

Anchor Structure Regularization Induced Multi-view Subspace Clustering via Enhanced Tensor Rank Minimization

Jintian Ji^{1,2}, Songhe Feng^{1,2,*}

¹Key Laboratory of Big Data & Artificial Intelligence in Transportation,
Ministry of Education, Beijing Jiaotong University, Beijing, 100044, China

²School of Computer and Information Technology, Beijing Jiaotong University, Beijing, 100044, China

{22120385, shfeng}@bjtu.edu.cn

Abstract

The tensor-based multi-view subspace clustering algorithms have received widespread attention due to the powerful ability to capture high-order correlation across views. Although such algorithms have achieved remarkable success, they still suffer from three main issues: 1)The extremely high computational complexity makes tensor-based methods difficult to handle large-scale data sets. 2)The subspace-based methods usually ignore the local geometric structure of the original data. 3)The commonly used Tensor Nuclear Norm (TNN) treats different singular values equally and under-penalizes the noise components, resulting in a sub-optimal representation tensor. Being aware of these, we propose Anchor Structure Regularization Induced Multi-view Subspace Clustering via Enhanced Tensor Rank Minimization (ASR-ETR). Specifically, an anchor-representation tensor is constructed by using the anchor representation strategy rather than the self-representation strategy to reduce the time complexity, and the local geometric structure in the learned anchor-representation tensor is enhanced by adopting the Anchor Structure Regularization (ASR). We further devise an Enhanced Tensor Rank (ETR), which is a tighter surrogate of the tensor rank to effectively capture the multi-view high-order correlation. An efficient iterative optimization algorithm is designed to solve the ASR-ETR, which is time-economical and enjoys favorable convergence. Extensive experimental results on various data sets demonstrate the superiority of the proposed algorithm as compared to state-of-the-art methods.

1. Introduction

Clustering is an important and fundamental task in unsupervised learning, which has been widely used in many

fields, such as pattern recognition, computer vision, objection recognition, etc. [2, 33]. However, with the development of technology, the storage and representation of data have become more diverse, for example, human fingerprints can be recorded via multiple sensors. Such data is known as multi-view data, and the emergence of multi-view data also makes it difficult for traditional clustering algorithms to meet current needs. Traditional clustering algorithms such as spectral clustering [32] can only deal with single-view data, which is difficult to utilize the rich complementary information among different views, so multi-view clustering is born to break this limitation by exploring the consensus information and complementary information hidden in different views.

The current multi-view clustering algorithms can be categorized into graph-based methods [41, 28, 35, 19], co-training-based methods [15], and subspace-based methods [22, 17, 6, 21, 44, 4, 7]. The subspace-based methods have received wide attention due to their efficient performance and noise robustness. They first construct the representation matrix of multiple views via self-representation and send the fused affinity matrix to spectral clustering to obtain the final results. The self-representation strategy always focuses on the global information among data points, leading to poor discriminability of the learned representation matrices. Therefore, many efforts have been made [1, 36, 11] to improve the discriminability of the representation matrix. For example, the Hilbert Schmidt Independence Criterion is employed in [1] to diversify the representation matrix of each view. However, such methods simply explore the relationship between point pairs and fail to effectively explore the higher-order correlation across views.

More recently, tensor-based algorithms [43, 25, 45, 26, 40], which aim to constrain the low rankness of the representation tensor composed by multi-view affinity matrices to capture the higher-order correlation and complementary information across views, are proposed and achieve promis-

*Corresponding Author

ing performances. However, such methods also bring several sharp problems: 1) Tensor-related operations such as Fourier transformation and tensor singular value decomposition (t-SVD) [14] result in extremely high time complexity, which limits the application to large-scale data sets. 2) Subspace-based methods usually consider the global consensus between multiple views and ignore the local geometry of a particular view, leading to a poorly learned clustering structure. 3) The commonly used tensor nuclear norm (TNN) [40] treats different singular values equally, which over-shrinks large singular values and under-penalizes small singular values during TNN minimization process. Since the large singular values usually characterize important information and small singular values encode noises or redundant information, TNN cannot effectually remove the redundancy of the learned representation tensor.

To overcome the above drawbacks simultaneously, we propose a novel multi-view clustering method, termed Anchor Structure Regularization Induced Multi-view Subspace Clustering via Enhanced Tensor Rank Minimization (ASR-ETR). Unlike existing methods that adopt the self-representation strategy to construct the tensor representation, ASR-ETR designs an anchor-representation strategy to construct the anchor-representation tensor, which can greatly reduce the complexity of tensor-related operations. Meanwhile, an Anchor Structure Regularization (ASR) is proposed to enhance the local geometric structure of the learned anchor-representation tensor. Furthermore, an Enhanced Tensor Rank (ETR) is devised to approximate the tensor rank, which can better characterize high-order correlation among multi-view representations, as well as achieves a stronger penalty for the noise against TNN. The novelty and contributions of this work are summarized as follows:

- ASR-ETR adopts an anchor-representation strategy to construct an anchor-representation tensor with a much lower dimension than the self-representation tensor in existing tensor-based methods, which greatly reduces the complexity of tensor-related operations and allows ASR-ETR to effectively handle large-scale data sets.
- ASR-ETR designs an Anchor Structure Regularization (ASR) to keep the local geometric structure and enhance the inter-class structure differences in the learned anchor-representation tensor.
- ASR-ETR proposes a novel tensor rank approximation termed Enhanced Tensor Rank (ETR), which leans to protect larger singular values and pushes smaller singular values to zero, thus a compact low-rank representation tensor can be desired.
- An efficient algorithm is proposed to solve the proposed objective function, which is time-economical and enjoys theoretically proven convergence.

