\mathcal{Z}_k,\mathcal{E}_k;\Gamma_k). \quad (25)$$

This means that $\mathcal{L}_\mu(\mathcal{Z},\mathcal{E};\Gamma)$ is monotonically decreasing with respect to variable $\mathcal{Z}$. □

Lemma 3: The sequence $\{\Gamma_k\}$ is bounded.

Proof: The optimal $\mathcal{E}_{k+1}$ satisfies the first-order optimality condition, *i.e.*,

$$\begin{aligned}0 &\in \partial_{\mathcal{E}}\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E};\Gamma_k)|_{\mathcal{E}_{k+1}}\\ &= \partial_{\mathcal{E}}(\alpha\|\mathcal{E}\|_{2,1})|_{\mathcal{E}_{k+1}}-\Gamma_k-\mu_k(\mathcal{P}-\mathcal{Z}_{k+1}-\mathcal{E}_{k+1})\\ &= \partial_{\mathcal{E}}(\alpha\|\mathcal{E}\|_{2,1})|_{\mathcal{E}_{k+1}}-\Gamma_{k+1}.\end{aligned} \quad (26)$$

Solving $\mathcal{E}$ is transformed into matrix E. Besides, $\partial_E(\alpha\|E\|_{2,1})|_{E_{k+1}}$ is bounded since

$$\partial_E(\|E\|_{2,1})|_{E_{k+1}}=\begin{cases}0, & if\ \|E_{k+1}(:,l)\|_2=0\\ \dfrac{E_{k+1}(:,l)}{\|E_{k+1}(:,l)\|_2}, & otherwise.\end{cases}$$

Thus, $\partial_{\mathcal{E}}(\alpha\|\mathcal{E}\|_{2,1})|_{\mathcal{E}_{k+1}}$ is bounded. Finally, we obtain that $\{\Gamma_k\}$ is bounded. □

Lemma 4: Sequences $\{\mathcal{Z}_k\}$ and $\{\mathcal{E}_k\}$ are bounded if $\sum_{i=1}^{\infty}(\mu_i+\mu_{i-1})/(\mu_{i-1})^2<\infty$.

Proof: With simple manipulation, we have

$$\begin{aligned}&\mathcal{L}_{\mu_k}(\mathcal{Z}_k,\mathcal{E}_k;\Gamma_k)\\ &= \mathcal{L}_{\mu_{k-1}}(\mathcal{Z}_k,\mathcal{E}_k;\Gamma_{k-1})+\frac{\mu_k-\mu_{k-1}}{2}\|\mathcal{P}-\mathcal{Z}_k-\mathcal{E}_k\|_F^2\\ &\quad+\langle\Gamma_k-\Gamma_{k-1},\mathcal{P}-\mathcal{Z}_k-\mathcal{E}_k\rangle\\ &= \mathcal{L}_{\mu_{k-1}}(\mathcal{Z}_k,\mathcal{E}_k;\Gamma_{k-1})+\frac{\mu_k+\mu_{k-1}}{2(\mu_{k-1})^2}\|\Gamma_k-\Gamma_{k-1}\|_F^2,\end{aligned} \quad (27)$$

and

$$\begin{aligned}&\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E}_{k+1};\Gamma_k)\\ &\leq \mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E}_k;\Gamma_k)\\ &\leq \mathcal{L}_{\mu_k}(\mathcal{Z}_k,\mathcal{E}_k;\Gamma_k)\\ &= \mathcal{L}_{\mu_{k-1}}(\mathcal{Z}_k,\mathcal{E}_k;\Gamma_{k-1})+\frac{\mu_k+\mu_{k-1}}{2(\mu_{k-1})^2}\|\Gamma_k-\Gamma_{k-1}\|_F^2.\end{aligned} \quad (28)$$

The first two inequations are obtained since $\mathcal{E}_{k+1}$ is a global solution of Eq. (16) and Lemma 2. The last equation is yielded by Eq. (27). Based on Eq. (28), we obtain

$$\begin{aligned}&\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E}_{k+1};\Gamma_k)\\ &\leq \mathcal{L}_{\mu_0}(\mathcal{Z}_1,\mathcal{E}_1;\Gamma_0)+\sum_{i=1}^{k}\frac{\mu_i+\mu_{i-1}}{2(\mu_{i-1})^2}\|\Gamma_i-\Gamma_{i-1}\|_F^2\\ &\leq \mathcal{L}_{\mu_0}(\mathcal{Z}_1,\mathcal{E}_1;\Gamma_0)+C\sum_{i=1}^{k}\frac{\mu_i+\mu_{i-1}}{2(\mu_{i-1})^2},\end{aligned} \quad (29)$$

where C is the upper bound of $\{\Gamma_i-\Gamma_{i-1}\}$. Since $\sum_{i=1}^{\infty}(\mu_i+\mu_{i-1})/(\mu_{i-1})^2<\infty$, all of Eq. (29) are bounded. Thus, $\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E}_{k+1};\Gamma_k)$ is upper bounded. By adding $1/(2\mu_k)\|\Gamma_k\|_F^2$ into Eq. (9), we have

$$\begin{aligned}&\mathcal{L}_{\mu_k}(\mathcal{Z}_{k+1},\mathcal{E}_{k+1};\Gamma_k)+\frac{1}{2\mu_k}\|\Gamma_k\|_F^2\\ &=\|\mathcal{Z}_{k+1}\|_\theta+\alpha\|\mathcal{E}_{k+1}\|_{2,1}+\frac{\mu_k}{2}\|\mathcal{P}-\mathcal{Z}_{k+1}-\mathcal{E}_{k+1}+\frac{\Gamma_k}{\mu_k}\|_F^2.\end{aligned}$$

According to Lemma 3, these two terms of left side is bounded. We reach that each term of right side is also bounded, *i.e.*, $\{\mathcal{Z}_k\}$ and $\{\mathcal{E}_k\}$ are bounded. □

Theorem 1: Suppose $\{\mathcal{Z}_k,\mathcal{E}_k,\Gamma_k\}$ be the sequence by Algorithm 1 and $\{\mathcal{Z}^*,\mathcal{E}^*,\Gamma^*\}$ be an accumulation point. Then, $\{\mathcal{Z}^*,\mathcal{E}^*\}$ is stationary point of Eq. (8) as long as $\sum_{i=1}^{\infty}(\mu_i+\mu_{i-1})/(\mu_{i-1})^2<\infty$ and $\lim_{k\to\infty}(\mathcal{E}_k-\mathcal{E}_{k+1})/\mu_k=0$.

Proof: According to Lemmas 3 and 4, we yield that the sequence $\{\mathcal{Z}_k,\mathcal{E}_k,\Gamma_k\}$ is bounded. By Bolzano–Weierstrass theorem, the sequence has at least one accumulation point which is denoted as $\{\mathcal{Z}^*,\mathcal{E}^*,\Gamma^*\}$. Without loss of generality, we assume that $\{\mathcal{Z}_k,\mathcal{E}_k,\Gamma_k\}$ converges to $\{\mathcal{Z}^*,\mathcal{E}^*,\Gamma^*\}$. Then, we need only to prove that this accumulation point is a stationary point of Eq. (8).

According to Eq. (19) and the boundness of $\{\Gamma_k\}$, we get $\lim_{k\to\infty}\mathcal{P}-\mathcal{Z}_{k+1}-\mathcal{E}_{k+1}=\lim_{k\to\infty}(\Gamma_{k+1}-\Gamma_k)/\mu_k=0$; then

$$\mathcal{P}=\mathcal{Z}^*+\mathcal{E}^*. \quad (30)$$

TABLE II
INFORMATION OF ALL REAL MULTI-VIEW DATABASES

Database	Ins./Num.	View1	View2	View3	View4	View5
BBC4view	685/5	4659d	4633d	4655d	4684d	-
BBCSport	544/5	3183d	3203d	-	-	-
Flowers	1360/17	1360d	1360d	1360d	-	-
UCI	2000/10	240d	76d	6d	-	-
StillDB	467/6	200d	200d	200d	-	-
MITIndoor	5360/67	4096d	3600d	1770d	1240d	-
Reuters	600/6	21526d	24892d	34121d	15487d	11539d

Fig. 1. Samples (from top to bottom rows) of Flowers, UCI, StillDB, and MITIndoor datasets.