2. Problem Formulation

2.1. Preliminary of Representation Tensor Learning

Tensor-based multi-view subspace clustering methods [8, 39, 10, 5, 3, 30] adopt low-rank representation tensor learning to capture higher-order correlation and complementary information among views. Given a multi-view data set $\{\mathbf{X}^1, \dots, \mathbf{X}^m\}$ with m views, $\forall v = 1, \dots, m$, $\mathbf{X}^v \in \mathbb{R}^{n \times d^v}$. Such methods usually take the following form:

$$\begin{aligned} \min_{\{\mathbf{Z}_S^v, \mathbf{E}^v\}} \mathcal{T}(\mathcal{Z}_S) + \lambda_1 \mathcal{L}(\{\mathbf{E}^v\}) + \lambda_2 \mathcal{R}(\{\mathbf{Z}_S^v\}), \\ s.t. \forall v, \mathbf{X}^v = \mathbf{Z}_S^v \mathbf{X}^v + \mathbf{E}^v, \\ \mathcal{Z}_S = \Phi(\mathbf{Z}_S^1, \dots, \mathbf{Z}_S^m), \end{aligned} \quad (1)$$

where λ_1 and λ_2 are two trade-off parameters. $\mathcal{T}(\cdot)$ is the approximation of tensor rank. $\mathcal{L}(\cdot)$ is designed to capture the reconstruction error \mathbf{E}^v , which usually adopts the $\ell_{2,1}$ -norm. $\mathcal{R}(\cdot)$ denotes some specific regularizations such as sparse regularization for self-representation matrix \mathbf{Z}_S^v . $\Phi(\cdot)$ denotes merging and rotating operation [40], which merges self-representation matrices $\mathbf{Z}_S^v \in \mathbb{R}^{n \times n}$ to a three-order tensor \mathcal{Z}_S with the dimension of $n \times m \times n$. With such self-representation tensor \mathcal{Z}_S , the tensor-related operations such as Fourier transformation and tensor singular value decomposition (t-SVD) will make the time complexity approximate to $\mathcal{O}(n^3)$.

As mentioned above, the model (1) suffers from three main problems: 1) The extremely high complexity of tensor-related operations. 2) Poor local geometric structure of the representation tensor. 3) The noise redundancy of the existing tensor rank approximation.

2.2. The Proposed ASR-ETR

The self-representation strategy adopted in the model (1) is essentially dictionary learning, which treats all samples as dictionary elements to represent the entire sample space, so the optimization time and storage cost related to the self-representation matrices restrict the scalability of the model (1). Besides, depicting one sample with a dictionary composed of all instances is unnecessary and redundant. Inspired by anchor-based algorithms [29, 23, 12], we take an anchor-representation strategy, which selects a small number of representative points called anchors or landmarks to serve as dictionary elements, to reduce the redundancy and dimension of representation matrices. Given the anchor dictionaries $\{\mathbf{A}^1, \dots, \mathbf{A}^m\}$ with m views and t anchors per view respectively, $\forall v$, $\mathbf{A}^v \in \mathbb{R}^{t \times d^v}$, then the model (1) with anchor-representation strategy can be formulated as:

$$\begin{aligned} \min_{\{\mathbf{Z}^v, \mathbf{A}^v, \mathbf{E}^v\}} \mathcal{T}(\mathcal{Z}) + \lambda_1 \mathcal{L}(\{\mathbf{E}^v\}) + \lambda_2 \mathcal{R}(\{\mathbf{Z}^v\}), \\ s.t. \forall v, \mathbf{X}^v = \mathbf{Z}^v \mathbf{A}^v + \mathbf{E}^v, \mathbf{A}^v (\mathbf{A}^v)^T = \mathbf{I}, \\ \mathcal{Z} = \Phi(\mathbf{Z}^1, \dots, \mathbf{Z}^m), \end{aligned} \quad (2)$$

where $\mathbf{Z}^v \in \mathbb{R}^{n \times t}$ is the learned anchor-representation matrix of v -th view. To avoid the local optimal solution brought by the initial anchor selection strategies [16], the anchor dictionary \mathbf{A}^v in Eq. (2) is regarded as variables to learn an optimal set of anchors $\{\mathbf{a}_1^v, \dots, \mathbf{a}_t^v\}$, where \mathbf{a}_i^v is the i -th row of \mathbf{A}^v , so it is unnecessary to focus on the selection of anchors but rather initialize \mathbf{A}^v to zero matrices. The anchor matrices are further imposed to be orthogonal that $\mathbf{A}^v(\mathbf{A}^v)^T = \mathbf{I}$ in each view to make the learned anchors more diverse and discriminative.

Since the dimension of learned anchor-representation tensor $\mathcal{Z} = \Phi(\mathbf{Z}^1, \dots, \mathbf{Z}^m) \in \mathbb{R}^{t \times m \times n}$ in the model (2) is much lower than that in self-representation tensor \mathcal{Z}_S , the time complexity of tensor-related operations can be greatly alleviated by employing the anchor-representation strategy, making it possible to address large-scale data sets.

To ensure favorable local geometric structure and noise robustness of the learned anchor-representation tensor, we design an Anchor Structure Regularisation (ASR) and an Enhanced Tensor Rank (ETR), respectively. They are defined as follows,

Definition 1. Given an anchor-representation tensor $\mathcal{Z} = \Phi(\mathbf{Z}^1, \dots, \mathbf{Z}^m) \in \mathbb{R}^{t \times m \times n}$, then the Anchor Structure Regularisation (ASR) is defined as:

$$\begin{aligned} \|\mathcal{Z}\|_{ASR} &= \sum_{v=1}^m \|\mathbf{Z}^v\|_{ASR} \\ &= \sum_{v=1}^m \text{Tr}(\mathbf{Z}^v \mathbf{L}^v (\mathbf{Z}^v)^T), \end{aligned} \quad (3)$$

where $\mathbf{L}^v = \mathbf{D}^v - \mathbf{B}^v$ is the Laplacian matrix of $\mathbf{B}^v \in \mathbb{R}^{t \times t}$, which is the adjacency matrix of v -th view anchors and is constructed with the same method in [18]. The degree matrix \mathbf{D}^v is a diagonal matrix whose i -th diagonal element is computed by $\mathbf{D}^v(i, i) = \sum_{j=1}^t \mathbf{B}^v(i, j)$.