According to the first-order optimality condition, we have

$$\begin{aligned}&\partial_{\mathcal{Z}}\mathcal{L}_{\mu_k}(\mathcal{Z},\mathcal{E}_k;\Gamma_k)|_{\mathcal{Z}_{k+1}}\\&\quad=\partial_{\mathcal{Z}}(\|\mathcal{Z}\|_\theta)|_{\mathcal{Z}_{k+1}}-\Gamma_k-\mu_k(\mathcal{P}-\mathcal{Z}_{k+1}-\mathcal{E}_k)\\&\quad=\partial_{\mathcal{Z}}(\|\mathcal{Z}\|_\theta)|_{\mathcal{Z}_{k+1}}-\Gamma_{k+1}+\mu_k(\mathcal{E}_k-\mathcal{E}_{k+1})\\&\quad=0.\end{aligned}$$

Under the assumption of $\lim_{k\to\infty}(\mathcal{E}_k-\mathcal{E}_{k+1})/\mu_k=0$, we have

$$\partial_{\mathcal{Z}}(\|\mathcal{Z}^*\|_\theta)-\Gamma^*=0. \tag{31}$$

From Eq. (26), we have

$$\partial_{\mathcal{E}}(\alpha\|\mathcal{E}^*\|_{2,1})-\Gamma^*=0. \tag{32}$$

According to Eqs. (30)-(32), we reach the conclusion that $\{\mathcal{Z}^*,\mathcal{E}^*,\Gamma^*\}$ satisfies the Karush—Kuhn–Tucker conditions of $\mathcal{L}_\mu(\mathcal{Z},\mathcal{E};\Gamma)$. Thus, $\{\mathcal{Z}^*,\mathcal{E}^*\}$ is a stationary point of the original problem (8). □

IV. EXPERIMENTAL RESULTS

In this section, we conducted several experiments on seven real-world and challenging datasets to investigate the performance of the proposed GNLTA.

A. Experimental Setting

We first introduce seven multi-view datasets including text and image data which are summarized in Table II. Some examples of our testing datasets are shown in Fig. 1. The details of these datasets are described as follows: **BBC4view** and **BBCSport**[1] are the News article datsets. The number of samples are 685 and 544, respectively. BBC4view extracted four types features while BBCSport has two views. **Flowers**[2] contains 1360 flower samples with 17 categories. Three different visual features, *i.e.*, 1360d color, 1360d texture, and 1360d shape are selected as three views. **UCI-digits**[3] is a handwritten digit database from the UC Irvine machine learning repository. There are 2000 digit images belonging to 10 classes. Three types of features including 240d Fourier coefficients, 76d pixel averages and 6d morphological features are extracted. **StillDB** [56] is a still image dataset including 467 still images with 6 classes of actions *i.e.*, running, walking, catching, throwing, crouching and kicking. 200d Sift Bow, 200d Color Sift Bow, and 200d Shape Context Bow are explored in StillDB dataset. **MITIndoor** [57] is a scene dataset which contains 5360 images with 67 categories. In this dataset, we extract four types of features including 4096d PHOW, 3600d LBP, 1770d CENTRIST, and 1240d deep features. **Reuters** is a newswire articles written in 5 languages including English, French, German, Italian and Spanish, leading to 5 views. Same to [58], we randomly sampled 100 documents from each class, resulting in a dataset of 600 documents.

The following three single-view clustering and seventeen multi-view clustering methods are compared with the proposed GNLTA: **SSC** [47]: SSC handles multi-view features independently and pursues a spare representation matrix; **LRR** [48]: LRR handles multi-view features independently and pursues a low-rank representation matrix; **S$_0$/L_0-LRSSC** [31]: it is the l_0-motivated low-rank sparse subspace clustering; **RMSC** [12]: RMSC uses the low-rank and sparse decomposition to recover a shared affinity matrix; **DiMSC** [13]: DiMSC uses the Hilbert Schmidt Independence Criterion to explore the complementary information; **LT-MSC** [5]: LT-MSC explores the convex tensor nuclear norm to pursue a low-rank representation tensor; **ECMSC** [49]: ECMSC exploits the data representation exclusivity and indicator consistency simultaneously; **MLAN** [50]: MLAN performs the clustering with adaptive neighbors; **GMC** [11]: GMC fuses all graph matrices to generate a unified graph matrix; **AWP** [51]: AWP ultilizes the adaptively weighted procrustes for multi-view clustering; **GLSR** [18]: GLSR integrates the least square regression with the graph learning for multi-view clustering; **LMSC** [14]: LMSC clusters data points based on the latent representation instead of the original features. **MCLES** [52]: Similar to [14], MCLES also performs the clustering in a learned latent embedding space; **tSVDMSC** [1]: tSVDMSC utilizes the t-SVD-based tensor nuclear norm to learn a low-rank representation tensor; **SM^{2}SC** [53]: SM^{2}SC exploits the joint strength of the multiplicative decomposition scheme and the variable splitting scheme for MVSC; **ETLMSC** [26]: ETLMSC extended the RMSC method from the matrix optimization into the tensor one; **LMVSC** [54]: LMVSC used the anchor graph for MVSC with linear order complexity; **DMF-MVC** [17]: DMF-MVC exploited the deep matrix factorization for MVC;

[1] http://mlg.ucd.ie/datasets/segment.html
[2] http://www.robots.ox.ac.uk/vgg/data/flowers/
[3] http://archive.ics.uci.edu/ml/datasets/Multiple+Features

TABLE III
COMPARISON RESULTS ON BBC4VIEW AND BBCSPORT DATABASES

Dataset	Method	ACC	NMI	AR	F-score	Precision	Recall
BBC4view	SSC[47]	0.660±0.002	0.494±0.005	0.470±0.001	0.599±0.001	0.578±0.001	0.622±0.001
	LRR[48]	0.802±0.000	0.568±0.000	0.621±0.000	0.712±0.000	0.697±0.000	0.727±0.000
	S_0/L_0-LRSSC[31]	0.858±0.001	0.648±0.001	0.680±0.001	0.755±0.001	0.763±0.001	0.747±0.001
	RMSC[12]	0.775±0.003	0.616±0.004	0.560±0.002	0.656±0.002	0.703±0.003	0.616±0.001
	DiMSC[13]	0.892±0.001	0.728±0.002	0.752±0.002	0.810±0.002	0.811±0.002	0.810±0.002
	LT-MSC[5]	0.591±0.000	0.442±0.005	0.400±0.001	0.546±0.000	0.525±0.000	0.570±0.001
	ECMSC[49]	0.308±0.028	0.047±0.009	0.008±0.018	0.322±0.017	0.239±0.009	0.497±0.064
	MLAN[50]	0.853±0.007	0.698±0.010	0.716±0.005	0.783±0.004	0.776±0.003	0.790±0.004
	GMC[11]	0.693±0.000	0.563±0.000	0.479±0.000	0.633±0.000	0.501±0.000	0.860±0.000
	AWP[51]	0.904±0.000	0.761±0.000	0.797±0.000	0.845±0.000	0.838±0.000	0.851±0.000
	GLSR[18]	0.917±0.000	0.780±0.000	0.811±0.000	0.856±0.000	0.835±0.000	0.878±0.000
	LMSC[14]	0.883±0.000	0.699±0.000	0.746±0.000	0.806±0.000	0.797±0.000	0.816±0.000
	MCLES[52]	0.819±0.000	0.637±0.000	0.662±0.000	0.742±0.000	0.735±0.000	0.750±0.000
	tSVDMSC[1]	0.858±0.001	0.685±0.002	0.725±0.002	0.789±0.001	0.800±0.001	0.778±0.002
	SM^2SC[53]	0.934±0.008	0.812±0.001	0.853±0.003	0.887±0.006	0.892±0.010	0.883±0.022
	ETLMSC[26]	0.872±0.094	0.826±0.028	0.811±0.082	0.855±0.063	0.866±0.073	0.846±0.057
	LMVSC[54]	0.480±0.000	0.242±0.000	0.103±0.000	0.380±0.000	0.289±0.000	0.554±0.000
	DMF-MVC[17]	0.321±0.001	0.075±0.001	0.037±0.001	0.264±0.001	0.262±0.001	0.266±0.001
	MvDSCN[55]	0.495±0.019	0.247±0.022	0.224±0.023	0.437±0.015	0.407±0.019	0.495±0.011
	RMSL[39]	0.943±0.009	0.831±0.015	0.862±0.004	0.894±0.002	0.901±0.005	0.888±0.004
	GNLTA-Sch	0.974±0.006	0.937±0.005	0.941±0.007	0.954±0.006	0.974±0.003	0.936±0.008
	GNLTA-Lap	0.967±0.066	0.952±0.043	0.947±0.084	0.960±0.062	0.961±0.085	0.959±0.035
	GNLTA-Gem	0.973±0.003	0.931±0.005	0.937±0.007	0.952±0.005	0.971±0.003	0.933±0.008
BBCSport	SSC[47]	0.627±0.003	0.534±0.008	0.364±0.007	0.565±0.005	0.427±0.004	0.834±0.004
	LRR[48]	0.836±0.001	0.698±0.002	0.705±0.001	0.776±0.001	0.768±0.001	0.784±0.001
	S_0/L_0-LRSSC[31]	0.902±0.004	0.780±0.001	0.751±0.004	0.809±0.003	0.820±0.002	0.799±0.009
	RMSC[12]	0.826±0.001	0.666±0.001	0.637±0.001	0.719±0.001	0.766±0.001	0.677±0.001
	DiMSC[13]	0.922±0.000	0.785±0.000	0.813±0.000	0.858±0.000	0.846±0.000	0.872±0.000
	LT-MSC[5]	0.460±0.046	0.222±0.028	0.167±0.043	0.428±0.014	0.328±0.028	0.629±0.053
	ECMSC[49]	0.285±0.014	0.027±0.013	0.009±0.011	0.267±0.020	0.244±0.007	0.297±0.045
	MLAN[50]	0.721±0.000	0.779±0.000	0.591±0.000	0.714±0.000	0.567±0.000	0.962±0.000
	GMC[11]	0.807±0.000	0.760±0.000	0.722±0.000	0.794±0.000	0.727±0.000	0.875±0.000
	AWP[51]	0.809±0.000	0.723±0.000	0.726±0.000	0.796±0.000	0.743±0.000	0.857±0.000
	GLSR[18]	0.873±0.000	0.781±0.000	0.803±0.000	0.851±0.000	0.837±0.000	0.865±0.000
	LMSC[14]	0.847±0.003	0.739±0.001	0.749±0.001	0.810±0.001	0.799±0.001	0.822±0.001
	MCLES[52]	0.921±0.000	0.802±0.000	0.795±0.000	0.845±0.000	0.827±0.000	0.865±0.000
	tSVDMSC[1]	0.879±0.000	0.765±0.000	0.784±0.000	0.834±0.000	0.863±0.000	0.807±0.000
	SM^2SC[53]	0.982±0.000	0.937±0.000	0.952±0.000	0.963±0.000	0.970±0.000	0.957±0.000
	ETLMSC[26]	0.959±0.086	0.972±0.058	0.949±0.107	0.961±0.081	0.963±0.078	0.960±0.085
	LMVSC[54]	0.517±0.000	0.382±0.000	0.151±0.000	0.394±0.000	0.329±0.000	0.491±0.000
	DMF-MVC[17]	0.320±0.002	0.061±0.001	0.035±0.001	0.258±0.001	0.266±0.001	0.250±0.002
	MvDSCN[55]	0.931±0.001	0.835±0.000	0.909±0.001	0.860±0.000	0.915±0.001	0.912±0.000
	RMSL[39]	0.972±0.002	0.905±0.005	0.931±0.002	0.947±0.004	0.958±0.00	0.937±0.003
	GNLTA-Sch	0.979±0.065	0.986±0.045	0.972±0.087	0.979±0.066	0.979±0.067	0.980±0.065
	GNLTA-Lap	0.981±0.061	0.989±0.034	0.986±0.044	0.989±0.033	0.989±0.034	0.989±0.033
	GNLTA-Gem	0.980±0.064	0.986±0.043	0.973±0.086	0.979±0.065	0.979±0.067	0.980±0.064

MvDSCN [55]: MvDSCN is the convolution neural network-based deep MVSC method; **RMSL** [39]: RMSL explored the reciprocal multi-layer subspace learning for MVSC.

One parameter of SSC and LRR was selected from $[0.01, 10]$; in S_0/L_0-LRSSC, λ and μ were optimized from $[0.1, 0.9]$ and $\{1, 3, 5, 10, 20\}$, respectively; in RMSC, λ selected from $[0.005, 1]$; in DiMSC, two parameters were chosen from $[0.01, 0.03]$ and $[20 : 20 : 180]$, respectively; λ of LT-MSC was tested from 0.01 to 100; three parameters of ECMSC were selected from $[0.1, 1]$, $[0.1, 1]$ and 1.2; in MLAN, a random parameter was located between 1 and 30; GMC set the trade-off parameter as 1; AWP is a parameter-free method; α, β, λ in GLSR were empirically selected from $[0.001, 0.5]$, $[0.01, 10]$ and $[0.01, 10]$, respectively; The dimensions of the latent representation and λ of LMSC were searched from the range $[10 : 10 : 100]$ and $[0.01, 0.1, 1, 10, 100]$, respectively; Four parameters in MCLES were varied in the range $[10, 100]$, $[0.2, 2]$, $[0.2, 2]$, and $[0.001, 0.01]$, respectively; λ in tSVDMSC and ETLMSC were set within the range $[0.1, 2]$ and $[0.008, 0.1]$, respectively; SM^2SC tunes three parameters in the range $[0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 1, 10, 40, 100]$, $[0.1, 0.5, 1, 1.5, 2]$ and $[0.05, 0.1, 0.4, 1, 5]$; LMVSC searched anchor number in the range $[k, 50, 100]$ and one parameter in $[0.001, 0.01, 0.1, 1, 10]$, where k denotes the number of clusters; DMF-MVC adopted $\{[100, 50], [500, 50], [500, 200]\}$ as the sizes of the last layer and other parameters used the default settings as recommended in [17]; MvDSCN fixed four parameters as $10, 1, 0.1, 0.1$; RMSL tuned α and β from $[0.1 : 0.1 : 1]$, λ and the dimension from $[0.001, 1]$ and $[20 : 20 : 200]$, respectively.

In summary, SSC, LRR and S_0/L_0-LRSSC belong to single-view clustering. All of them perform the subspace