Definition 2. Given a tensor $\mathcal{Z} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, then the Enhanced Tensor Rank (ETR) is defined as:

$$\begin{aligned} \|\mathcal{Z}\|_{ETR} &= \frac{1}{n_3} \sum_{k=1}^{n_3} \|\mathcal{Z}_f^k\|_{ETR} \\ &= \frac{1}{n_3} \sum_{k=1}^{n_3} \sum_{i=1}^h \left(\frac{e^{\delta^2 \mathcal{S}_f^k(i, i)}}{\delta + \mathcal{S}_f^k(i, i)} \right), \end{aligned} \quad (4)$$

where $0 < \delta \leq 1$, $h = \min(n_1, n_2)$ and \mathcal{S}_f is obtained by t -SVD of $\mathcal{Z}_f = \mathcal{U}_f \mathcal{S}_f \mathcal{V}_f^T$ in Fourier domain. \mathcal{Z}_f^k means the k -th frontal slice of the tensor \mathcal{Z}_f .

By simultaneously considering Eq. (2) (3) (4), the final objective function of ASR-ETR is formulated as:

$$\begin{aligned} \min_{\mathcal{Z}, \mathbf{E}, \{\mathbf{A}^v\}} & \|\mathcal{Z}\|_{ETR} + \alpha \|\mathbf{E}\|_{2,1} + \gamma \|\mathcal{Z}\|_{ASR}, \\ \text{s.t. } \forall v, & \mathbf{X}^v = \mathbf{Z}^v \mathbf{A}^v + \mathbf{E}^v, \mathbf{A}^v (\mathbf{A}^v)^T = \mathbf{I}, \\ & \mathcal{Z} = \Phi(\mathbf{Z}^1, \dots, \mathbf{Z}^m), \mathbf{E} = [\mathbf{E}^1, \dots, \mathbf{E}^m]^T, \end{aligned} \quad (5)$$

where α and γ are two trade-off parameters. $\mathbf{E}^v \in \mathbb{R}^{n \times d^v}$ is the error matrix of v -th view. $\mathbf{E} = [\mathbf{E}^1, \dots, \mathbf{E}^m]^T$ is obtained by horizontally concatenating along the row of \mathbf{E}^v . $\|\mathbf{E}\|_{2,1}$ denotes the $\ell_{2,1}$ -norm of \mathbf{E} , which depicts corruptions and outliers in each specific view.

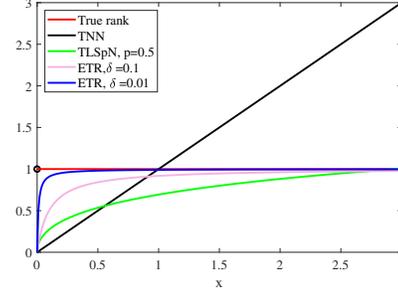


Figure 1: Comparison of different methods to approximate the true rank function.

Remark 1. [The superiority of ETR] The approximation function used by ETR is $f_{ETR}(x) = \frac{e^{\delta^2 x}}{\delta + x}$, which is inspired by the Geman function [9]. Basically, $f_{ETR}(0) = 0$ is satisfied, which is consistent with the true rank function. We compare ETR with other existing methods (i.e., TNN [40] and LTS_pN [10]), as shown in Fig. 1, ETR approximates the rank better than TNN and LTS_pN , especially for larger and near-zero singular values. When x is relatively large, we obtain $f_{ETR}(x) \rightarrow 1$, permitting larger singular values. If $x \rightarrow 0$, $f_{ETR}(x) \gg x$ and $f_{ETR}(x) \gg \log(1 + x^p)$, which means ETR achieves a stronger penalization on near-zero singular values than that in TNN and LTS_pN . Since smaller singular values usually come from noises, this indicates $\|\mathcal{Z}\|_{ETR}$ enjoys better noise-attention property. Such property helps to drive the noise out and make sure that \mathcal{Z} has a spatial low-rank structure, which helps to explore the high-order correlation and complementary information across views.

Remark 2. [The benefits of ASR] In the model (5), ASR is used to enhance the local geometric structure in anchor-representation tensor \mathcal{Z} . Unlike [27], which needs to compute the adjacency matrix of all samples, we enhance the local geometric structure of anchors to keep the clustering structure of the entire sample space, which brings the following benefits: 1) The adjacency matrix of anchors has low dimension, which can improve the efficiency of the algorithm. 2) Since an anchor point usually covers a class of samples, maintaining the local structure of anchors can improve the inter-class structural differences, thus enhancing the discriminability of the learned anchor-representation tensor.

After solving the model (5), the anchor-representation matrix $\mathbf{Z}^v = \Phi_v^{-1}(\mathcal{Z}) \in \mathbb{R}^{n \times t}$ of v -th view is obtained,

which cannot be directly employed for spectral clustering, so we recover a doubly-stochastic similarity affinity matrix $\mathbf{S}^v \in \mathbb{R}^{n \times n}$,

$$\mathbf{S}^v = \hat{\mathbf{Z}}^v (\hat{\mathbf{Z}}^v)^T = \mathbf{Z}^v (\Sigma^v)^{-1} (\mathbf{Z}^v)^T, \quad (6)$$

where Σ^v is a diagonal matrix with the entry $\Sigma^v(i, i) = \sum_{j=1}^n \mathbf{Z}^v(j, i)$. Then, the spectral clustering is implemented on the fused affinity matrix $\hat{\mathbf{S}} = \frac{1}{m} \sum_{v=1}^m \mathbf{S}^v$. Relying on proposition 1 in [13], the embedding $\mathbf{Q} \in \mathbb{R}^{n \times c}$ is obtained by performing SVD on $\hat{\mathbf{Z}} \in \mathbb{R}^{n \times tm}$,

$$\bar{\mathbf{Z}} = \frac{1}{\sqrt{m}} [\hat{\mathbf{Z}}^1, \dots, \hat{\mathbf{Z}}^m], \quad (7)$$

where $\hat{\mathbf{Z}}^v = \mathbf{Z}^v \Sigma^{v-\frac{1}{2}}$. Finally, k -means clustering is executed on \mathbf{Q} to achieve the final clustering results.