TABLE IV
COMPARISON RESULTS ON FLOWERS AND UCI-3VIEWS DATABASES

Dataset	Method	ACC	NMI	AR	F-score	Precision	Recall
Flowers	SSC[47]	0.398±0.004	0.439±0.003	0.240±0.004	0.285±0.004	0.280±0.004	0.290±0.004
	LRR[48]	0.396±0.009	0.419±0.006	0.239±0.003	0.284±0.003	0.279±0.004	0.290±0.003
	S_0/L_0-LRSSC[31]	0.392±0.014	0.413±0.006	0.242±0.008	0.289±0.007	0.269±0.010	0.313±0.005
	RMSC[12]	0.385±0.016	0.396±0.014	0.231±0.019	0.249±0.011	0.234±0.012	0.256±0.010
	DiMSC[13]	0.434±0.014	0.442±0.011	0.266±0.009	0.310±0.008	0.302±0.007	0.318±0.010
	LT-MSC[5]	0.476±0.012	0.478±0.008	0.313±0.009	0.354±0.008	0.347±0.009	0.361±0.008
	ECMSC[49]	0.446±0.002	0.423±0.001	0.260±0.001	0.304±0.001	0.298±0.001	0.311±0.001
	MLAN[50]	0.501±0.008	0.532±0.003	0.331±0.010	0.373±0.009	0.345±0.010	0.404±0.006
	GMC[11]	0.177±0.000	0.247±0.000	0.020±0.000	0.125±0.000	0.068±0.000	0.822±0.000
	AWP[51]	0.435±0.000	0.430±0.000	0.246±0.000	0.292±0.000	0.277±0.000	0.309±0.000
	GLSR[18]	0.473±0.006	0.478±0.004	0.316±0.007	0.358±0.006	0.341±0.006	0.376±0.006
	LMSC[14]	0.442±0.009	0.444±0.009	0.275±0.007	0.318±0.012	0.312±0.011	0.325±0.011
	MCLES[52]	0.469±0.000	0.516±0.000	0.337±0.000	0.390±0.000	0.342±0.000	0.462±0.000
	tSVDMSC[1]	0.836±0.005	0.852±0.002	0.766±0.002	0.780±0.002	0.772±0.002	0.789±0.002
	SM^2SC[53]	0.442±0.008	0.453±0.005	0.276±0.007	0.319±0.007	0.313±0.007	0.325±0.006
	ETLMSC[26]	0.811±0.066	0.874±0.025	0.763±0.057	0.778±0.054	0.748±0.064	0.810±0.041
	LMVSC[54]	0.360±0.000	0.385±0.000	0.198±0.000	0.246±0.000	0.241±0.000	0.250±0.000
	DMF-MVC[17]	0.240±0.011	0.223±0.006	0.080±0.005	0.139±0.004	0.126±0.004	0.155±0.005
	MvDSCN[55]	0.276±0.012	0.285±0.012	0.108±0.010	0.182±0.008	0.151±0.010	0.300±0.002
	RMSL[39]	0.511±0.006	0.490±0.007	0.332±0.010	0.372±0.005	0.361±0.008	0.384±0.011
	GNLTA-Sch	0.823±0.055	0.875±0.020	0.773±0.046	0.787±0.043	0.760±0.055	0.816±0.032
	GNLTA-Lap	0.845±0.058	0.895±0.029	0.808±0.058	0.819±0.055	0.791±0.062	0.850±0.049
	GNLTA-Gem	0.847±0.038	0.897±0.019	0.810±0.040	0.822±0.037	0.790±0.044	0.857±0.030
UCI-3views	SSC[47]	0.815±0.011	0.840±0.001	0.770±0.005	0.794±0.004	0.747±0.010	0.848±0.004
	LRR[48]	0.871±0.001	0.768±0.002	0.736±0.002	0.763±0.002	0.759±0.002	0.767±0.002
	S_0/L_0-LRSSC[31]	0.927±0.003	0.861±0.001	0.847±0.001	0.862±0.001	0.860±0.001	0.864±0.001
	RMSC[12]	0.915±0.024	0.822±0.008	0.789±0.014	0.811±0.012	0.797±0.017	0.826±0.006
	DiMSC[13]	0.703±0.010	0.772±0.006	0.652±0.006	0.695±0.006	0.673±0.005	0.718±0.007
	LT-MSC[5]	0.803±0.001	0.775±0.001	0.725±0.001	0.753±0.001	0.739±0.001	0.767±0.001
	ECMSC[49]	0.718±0.001	0.780±0.001	0.672±0.001	0.707±0.001	0.660±0.001	0.760±0.001
	MLAN[50]	0.874±0.000	0.910±0.000	0.847±0.000	0.864±0.000	0.797±0.000	0.943±0.000
	GMC[11]	0.736±0.000	0.815±0.000	0.678±0.000	0.713±0.000	0.644±0.000	0.799±0.000
	AWP[51]	0.806±0.000	0.842±0.000	0.759±0.000	0.785±0.000	0.734±0.000	0.842±0.000
	GLSR[18]	0.924±0.001	0.859±0.001	0.841±0.001	0.857±0.001	0.854±0.001	0.860±0.001
	LMSC[14]	0.893±0.000	0.815±0.000	0.783±0.000	0.805±0.000	0.798±0.000	0.812±0.000
	MCLES[52]	0.941±0.004	0.891±0.008	0.877±0.009	0.889±0.008	0.885±0.008	0.894±0.007
	tSVDMSC[1]	0.830±0.000	0.884±0.005	0.786±0.003	0.800±0.004	0.785±0.007	0.808±0.001
	SM^2SC[53]	0.961±0.001	0.914±0.001	0.914±0.001	0.923±0.001	0.922±0.001	0.924±0.001
	ETLMSC[26]	0.958±0.078	0.977±0.028	0.953±0.069	0.958±0.062	0.940±0.088	0.980±0.029
	LMVSC[54]	0.790±0.000	0.756±0.000	0.643±0.000	0.681±0.000	0.643±0.000	0.722±0.000
	MvDSCN[55]	0.308±0.011	0.299±0.013	0.158±0.009	0.281±0.006	0.204±0.008	0.540±0.008
	RMSL[39]	0.578±0.013	0.511±0.014	0.407±0.017	0.474±0.007	0.423±0.008	0.538±0.011
	GNLTA-Sch	0.968±0.089	0.981±0.030	0.966±0.077	0.970±0.068	0.961±0.097	0.982±0.031
	GNLTA-Lap	0.981±0.036	0.979±0.012	0.972±0.036	0.975±0.032	0.968±0.052	0.983±0.008
	GNLTA-Gem	0.967±0.057	0.977±0.019	0.961±0.052	0.965±0.047	0.954±0.070	0.978±0.019

clustering on each view with the best performance reported. The others belong to multi-view clustering. GMC, AWP and MLAN learn the affinity matrix by using the original features, while the others by using the learnt representation matrix or tensor. LRR, RMSC and LMSC utilize the convex matrix nuclear norm to pursue a low-rank representation matrix while LT-MSC, tSVDMSC and ETLMSC use the convex tensor nuclear norm to learn a low-rank representation tensor. DMF-MVC, MSCN and RMSL are recently proposed deep leanring-based MVC methods. In the proposed GNLTA, we develop a novel nonconvex low-rank tensor approximation to learn a better representation tensor.

Same to References [1], [5], [59], we also select six widely used metrics to investigate the clustering performance of all methods. These six metrics include accuracy (ACC), normalized mutual information (NMI), adjusted rank index (AR), F-score, Precision, and Recall. All of them are located in the interval $[0, 1]$ and the higher the values of them are, the better performance is. For a fair comparison, we perform all methods ten times and report the average values with standard deviations. Following [12], [26], we also adapted the Markov chain [43] to compute the initialized similarity matrix $P^{(v)}$ of the v-th view. All view features are normalized into the same level to guarantee the experimental fairness. For simplicity, we set the weights of Eq. (7) to 1.

B. Experimental Results

The clustering results of all methods are reported in Tables III-V in which the best results are marked in red while the second-best ones are marked in blue. We denote GNLTA using the weighted Schatten function, Laplace function, and Geman function as GNLTA-Sch, GNLTA-Lap, and GNLTA-Gem, respectively. Generally speaking, the proposed GNLTA has achieved the top three performance in most cases. For example, on BBCview, UCI-3views, and Flowers, all of GNLTA-Sch, GNLTA-Lap, and GNLTA-Gem have the better performance than all competing methods. On BBC4view,

TABLE V
COMPARISON RESULTS ON STILLDB AND MITINDOOR DATABASES

Dataset	Method	ACC	NMI	AR	F-score	Precision	Recall
StillDB	SSC[47]	0.328±0.002	0.138±0.003	0.087±0.001	0.295±0.001	0.222±0.001	0.440±0.007
	LRR[48]	0.306±0.004	0.109±0.003	0.061±0.004	0.219±0.000	0.223±0.003	0.212±0.004
	S_0/L_0-LRSSC[31]	0.339±0.002	0.138±0.001	0.091±0.001	0.264±0.001	0.237±0.001	0.298±0.001
	RMSC[12]	0.305±0.010	0.089±0.009	0.073±0.011	0.221±0.002	0.231±0.004	0.219±0.002
	DiMSC[13]	0.323±0.002	0.122±0.008	0.083±0.001	0.249±0.000	0.235±0.004	0.256±0.002
	LT-MSC[5]	0.342±0.002	0.136±0.002	0.090±0.001	0.252±0.002	0.243±0.001	0.261±0.003
	ECMSC[49]	0.320±0.008	0.111±0.007	0.090±0.001	0.264±0.010	0.237±0.006	0.300±0.031
	MLAN[50]	0.349±0.000	0.138±0.000	0.098±0.000	0.272±0.000	0.242±0.000	0.310±0.000
	GMC[11]	0.251±0.000	0.078±0.000	0.005±0.000	0.278±0.000	0.174±0.000	0.701±0.000
	AWP[51]	0.306±0.000	0.093±0.000	0.058±0.000	0.223±0.000	0.219±0.000	0.224±0.000
	GLSR[18]	0.368±0.003	0.142±0.002	0.105±0.003	0.279±0.003	0.247±0.002	0.320±0.005
	LMSC[14]	0.327±0.003	0.136±0.003	0.084±0.011	0.269±0.005	0.235±0.007	0.247±0.012
	MCLES[52]	0.338±0.000	0.153±0.000	0.098±0.000	0.264±0.000	0.246±0.000	0.286±0.000
	tSVDMSC[1]	0.347±0.010	0.130±0.004	0.088±0.003	0.255±0.004	0.239±0.002	0.273±0.006
	SM^2SC[53]	0.452±0.002	0.336±0.002	0.312±0.002	0.370±0.002	0.381±0.003	0.378±0.002
	ETLMSC[26]	0.604±0.043	0.520±0.015	0.423±0.029	0.523±0.024	0.518±0.022	0.528±0.027
	LMVSC[54]	0.317±0.000	0.189±0.000	0.066±0.000	0.239±0.000	0.221±0.000	0.259±0.000
	DMF-MVC[17]	0.264±0.001	0.055±0.000	0.026±0.001	0.198±0.000	0.192±0.001	0.203±0.001
	MvDSCN[55]	0.377±0.023	0.245±0.020	0.169±0.003	0.320±0.015	0.272±0.005	0.414±0.010
	RMSL[39]	0.356±0.003	0.131±0.001	0.090±0.002	0.243±0.001	0.247±0.001	0.240±0.002
	GNLTA-Sch	0.613±0.039	0.522±0.025	0.427±0.034	0.530±0.029	0.522±0.026	0.530±0.032
	GNLTA-Lap	0.630±0.030	0.528±0.019	0.444±0.022	0.541±0.018	0.530±0.018	0.552±0.019
	GNLTA-Gem	0.622±0.017	0.518±0.020	0.432±0.019	0.530±0.016	0.524±0.013	0.537±0.020
MITIndoor	SSC[47]	0.475±0.008	0.615±0.003	0.332±0.006	0.343±0.006	0.314±0.007	0.377±0.007
	LRR[48]	0.120±0.004	0.226±0.006	0.031±0.007	0.045±0.004	0.044±0.006	0.047±0.004
	S_0/L_0-LRSSC[31]	0.528±0.006	0.639±0.002	0.370±0.006	0.379±0.005	0.367±0.007	0.393±0.005
	RMSC[12]	0.232±0.009	0.342±0.004	0.110±0.003	0.123±0.002	0.121±0.003	0.125±0.003
	LT-MSC[5]	0.431±0.002	0.546±0.004	0.280±0.008	0.290±0.002	0.279±0.006	0.306±0.005
	ECMSC[49]	0.353±0.002	0.489±0.001	0.216±0.002	0.228±0.001	0.213±0.001	0.247±0.002
	MLAN[50]	0.232±0.010	0.408±0.012	0.012±0.009	0.041±0.003	0.021±0.001	0.662±0.015
	GMC[11]	0.099±0.000	0.204±0.000	0.003±0.000	0.032±0.000	0.016±0.000	0.838±0.000
	AWP[51]	0.499±0.000	0.629±0.000	0.317±0.000	0.329±0.000	0.271±0.000	0.419±0.000
	LMSC[14]	0.384±0.006	0.506±0.005	0.243±0.005	0.254±0.004	0.245±0.005	0.264±0.004
	tSVDMSC[1]	0.684±0.005	0.750±0.007	0.555±0.005	0.562±0.008	0.543±0.005	0.582±0.004
	SM^2SC[53]	0.477±0.009	0.583±0.006	0.321±0.006	0.332±0.006	0.319±0.006	0.346±0.006
	ETLMSC[26]	0.775±0.040	0.899±0.011	0.729±0.037	0.733±0.036	0.709±0.046	0.758±0.025
	LMVSC[54]	0.371±0.000	0.522±0.000	0.112±0.000	0.132±0.000	0.087±0.000	0.278±0.000
	DMF-MVC[17]	0.101±0.002	0.201±0.001	0.020±0.001	0.035±0.001	0.034±0.001	0.038±0.001
	MvDSCN[55]	0.084±0.003	0.182±0.004	0.014±0.002	0.037±0.001	0.026±0.002	0.134±0.005
	RMSL[39]	0.279±0.004	0.372±0.003	0.125±0.005	0.139±0.002	0.126±0.004	0.155±0.002
	GNLTA-Sch	0.806±0.036	0.915±0.012	0.762±0.035	0.766±0.035	0.746±0.038	0.788±0.033
	GNLTA-Lap	0.791±0.047	0.907±0.016	0.740±0.050	0.744±0.050	0.727±0.050	0.762±0.049
	GNLTA-Gem	0.805±0.030	0.912±0.010	0.758±0.031	0.762±0.031	0.741±0.031	0.784±0.031

DiMSC, GLSR and MCLES run out of memory on MITIndoor.

the improvement of GNLTA is around 3.3%, 10.5%, 8.4%, 6.5%, 6.9% and 5.0% over SM^2SC in terms of ACC, NMI, AR, F-score, Precision, and Recall, respectively. This validates the effectiveness and superiority of the generalized nonconvex low-rank tensor approximation for MVSC. The proposed GLNTA has significantly improved the clustering results over the convex nuclear norm-based methods, such as RMSC, LMSC, LT-MSC, tSVDMSC, and ETLMSC. For example, the proposed GNLTA-Sch improves six metrics on MITIndoor over ETLMSC by 3.1%, 1.6%, 3.3%, 3.3%, 3.7%, and 3.0%, respectively. This is because our proposed noncovex low-rank tensor approximation has less estimation bias and better generalization for the pursue of low-rank property over the existing convex low-rank matrix and tensor approximation.

We can observe that the low-rank tensor approximation-based approaches (the proposed GNLTA, tSVDMSC, ETLMSC) outperform the low-rank matrix approximation-based ones (RMSC, ECMSC, LMSC), indicating that the intrinsic low-rank property of the representation tensor is superior to that of the representation matrix. These tensor-based clustering methods not only explore the spatial correlation of multi-view data points, but also the view correlation of them. Most of multi-view clustering methods consistently achieve better performance over those two single-view clustering methods, LRR and SSC on most databases. Particularly, except for RMSC and GMC, other multi-view clustering methods outperform single-view clustering ones LRR and SSC on flowers database. This directly demonstrates that those multi-view complementary and compatible information are beneficial for the performance improvement.

Obviously, deep MVC methods MvDSCN and RMSL have yielded the competitive performance on BBCSport and BBC4view compared with our GNLTA, while they were worse than that of our GNLTA on other databases. This happens for the following reasons: (1) As stated in [41], the performance