3. Optimization

Inspired by the alternation direction method of multipliers (ADMM) [20], we introduce the auxiliary tensor variable \mathcal{G} and rewrite the model (5) as the following unconstrained problem,

$$\begin{aligned} \mathcal{L}(\{\mathbf{Z}^v\}_{v=1}^m, \{\mathbf{A}^v\}_{v=1}^m \mathcal{G}, \mathbf{E}, \{\mathbf{Y}^v\}_{v=1}^m, \mathcal{W}) \\ = \|\mathcal{G}\|_{\text{ETR}} + \alpha \|\mathbf{E}\|_{2,1} + \gamma \sum_{v=1}^m \text{Tr}(\mathbf{Z}^v \mathbf{L}^v (\mathbf{Z}^v)^T) \\ + \sum_{v=1}^m \left(\langle \mathbf{Y}^v, \mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v - \mathbf{E}^v \rangle + \frac{\mu}{2} \|\mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v - \mathbf{E}^v\|_F^2 \right) \\ + \langle \mathcal{W}, \mathcal{Z} - \mathcal{G} \rangle + \frac{\rho}{2} \|\mathcal{Z} - \mathcal{G}\|_F^2, \end{aligned} \quad (8)$$

where the tensor \mathcal{W} and the matrix $\{\mathbf{Y}^v\}_{v=1}^m$ are Lagrange multipliers, and μ and ρ are penalty parameters to control convergence.

Then, the optimization problem can be divided into four subproblems.

3.1. \mathbf{Z}^v -Subproblem

Fixing the other variables leads to the following problem for \mathbf{Z}^v ,

$$\begin{aligned} \arg \min_{\{\mathbf{Z}^v\}} \gamma \sum_{v=1}^m \text{Tr}(\mathbf{Z}^v \mathbf{L}^v (\mathbf{Z}^v)^T) + \langle \mathcal{W}, \mathcal{Z} - \mathcal{G} \rangle \\ + \sum_{v=1}^m \left(\langle \mathbf{Y}^v, \mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v - \mathbf{E}^v \rangle \right. \\ \left. + \frac{\mu}{2} \|\mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v - \mathbf{E}^v\|_F^2 \right) + \frac{\rho}{2} \|\mathcal{Z} - \mathcal{G}\|_F^2. \end{aligned} \quad (9)$$

As the problem (9) is quadratic, smooth, and convex, it can be solved by the first-order optimality condition as follows,

$$\begin{aligned} \mathbf{Z}^v = (\mathbf{Y}^v (\mathbf{A}^v)^T + \mu \mathbf{X}^v (\mathbf{A}^v)^T + \rho \mathbf{G}^v \\ - \mu \mathbf{E}^v (\mathbf{A}^v)^T - \mathbf{W}^v) (2\gamma \mathbf{L}^v + (\rho + \mu) \mathbf{I})^{-1}. \end{aligned} \quad (10)$$

3.2. \mathbf{E} -Subproblem

Fixing the other variables, the problem with \mathbf{E} is formulated as,

$$\arg \min_{\mathbf{E}} \frac{\alpha}{\mu} \|\mathbf{E}\|_{2,1} + \frac{1}{2} \|\mathbf{E} - \hat{\mathbf{E}}\|_F^2, \quad (11)$$

where $\hat{\mathbf{E}}$ is constructed by horizontally concatenating the matrices $\mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v + \frac{1}{\mu} \mathbf{Y}^v$ together along row. Its solution can be obtained by $\ell_{2,1}$ minimization thresholding operator as in [24],

$$\mathbf{E}_{i,:} = \begin{cases} \frac{\|\hat{\mathbf{E}}_{i,:}\|_2 - \frac{\alpha}{\mu} \hat{\mathbf{E}}_{i,:}}{\|\hat{\mathbf{E}}_{i,:}\|_2} \hat{\mathbf{E}}_{i,:}, & \|\hat{\mathbf{E}}_{i,:}\|_2 > \frac{\alpha}{\mu}, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

where $\hat{\mathbf{E}}_{i,:}$ represents the i -th row of $\hat{\mathbf{E}}$.

3.3. \mathcal{G} -Subproblem

When other variables are fixed, the subproblem for \mathcal{G} is formulated as,

$$\arg \min_{\mathcal{G}} \frac{1}{\rho} \|\mathcal{G}\|_{\text{ETR}} + \frac{1}{2} \left\| \mathcal{G} - \left(\mathcal{Z} + \frac{\mathcal{G}}{\rho} \right) \right\|_F^2. \quad (13)$$

We refer to this problem as the Enhanced Tensorial Rank Minimization problem (ETRM), which can be solved by the following theorem.

Theorem 1. Suppose $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ with t -SVD $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ and $\beta > 0$. The Enhanced Tensorial Rank Minimization problem (ETRM) can be described as follows,

$$\arg \min_{\mathcal{G}} \beta \|\mathcal{G}\|_{\text{ETR}} + \frac{1}{2} \|\mathcal{G} - \mathcal{A}\|_F^2. \quad (14)$$

Then, optimal solution \mathcal{G}^* is obtained as,

$$\mathcal{G}^* = \mathcal{U} * \text{ifft}(\text{Prox}_{f,\beta}(\mathcal{S}_f), [], 3) * \mathcal{V}^T, \quad (15)$$

where $\text{ifft}(\text{Prox}_{f,\beta}(\mathcal{S}_f), [], 3) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is a f -diagonal tensor, and $\text{Prox}_{f,\beta}(\mathcal{S}_f^k(i, i))$ satisfies the following equation,

$$\text{Prox}_{f,\beta}(\mathcal{S}_f^k(i, i)) = \arg \min_{x \geq 0} \frac{1}{2} (x - \mathcal{S}_f^k(i, i))^2 + \beta f(x), \quad (16)$$

where $f(x) = \frac{e^{\delta^2 x}}{\delta + x}$.