TABLE VI
COMPARISON RESULTS ON REUTERS DATABASE

Dataset	Method	ACC	NMI	AR	F-score	Precision	Recall
Reuters	SSC[47]	0.433±0.002	0.278±0.002	0.134±0.002	0.324±0.001	0.244±0.001	0.481±0.002
	LRR[48]	0.512±0.001	0.361±0.001	0.246±0.001	0.393±0.001	0.332±0.001	0.482±0.001
	S_0/L_0-LRSSC[31]	0.523±0.011	0.377±0.013	0.278±0.009	0.416±0.007	0.359±0.009	0.496±0.017
	RMSC[12]	0.496±0.039	0.296±0.020	0.238±0.020	0.372±0.015	0.350±0.017	0.398±0.014
	DiMSC[13]	0.476±0.007	0.307±0.002	0.175±0.006	0.345±0.003	0.276±0.005	0.459±0.007
	LT-MSC[5]	0.435±0.001	0.273±0.003	0.171±0.005	0.337±0.004	0.279±0.011	0.431±0.035
	MLAN[50]	0.490±0.003	0.355±0.001	0.250±0.003	0.389±0.002	0.346±0.002	0.444±0.002
	GMC[11]	0.277±0.000	0.235±0.000	0.027±0.000	0.287±0.000	0.177±0.000	0.748±0.000
	AWP[51]	0.398±0.000	0.289±0.000	0.169±0.000	0.359±0.000	0.260±0.000	0.580±0.000
	GLSR[18]	0.569±0.001	0.384±0.001	0.292±0.001	0.415±0.001	0.395±0.001	0.438±0.001
	LMSC[14]	0.512±0.004	0.317±0.005	0.238±0.004	0.375±0.003	0.346±0.003	0.408±0.004
	MCLES[52]	0.442±0.000	0.298±0.000	0.176±0.000	0.350±0.000	0.274±0.000	0.482±0.000
	tSVDMSC[1]	0.907±0.001	0.841±0.001	0.801±0.001	0.835±0.001	0.823±0.001	0.846±0.001
	SM^2SC[53]	0.554±0.004	0.400±0.001	0.306±0.001	0.433±0.001	0.391±0.001	0.484±0.003
	ETLMSC[26]	0.932±0.049	0.957±0.050	0.924±0.073	0.938±0.060	0.914±0.066	0.968±0.055
	LMVSC[54]	0.435±0.000	0.318±0.000	0.140±0.000	0.338±0.000	0.243±0.000	0.559±0.000
	DMF-MVC[17]	0.289±0.001	0.096±0.001	0.055±0.001	0.214±0.001	0.210±0.001	0.218±0.001
	MvDSCN[55]	0.316±0.105	0.112±0.127	0.072±0.095	0.236±0.085	0.217±0.066	0.263±0.018
	RMSL[39]	0.570±0.005	0.394±0.008	0.312±0.004	0.433±0.006	0.407±0.007	0.462±0.005
	GNLTA-Sch	0.966±0.076	0.963±0.040	0.953±0.077	0.961±0.062	0.953±0.085	0.970±0.034
	GNLTA-Lap	0.965±0.077	0.960±0.041	0.950±0.077	0.958±0.063	0.951±0.086	0.968±0.035
	GNLTA-Gem	0.967±0.077	0.967±0.042	0.955±0.077	0.963±0.063	0.956±0.086	0.972±0.035

TABLE VII
AVERAGE RUNNING TIME (IN SECONDS) ON ALL DATABASES

Method	DiMSC	LT-MSC	ECMSC	GLSR	LMSC	MCLES	tSVDMSC	ETLMSC	LMVSC	GNLTA
BBC4view	207.21	335.51	1238.70	68.73	180.38	295.89	97.99	48.06	**2.69**	6.74
BBCSport	38.15	77.23	266.86	24.92	59.28	107.02	19.59	4.21	**1.59**	2.21
StillDB	6.23	18.68	8.82	9.03	12.30	121.02	6.12	2.29	**0.96**	1.91
Flowers	110.21	322.19	193.66	156.24	330.22	9048.92	118.32	26.89	**4.19**	10.77
UCI	296.66	725.50	305.59	302.86	634.30	52909.01	158.26	58.58	**7.67**	33.95
MITIndoor	N/A	117720.7	N/A	N/A	11346.48	N/A	5143.26	3235.75	**52.75**	2311.0
Reuters	99.31	833.87	N/A	125.05	201.54	2705.25	101.01	31.52	**2.60**	10.24

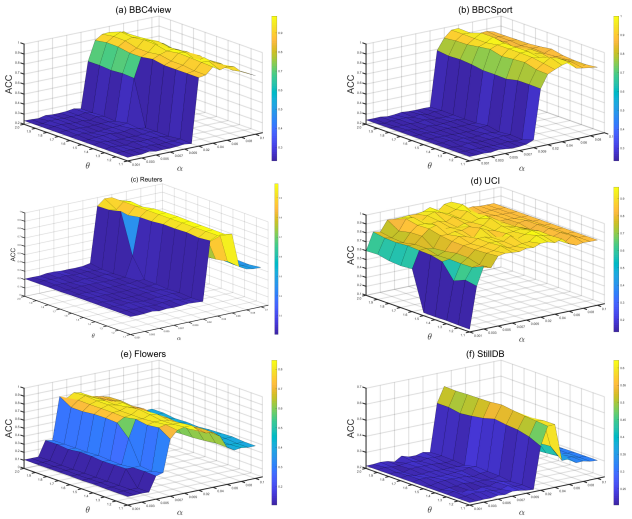

Fig. 2. Parameter tuning with respect to α and θ on (a) BBC4view, (b) BBCSport, (c) Reuters, (d) UCI-digits, (e) Flowers and (f) StillDB databases.

of the deep MVC methods is highly dependent on the training samples. Only with a large number of training samples, they may achieve high performance. (2) Essentially, these three deep MVC methods belong to the matrix optimization technique which considers only spatial two-dimensional correlation. However, our GNLTA employs the tensor optimization technique that explores not only the spatial two-dimensional correlation, but also the view relation among multiple views.

Additionally, the nonconvex single-view clustering method S_0/L_0-LRSSC consistently achieves better performance than the other two single-view methods LRR and SSC. This demonstrates that the nonconvex surrogate can achieve more accurate approximation than the convex nuclear norm and l_1 norm for clustering. However, all of them still perform worse than the proposed GNLTA on all databases, indicating that our generalized nonconvex low-rank tensor approximation is better for preserving the high-order correlation across views. The proposed GNLTA has robust performance on different types of multi-view databases, including the News (BBC4view and BBCSport), flower, handwritten digit (UCI-digit), still image, scene (MITIndoor), and Reuters databases. However, some of multi-view clustering methods perform unstable. Specifically, SM^2SC is the best one of all competing methods on BBC4view database while it achieves the similar performance to those of two single-view clustering methods on Flowers database.

C. Model Analysis

Here, we give a comprehensive model analysis of the proposed GNLTA with respect to parameter analysis, numerical convergence and running times.

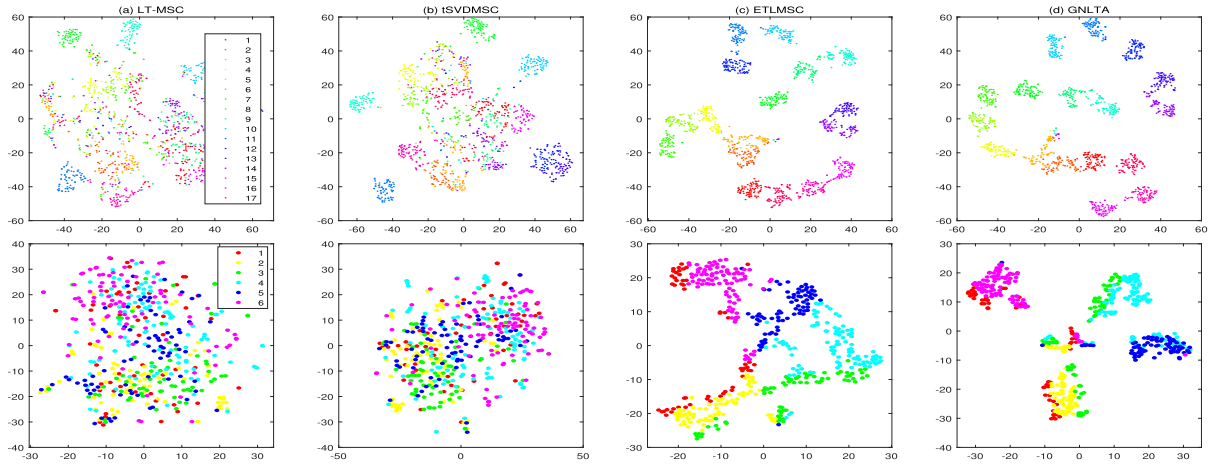

Fig. 3. Visualization of the embedding results on Flowers and StillDB databases (a) LT-MSC, (b) tSVDMSC, (c) ETLMSC, and (d) GNLTA-Gem.

The proposed GNLTA includes two free parameters α and θ. Parameter α is to balance the effect of the low-rank tensor term and the noise term. While parameter θ aims to measure the low-rank property of $\mathcal{Z}$. We tune them from interval [0.001, 0.1] and [1.1, 2], respectively. GNLTA-Gem is taken as the example of parameter setting. ACC values with different combinations of α and θ are shown in Fig. 2. We can observe that the clustering performance of the proposed GNLTA becomes better when α increases and then decreases. This is because the low-rank structure may be destroyed by the heavy noise term. In addition, the proposed GNLTA is insensitive to parameter θ. For example, when α is set to 0.06, GNLTA has the stable clustering results with different θ on BBC4view. This means that the parameter settings of GNLTA are easy to choose. The best performance is achieved by setting α and θ within $[0.005, 0.06]$ and $[1.1, 2]$, respectively.

We also investigate the numerical convergence of the proposed algorithm. Fig. 4 shows the relative errors defined in Eq. (21), ACC and NMI values versus iterations on the first six databases. It is easy to see that (1) as the number of iterations increases, ACC and NMI values tend to be stable; (2) the relative errors yield stable values after 10 iterations. Both of them indicate that GNLTA converges fast. This is also demonstrated in Table VII which reports the running times of some representative multi-view clustering methods on all databases. We find that GNLTA costs the second-shortest running time on all databases. LMVSC is more efficient than GNLTA but it does not produce very competitive performance on all databases compared with our GNLTA. The reason may be that this anchor graph strategy cannot well elaborate the similarity of all samples when the subsample is not representative of the original data. While the convex nuclear norm-based LT-MSC, LMSC and tSVDMSC cost more computation times because they usually need more iterations and perform the matrix inversion inevitably. This directly demonstrates the efficiency of the proposed GNLTA. We can conclude that the proposed GNLTA has boosted the clustering performance and costed less running times over several state-of-the-art multi-view clustering methods.

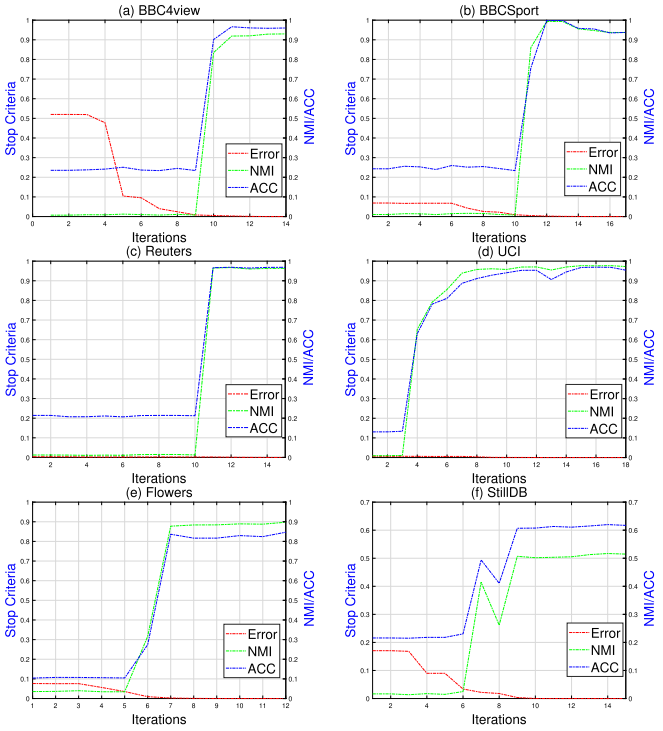

Fig. 4. Relative errors, ACC, and NMI versus iterations on (a) BBC4view, (b) BBCSport, (c) Reuters, (d) UCI-digits, (e) Flowers and (f) StillDB databases.

To investigate the superiority of our GNLTA visually, the embedding results of Flowers and StillDB databases are shown in Fig. 3. One can see that our GNLTA has revealed clearer cluster structure against other three nuclear norm-based methods. This consistently agrees with the clustering results in Tables III-VI.

V. Conclusion

In this paper, we developed a generalized nonconvex low-rank tensor approximation (GNLTA) model for MVSC. Unlike the existing low-rank tensor-based MVSC methods which use the convex nuclear norm to measure the low-rank property, GNLTA pursued a general and less bias approximation for the tensor rank. Based on the alternating

direction method of multipliers and the difference of convex method, an iteratively efficient algorithm was designed. The performance of the proposed GNLTA was investigated by conducting extensive experiments on seven real-world multi-view databases. In the future, we will investigate the generalized nonconvex low-rank tensor approximation with applications to hyperspectral image restoration, multi-view classification and so on.

REFERENCES

[1] Y. Xie, D. Tao, W. Zhang, Y. Liu, L. Zhang, and Y. Qu, "On unifying multi-view self-representations for clustering by tensor multi-rank minimization," *Int. J. Comput. Vis.*, vol. 126, no. 11, pp. 1157–1179, Nov. 2018.

[2] H. Zhao, H. Liu, Z. Ding, and Y. Fu, "Consensus regularized multi-view outlier detection," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 236–248, Jan. 2018.

[3] Z. Zhang, Y. Xie, W. Zhang, Y. Tang, and Q. Tian, "Tensor multi-task learning for person re-identification," *IEEE Trans. Image Process.*, vol. 29, pp. 2463–2477, 2020.

[4] L. Wang, Z. Ding, Z. Tao, Y. Liu, and Y. Fu, "Generative multi-view human action recognition," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 6212–6221.

[5] C. Zhang, H. Fu, S. Liu, G. Liu, and X. Cao, "Low-rank tensor constrained multiview subspace clustering," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1582–1590.

[6] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 849–856.

[7] F. Nie, X. Wang, M. I. Jordan, and H. Huang, "The constrained Laplacian rank algorithm for graph-based clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2016, pp. 1969–1976.

[8] F. Nie, G. Cai, and X. Li, "Multi-view clustering and semi-supervised classification with adaptive neighbours," in *Proc. AAAI Conf. Artif. Intell.*, 2017, pp. 2408–2414.

[9] K. Zhan, C. Zhang, J. Guan, and J. Wang, "Graph learning for multiview clustering," *IEEE Trans. Cybern.*, vol. 48, no. 10, pp. 2887–2895, Oct. 2018.

[10] K. Zhan, F. Nie, J. Wang, and Y. Yang, "Multiview consensus graph clustering," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1261–1270, Mar. 2019.

[11] H. Wang, Y. Yang, and B. Liu, "GMC: Graph-based multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 6, pp. 1116–1129, Jun. 2020.

[12] R. Xia, Y. Pan, L. Du, and J. Yin, "Robust multi-view spectral clustering via low-rank and sparse decomposition," in *Proc. AAAI Conf. Artif. Intell.*, 2014, pp. 2149–2155.

[13] X. Cao, C. Zhang, H. Fu, S. Liu, and H. Zhang, "Diversity-induced multi-view subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 586–594.

[14] C. Zhang *et al.*, "Generalized latent multi-view subspace clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 1, pp. 86–99, Jan. 2020.

[15] T. Zhou, C. Zhang, C. Gong, H. Bhaskar, and J. Yang, "Multiview latent space learning with feature redundancy minimization," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1655–1668, 2018.

[16] Y. Wang, L. Wu, X. Lin, and J. Gao, "Multiview spectral clustering via structured low-rank matrix factorization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4833–4843, Oct. 2018.

[17] H. Zhao, Z. Ding, and Y. Fu, "Multi-view clustering via deep matrix factorization," in *Proc. AAAI Conf. Artif. Intell.*, 2017, pp. 2921–2927.

[18] Y. Chen, S. Wang, F. Zheng, and Y. Cen, "Graph-regularized least squares regression for multi-view subspace clustering," *Knowl.-Based Syst.*, vol. 194, Apr. 2020, Art. no. 105482.

[19] S. Wang, Y. Chen, Y. Jin, Y. Cen, Y. Li, and L. Zhang, "Error-robust low-rank tensor approximation for multi-view clustering," *Knowl.-Based Syst.*, vol. 215, Mar. 2021, Art. no. 106745.

[20] Y. Chen, X. Xiao, C. Peng, G. Lu, and Y. Zhou, "Low-rank tensor graph learning for multi-view subspace clustering," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Feb. 3, 2021, doi: 10.1109/TCSVT.2021.3055625.

[21] Y. Xie, W. Zhang, Y. Qu, L. Dai, and D. Tao, "Hyper-Laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning," *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 572–586, Feb. 2020.

[22] Y. Chen, X. Xiao, and Y. Zhou, "Jointly learning kernel representation tensor and affinity matrix for multi-view clustering," *IEEE Trans. Multimedia*, vol. 22, no. 8, pp. 1985–1997, Aug. 2020.