The proof of Theorem 1 is given in the supplementary. Eq. (16) is a combination of concave and convex functions, so we can use the difference of convex (DC) programming [31] to acquire a closed-form solution until iteration converges,

$$\tau^{iter+1} = \left(\mathcal{S}_f^k(i, i) - \frac{\partial f(\tau^{iter})}{\rho} \right)_+, \quad (17)$$

where $\tau = \text{Prox}_{f,\beta}(\mathcal{S}_f^k(i, i))$, $f(x) = \frac{e^{\delta^2 x}}{\delta + x}$ and $iter$ is the number of iterations.

3.4. A^v -Subproblem

Fixing the other variables, A^v can be updated by,

$$A^{v*} = \arg \max_{A^v (\mathbf{A}^v)^T = \mathbf{I}} Tr((\mathbf{A}^v)^T \mathbf{M}^v), \quad (18)$$

where $\mathbf{M}^v = (\mathbf{Z}^v)^T (\mu \mathbf{X}^v + \mathbf{Y}^v - \mu \mathbf{E}^v)$. The optimal solution of A^v is $\mathbf{U}_v \mathbf{V}_v^T$, where \mathbf{U}_v and \mathbf{V}_v are the left and right singular matrix of \mathbf{M}^v .

At last, the Lagrange multipliers and penalty parameters are updated as follows,

$$\begin{cases} \mathbf{Y}^v = \mathbf{Y}^v + \mu(\mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v - \mathbf{E}^v), \\ \mathcal{W} = \mathcal{W} + \rho(\mathcal{Z} - \mathcal{G}), \\ \mu = \eta_\mu \mu, \rho = \eta_\rho \rho, \end{cases} \quad (19)$$

where $\eta_\mu, \eta_\rho > 1$ are used to accelerate convergence. The detailed algorithm of our model is summarized in Algorithm 1.

Algorithm 1 Optimization Algorithm for Eq. (5)

Input: Multi-view data matrix $\{\mathbf{X}^1, \dots, \mathbf{X}^m\}$, anchor number t , cluster number c , parameters γ and α ,

Initialize: Initialize $\mathbf{B}^v, \mathbf{Z}^v, \mathbf{E}^v, \mathbf{Y}^v, \mathbf{A}^v$ to zero matrix, $\mathcal{G} = \mathcal{W} = \mathbf{0}_{n \times t \times m}$, $\mu = 10^{-5}, \rho = 10^{-4}, \eta_\mu = \eta_\rho = 2, \mu_{max} = \rho_{max} = 10^{10}, \epsilon = 10^{-7}$;

- 1: **while** not converge **do**
 - 2: Update \mathbf{Z}^v by Eq. (9);
 - 3: Update \mathbf{E}^v by Eq. (12);
 - 4: Update \mathcal{G} by Eq. (17);
 - 5: Update \mathbf{A}^v by Eq. (18);
 - 6: Update \mathbf{B}^v with the updated \mathbf{A}^v ;
 - 7: Update \mathbf{Y}^v and \mathcal{W} by Eq. (19);
 - 8: Update parameters μ and ρ by Eq. (19) and $\mu = \min(\mu, \mu_{max}), \rho = \min(\rho, \rho_{max})$;
 - 9: Check the convergence conditions:
 $\|\mathbf{X}^v - \mathbf{Z}^v \mathbf{A}^v - \mathbf{E}^v\|_\infty < \epsilon$ & $\|\mathcal{Z} - \mathcal{G}\|_\infty < \epsilon$
 - 10: **end while**
 - 11: Compute \mathbf{Q} by performing SVD on $\bar{\mathbf{Z}}$ in Eq. (7), and output clustering results via performing k -means clustering on \mathbf{Q} .
-

3.5. Convergence Analysis

The convergence of Algorithm 1 is ensured by the Theorem 2 and the proof has been given in the supplementary.

Theorem 2. Let $\{\mathcal{P}_k = (\mathbf{Z}_k^v, \mathbf{E}_k^v, \mathbf{A}_k^v, \mathbf{Y}_k^v, \mathcal{W}_k, \mathcal{G}_k)\}_{k=1}^\infty$ be the sequence generated by Algorithm 1, then the sequence $\{\mathcal{P}_k\}_{k=1}^\infty$ satisfies the following two principles:

- 1). $\{\mathcal{P}_k\}_{k=1}^\infty$ is bounded.
- 2). Any accumulation point of $\{\mathcal{P}_k\}_{k=1}^\infty$ is a KKT point of Eq. (8).

Table 1: Summary of the benchmark data sets.

Data set	Type	Sam./Clu.	Views
BBCSport	Text	544 / 5	2
NGs	Text	500 / 5	3
CCV	Video	6773 / 20	3
Caltech101-all	Object	9144 / 102	6
Aloi-100	Object	11025 / 100	4
CIFAR10	Object	50000 / 10	3
Noisy MNIST	Digit	50000 / 10	2

3.6. Complexity Analysis

ASR-ETR consists of two stages: 1) optimization by iterative solving Eq. (8). 2) clustering stage. The first stage mainly focuses on the updates of five variables $\{\mathbf{Z}^v, \mathbf{E}^v, \mathcal{G}, \mathbf{A}^v, \mathbf{B}^v\}$, and the time complexity spent on these variables is $\mathcal{O}(nt^2 + ntd)$, $\mathcal{O}(nd)$, $\mathcal{O}(mnt \log(mn) + nm^2t)$, $\mathcal{O}(ntd^v + t^2d^v)$, $\mathcal{O}(t^2(d^v)^2)$, respectively. The second stage takes $\mathcal{O}(nt^2m^2)$, where $d = \sum_{v=1}^m d^v$. As $t \ll n$, the total time complexity of our ASR-ETR is $\mathcal{O}(ntd + mnt \log(mn))$. The space complexity is $\mathcal{O}(ntm + nd_{max})$, $d_{max} = \max(d^v)$, which is linear to n .

4. Experiment

In this section, extensive experiments are conducted to verify the effectiveness and superiority of our ASR-ETR. Due to the page limit, we show partial experimental results, for more experimental results, please refer to the supplementary. All the experiments are implemented on a computer with a 2.50GHz i7-11700 CPU and 64GB RAM, Matlab R2021a.

4.1. Experimental Settings

Data sets: Seven challenging data sets are adopted for the validation of our ASR-ETR, including **NGs**, **BBC-Sport**, **CCV**, **Caltech101-all**, **Aloi-100**, **CIFAR10**, and **Noisy MNIST**. More details can be found in Table 1.

Baselines: To verify the superiority of our model, we compare ASR-ETR with nine state-of-the-art multi-view clustering methods, including MVGL(2017) [42], GMC(2019) [34], EOMSC-CA(2022) [23], SFMC(2020) [18], SMVSC(2021) [29], t-SVD-MS(2018) [40], ETLMSC(2019) [37], and TBGL(2022) [38], respectively. Furthermore, the standard single-view spectral clustering (SC) algorithm [32] is included as a baseline, and we perform spectral clustering on each view and show the best results.

Evaluation Metrics: To comprehensively measure the clustering quality, we adopt four commonly used metrics, including accuracy (ACC), normalized mutual information

Table 2: Results (mean(std)) of our proposed method and other compared methods on seven data sets. 'OM' indicates the "out-of-memory error".

Data set	Metric (%)	SC _{best}	MVGL	GMC	EOMSC-CA	SFMC	SMVSC	t-SVD-MSC	ETLMSC	TBGL	Ours
NGs	ACC	25.90(0.25)	22.80(0)	<u>98.20(0)</u>	63.00(0)	22.20(0)	74.20(0)	90.00(0)	24.46(1.38)	32.20(0)	100(0)
	NMI	1.81(0.13)	6.87(0)	<u>93.92(0)</u>	47.85(0)	2.87(0)	52.38(0)	76.64(0)	5.80(1.66)	9.35(0)	100(0)
	PUR	25.90(0.25)	23.80(0)	<u>98.20(0)</u>	64.80(0)	22.80(0)	74.20(0)	81.84(0)	20.06(0.25)	32.80(0)	100(0)
	ARI	0.52(0.11)	0.21(0)	<u>95.54(0)</u>	41.30(0)	0.18(0)	47.25(0)	77.57(0)	0.50(0.57)	9.32(0)	100(0)
BBCSport	ACC	50.35(0.35)	22.8(0)	<u>80.70(0)</u>	46.32(0)	36.40(0)	59.01(0)	34.38(0)	50.17(2.64)	54.78(0)	100(0)
	NMI	21.00(0.22)	6.87(0)	<u>72.26(0)</u>	21.73(0)	1.75(0)	30.07(0)	1.45(0)	31.97(0.98)	27.75(0)	100(0)
	PUR	55.31(0.35)	23.80(0)	<u>84.38(0)</u>	50.37(0)	36.95(0)	61.40(0)	23.60(0)	40.65(0.71)	54.96(0)	100(0)
	ARI	15.93(0.45)	0.21(0)	<u>72.18(0)</u>	15.04(0)	0.56(0)	27.63(0)	-0.65(0)	22.64(1.71)	18.75(0)	100(0)
CCV	ACC	13.69(0.22)	11.32(0)	10.63(0)	24.32(0)	15.49(0)	22.52(0)	<u>47.80(0.04)</u>	19.23(0.49)	15.13(0)	78.92(0)
	NMI	9.51(0.13)	3.15(0)	0.38(0)	18.70(0)	1.75(0)	16.18(0)	<u>41.34(0.05)</u>	15.42(0.34)	7.42(0)	75.71(0)
	PUR	17.50(0.15)	11.68(0)	10.79(0)	26.68(0)	17.05(0)	25.41(0)	<u>35.25(0.05)</u>	12.02(0.16)	16.34(0)	79.14(0)
	ARI	3.25(0.13)	0.06(0)	0.02(0)	8.52(0)	3.17(0)	7.17(0)	<u>29.61(0.05)</u>	6.30(0.14)	1.45(0)	64.67(0)
Caltech101-all	ACC	19.50(0.86)	14.34(0)	19.50(0)	24.70(0)	23.15(0)	29.18(0)	<u>47.45(1.01)</u>	21.65(1.11)	23.20(0)	65.78(0)
	NMI	40.57(0.33)	25.14(0)	23.79(0)	27.09(0)	19.02(0)	40.39(0)	<u>71.49(0.44)</u>	43.93(0.27)	21.39(0)	87.27(0)
	PUR	40.61(0.52)	22.86(0)	30.12(0)	27.64(0)	28.07(0)	38.17(0)	<u>53.74(1.88)</u>	36.71(2.76)	29.45(0)	83.83(0)
	ARI	13.44(0.92)	-0.71(0)	-0.42(0)	10.02(0)	0(0)	23.27(0)	<u>31.60(1.06)</u>	19.21(2.01)	0.16(0)	54.15(0)
Aloi-100	ACC	64.54(1.15)	53.44(0)	<u>72.07(0)</u>	23.70(0)	65.53(0)	34.35(0)	71.99(1.44)	66.82(2.25)	68.58(0)	83.25(0)
	NMI	80.03(0.33)	67.22(0)	74.36(0)	57.93(0)	68.72(0)	60.99(0)	83.92(0.43)	81.67(0.53)	72.20(0)	92.63(0)
	PUR	67.28(0.94)	55.94(0)	<u>73.07(0)</u>	24.92(0)	66.38(0)	35.58(0)	<u>58.03(1.56)</u>	53.05(2.73)	69.81(0)	85.86(0)
	ARI	54.29(1.08)	4.55(0)	15.84(0)	7.42(0)	11.31(0)	20.39(0)	<u>61.71(1.13)</u>	57.78(2.06)	14.49(0)	77.64(0)
CIFAR10	ACC	89.41(4.06)	OM	OM	<u>99.08(0)</u>	97.77(0)	98.35(0)	OM	OM	OM	100(0)
	NMI	80.08(1.41)	OM	OM	<u>97.48(0)</u>	94.63(0)	97.03(0)	OM	OM	OM	100(0)
	PUR	89.67(3.25)	OM	OM	<u>99.08(0)</u>	97.77(0)	98.85(0)	OM	OM	OM	100(0)
	ARI	79.79(3.00)	OM	OM	<u>97.97(0)</u>	95.14(0)	97.50(0)	OM	OM	OM	100(0)
Noisy MNIST	ACC	61.10(2.78)	OM	OM	<u>58.65(0)</u>	41.12(0)	18.44(0)	OM	OM	OM	99.84(0)
	NMI	61.82(1.58)	OM	OM	<u>49.08(0)</u>	34.81(0)	11.77(0)	OM	OM	OM	99.48(0)
	PUR	68.34(1.23)	OM	OM	<u>58.83(0)</u>	45.35(0)	22.97(0)	OM	OM	OM	99.84(0)
	ARI	50.02(2.52)	OM	OM	<u>40.44(0)</u>	16.55(0)	6.02(0)	OM	OM	OM	99.63(0)