[23] Y. Xie *et al.*, "Robust kernelized multiview self-representation for subspace clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 868–881, Feb. 2021.

[24] M. Yin, J. Gao, S. Xie, and Y. Guo, "Multiview subspace clustering via tensorial t-product representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 851–864, Mar. 2019.

[25] Y. Chen, X. Xiao, and Y. Zhou, "Multi-view clustering via simultaneously learning graph regularized low-rank tensor representation and affinity matrix," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2019, pp. 1348–1353.

[26] J. Wu, Z. Lin, and H. Zha, "Essential tensor learning for multi-view spectral clustering," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 5910–5922, Dec. 2019.

[27] Y. Chen, X. Xiao, Z. Hua, and Y. Zhou, "Adaptive transition probability matrix learning for multiview spectral clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 2, 2021, doi: 10.1109/TNNLS.2021.3059874.

[28] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jan. 2017.

[29] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5366–5380, Sep. 2017.

[30] Y. Chen, X. Xiao, and Y. Zhou, "Low-rank quaternion approximation for color image processing," *IEEE Trans. Image Process.*, vol. 29, pp. 1426–1439, 2020.

[31] M. Brbić and I. Kopriva, "ℓ_0-motivated low-rank sparse subspace clustering," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1711–1725, Apr. 2020.

[32] Y. Yang, J. Feng, N. Jojic, J. Yang, and S. T. Huang, "ℓ_0-sparse subspace clustering," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 731–747.

[33] Y.-P. Zhao, X. Lu, L. Chen, J. Tian, and C. L. P. Chen, "Robust nonconvex nonnegative low-rank representation," in *Proc. Int. Conf. Fuzzy Theory Its Appl. (iFUZZY)*, Nov. 2019, pp. 226–231.

[34] X. Peng, L. Zhang, and Z. Yi, "Scalable sparse subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 430–437.

[35] S. Matsushima and M. Brbic, "Selective sampling-based scalable sparse subspace clustering," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 12416–12425.

[36] J. Shen, P. Li, and H. Xu, "Online low-rank subspace clustering by basis dictionary pursuit," in *Proc. 33rd Int. Conf. Mach. Learn.*, 2016, pp. 622–631.

[37] P. Ji, T. Zhang, H. Li, M. Salzmann, and I. Reid, "Deep subspace clustering networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 24–33.

[38] M. Abavisani and V. M. Patel, "Deep multimodal subspace clustering networks," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1601–1614, Dec. 2018.

[39] R. Li, C. Zhang, H. Fu, X. Peng, J. T. Zhou, and Q. Hu, "Reciprocal multi-layer subspace learning for multi-view clustering," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 8172–8180.

[40] Y. Xie *et al.*, "Joint deep multi-view learning for image clustering," *IEEE Trans. Knowl. Data Eng.*, early access, Feb. 14, 2020, doi: 10.1109/TKDE.2020.2973981.

[41] P. Zhou, C. Lu, J. Feng, Z. Lin, and S. Yan, "Tensor low-rank representation for data recovery and clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Nov. 21, 2019, doi: 10.1109/TPAMI.2019.2954874.

[42] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, Aug. 2011.

[43] D. Zhou, J. Huang, and B. Schölkopf, "Learning from labeled and unlabeled data on a directed graph," in *Proc. 22nd Int. Conf. Mach. Learn. (ICML)*, 2005, pp. 1036–1043.

[44] S. Gaïffas and G. Lecué, "Weighted algorithms for compressed sensing and matrix completion," 2011, *arXiv:1107.1638*. [Online]. Available: http://arxiv.org/abs/1107.1638

[45] C. Lu, C. Zhu, C. Xu, S. Yan, and Z. Lin, "Generalized singular value thresholding," in *Proc. AAAI Conf. Artif. Intell.*, 2015, pp. 1–7.

[46] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[47] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[48] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
[49] X. Wang, X. Guo, Z. Lei, C. Zhang, and S. Z. Li, "Exclusivity-consistency regularized multi-view subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 923–931.
[50] F. Nie, G. Cai, J. Li, and X. Li, "Auto-weighted multi-view learning for image clustering and semi-supervised classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1501–1511, Mar. 2018.
[51] F. Nie, L. Tian, and X. Li, "Multiview clustering via adaptively weighted procrustes," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 2022–2030.
[52] M.-S. Chen, L. Huang, C.-D. Wang, and D. Huang, "Multi-view clustering in latent embedding space," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 1–8.
[53] Z. Yang, Q. Xu, W. Zhang, X. Cao, and Q. Huang, "Split multiplicative multi-view subspace clustering," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5147–5160, Oct. 2019.
[54] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, and Z. Xu, "Large-scale multi-view subspace clustering in linear time," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 4412–4419.
[55] P. Zhu, B. Hui, C. Zhang, D. Du, L. Wen, and Q. Hu, "Multi-view deep subspace clustering networks," 2019, *arXiv:1908.01978*. [Online]. Available: http://arxiv.org/abs/1908.01978
[56] N. Ikizler, R. G. Cinbis, S. Pehlivan, and P. Duygulu, "Recognizing actions from still images," in *Proc. 19th Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.
[57] A. Quattoni and A. Torralba, "Recognizing indoor scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 413–420.
[58] C. Lu, S. Yan, and Z. Lin, "Convex sparse spectral clustering: Single-view to multi-view," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2833–2843, Jun. 2016.
[59] Y. Chen, X. Xiao, and Y. Zhou, "Multi-view subspace clustering via simultaneously learning the representation tensor and affinity matrix," *Pattern Recognit.*, vol. 106, Oct. 2020, Art. no. 107441.

Yongyong Chen received the B.S. and M.S. degrees from the Shandong University of Science and Technology, Qingdao, China, in 2014 and 2017, respectively, and the Ph.D. degree from the University of Macau, Macau, in 2020. He is currently an Assistant Professor with the School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China. His research interests include image processing, data mining, and computer vision.

Shuqin Wang received the M.S. degree from the Shandong University of Science and Technology, Qingdao, China, in 2019. She is currently pursuing the Ph.D. degree with the Institute of Information Science, Beijing Jiaotong University. Her research interest includes multi-view learning.

Chong Peng received the Ph.D. degree in computer science from Southern Illinois University, Carbondale, IL, USA, in 2017. He is currently an Associate Professor with the College of Computer Science and Technology, Qingdao University. He has published more than 40 research papers in top-tier conferences and journals, including AAAI, ICDE, CVPR, SIGKDD, ICDM, CIKM, ACM TIST, TKDD, TOMM, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (TNNLS), TIP, TMM, TGRS, *Pattern Recognition*, *Neural Networks*, and *Information Sciences*. His research interests include pattern recognition, data mining, and machine learning.

Zhongyun Hua (Member, IEEE) received the B.S. degree in software engineering from Chongqing University, Chongqing, China, in 2011, and the M.S. and Ph.D. degrees in software engineering from the University of Macau, Macau, in 2013 and 2016, respectively.

He is currently an Associate Professor with the School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China. He has published over 30 technical papers at prestigious international journals and conferences, including IEEE TSP, TCYB, TSMC-S, TCAS-I, TMM, and TIE. His current research interests include chaotic systems, multimedia security, and image processing.

Yicong Zhou (Senior Member, IEEE) received the B.S. degree in electrical engineering from Hunan University, Changsha, China, and the M.S. and Ph.D. degrees in electrical engineering from Tufts University, Medford, MA, USA.

In 2011, he joined the University of Macau as an Assistant Professor with the Department of Computer and Information Science, where he is currently an Associate Professor and the Director of the Vision and Image Processing Laboratory. His research interests include image processing, computer vision, machine learning, and multimedia security. He is a Fellow of the Society of Photo-Optical Instrumentation Engineers (SPIE). He received the Third Prize of the Macao Natural Science Award as a sole winner in 2020 and a co-winner in 2014, and the Best Editor Award in 2020 to recognize his contribution to *Journal of Visual Communication and Image Representation*. Since 2015, he has been a leading Co-Chair of Technical Committee on Cognitive Computing in the IEEE Systems, Man, and Cybernetics Society. He serves as an Associate Editor for IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and four other journals. He was recognized as the "Highly Cited Researcher" in Web of Science in 2020.