Table 3: Average running time (sec.) comparison of multi-view clustering methods under data sets with more than 6,000 samples. 'OM' indicates the "out-of-memory error".

Data set	MVGL	GMC	EOMSC-CA	SFMC	SMVSC	t-SVD-MSC	ETLMSC	TBGL	Ours
CCV	1613.51	476.03	37.94	<u>10.99</u>	221.96	2630.32	3201.42	2255.34	3.18
Caltech101-all	6006.82	461.73	<u>98.82</u>	59.06	980.72	13841.21	15768.25	10401.31	288.49
Aloi-100	2808.92	633.14	<u>86.39</u>	46.11	658.18	14494.12	12723.68	11000.23	115.07
CIFAR10	OM	OM	<u>97.29</u>	72.18	547.36	OM	OM	OM	<u>133.05</u>
Noisy MNIST	OM	OM	99.57	74.30	528.41	OM	OM	OM	<u>88.78</u>

(NMI), purity (PUR), and adjusted rand index (ARI). For all metrics, the larger score, the higher clustering quality.

Parameter Setting: For the parameters in the baselines, we follow the settings in the corresponding paper and report the best results. For fairness, the number of anchors is searched in $\{2c, 3c, \dots, 8c\}$ for all anchor-based algorithms, where c is the number of clusters. For our ASR-ETR, the parameters α and γ are both searched in $\{10^{-6}, 10^{-5}, \dots, 0\}$, and the search range for δ is $\{0.0001, \dots, 1\}$. We run all the methods 20 times independently on each data set and report their averages and standard deviations.

4.2. Experimental Results

Table 2 and Table 3 show the comparison in metrics and running time of the above methods on seven data sets, where the best values and second best values are denoted by **bold**

values and underlined values. From the results, we derive the following interesting observations.

1) Single-view spectral clustering (SC_{best}) is inferior to multi-view clustering algorithms in most cases since all the multi-view clustering methods explore the complementary information among the views, while SC_{best} is the exception.

2) Tensor-based methods outperform non-tensor-based algorithms in most cases. It is mainly attributed to the fact that the tensor-based methods use the low rankness of the representation tensor to capture the higher-order correlation across views, which is lacking in non-tensor algorithms.

3) Our ASR-ETR consistently outperforms all other comparison methods over all metrics and all data sets. Especially, ASR-ETR achieves the **ideal clustering performance** on NGs, BBCSport, and CIFAR10 data sets. Furthermore, the improvement of our method on some chal-

lenging data sets is remarkable. For example, on the CCV data set, ASR-ETR improves around **31.12%**, **34.37%**, **43.89%**, **35.06%** in terms of four metrics over the second-best t-SVD-MSc method. For the Noisy MNIST data set, ASR-ETR outperforms the second-best EOMSC-CA method by **41.19%**, **50.4%**, **41.01%**, **59.15%**, respectively. These results consistently validate the effectiveness of ASR-ETR. And, under different applications (e.g., text clustering, object clustering, etc), ASR-ETR always achieves satisfactory performance, which further verifies its robustness of complex applications.

4) Despite ASR-ETR takes a little longer time as compared with the anchor-based and non-tensor-based methods, e.g., EOMSC-CA, SFMC, and SMVSC, it shows excellent performance that significantly beats these methods on all challenging data sets.

5) Compared to tensor-based methods such as t-SVD-MSc, ETLMSc, and TBGL that suffer from out-of-memory issues, both the time complexity and space complexity of ASR-ETR are significantly reduced as linear to the number of samples, making it possible to deal with large-scale data sets.

4.3. Ablation Studies

Influence of Enhanced Tensor Rank: To analyze the effect of our proposed Enhanced Tensor Rank (ETR), we devise two ablation experiments, in the first one we analyze the parameter δ of the Enhanced Tensor Rank with the search range of $\{10^{-4}, \dots, 1\}$. As shown in Fig. 2, we can observe that the parameter has a significant effect on the clustering results. The best clustering results of NGs are obtained when $\delta = 0.1$, but CCV peaks at $\delta = 10^{-4}$. The reason for this phenomenon is that since δ determines the strength of the penalty for different singular values and the distribution of singular values varies differs for different datasets, a suitable parameter is desired to provide better discriminability of the learned low-rank representation tensor.

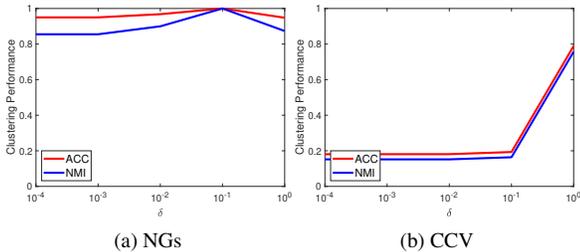


Figure 2: The performance (i.e., ACC and NMI) of ASR-ETR with varying parameter δ on two data sets.

The second experiment is to validate the superiority of the proposed Enhanced Tensor Rank (ETR). We replace the Enhanced Tensor Rank in the model (5) with the ten-

sor nuclear norm (TNN) and perform the experiments on five data sets. As shown in Fig. 3(a), it can be seen that ASR-ETR completely beats ASR-TNN, the main reason for this phenomenon is that TNN treats different singular values equally, leading to the over-punishment of important components and the under-punishment of noises, while ETR imposes different penalties on different singular values to ensure the removal of noise and the retention of important information.

Influence of Anchor Structure Regularization: In this paper, we propose an Anchor Structure Regularization (ASR) to keep the local geometric structure in different views. To verify the impact of this term, Fig. 3(b) shows the clustering results of the proposed method with/without the ASR on five datasets. It can be seen that the clustering performance is boosted via the Anchor Structure Regularization, which strongly demonstrates the effectiveness of our proposed ASR.

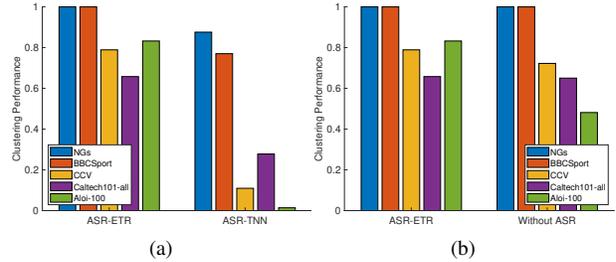


Figure 3: (a).The clustering performance of ASR-ETR and ASR-TNN on five data sets. (b).The clustering performance of ASR-ETR and without ASR on five data sets.

4.4. Model Analysis

Parameters Analysis: Our ASR-ETR has two parameters α and γ that need to be tuned, we empirically search for the optimal parameters in each data set, where the search range for both parameters is $\{10^{-6}, 10^{-5}, \dots, 1\}$. Due to the page limit, we only show the clustering performance (ACC) on the two data sets BBCSport and NGs in Fig. 4. We can observe that when α takes $\{10^{-3}, \dots, 1\}$ and γ in $\{10^{-6}, \dots, -3\}$, the ACC of both BBCSport and NGs keep a high level. This phenomenon shows that the performance of our ASR-ETR is stable across a wide range of parameters, and the stability of the parameters also directly illustrates the efficiency and robustness of ASR-ETR.

Anchor Analysis: We empirically analyze the impact of the anchor numbers on four data sets (i.e., NGs, BBCSport, Caltech101-all, and AloI-100). To this end, We set the variation range of anchor points to $\{2c, 3c, \dots, 8c\}$, then the clustering metrics ACC and NMI are shown in Fig. 5, we observe that different data set achieve the best clustering results at different numbers of anchor points, such as NGs achieves the highest performance at $8c$ and Caltech101-all

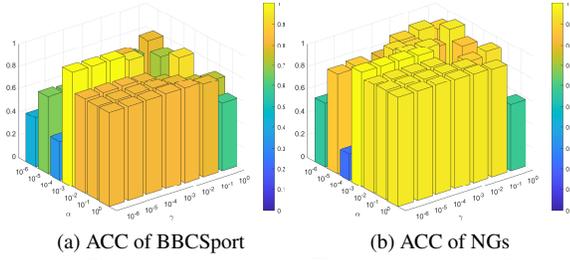


Figure 4: Parameters analysis: The clustering performances (ACC) with different parameters α and γ on the BBCSport and NGs data sets.

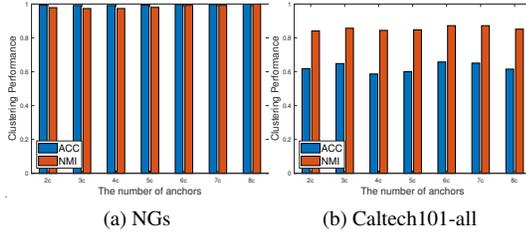


Figure 5: Anchor Analysis: The performance (i.e., ACC and NMI) of ASR-ETR on two data sets by varying the number of anchors.

performs best at $6c$. However, in a global view, the clustering results are stable under different anchors on the two data sets, which demonstrates that the anchor-representation strategy is robust to the number of anchors, and it is not necessary to use numerous anchors for clustering.

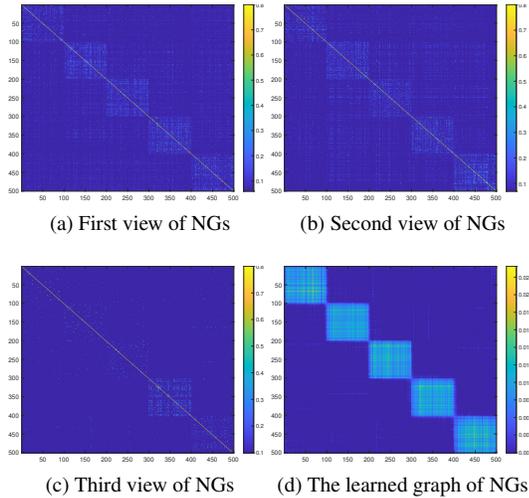


Figure 6: Graph Visualization: The visualization of the input graph and the learned graph on NGs data set

Graph Visualization: We present three input graphs of the data set NGs and the learned graphs in Fig. 6, where (a)-(c) are the input graphs of three views, and (d) is the learned graph of our ASR-ETR. It is obvious that the clustering structures of the input graphs are not clear and con-

tain many noises. After adopting ASR-ETR, the learned graph exhibits clear five components, exactly for the five categories. This indicates that our proposed algorithm can well characterize the cluster structure and achieve a favorable noise removal capability.

Convergence Analysis: The convergence of our ASR-ETR is guaranteed by the Theorem 2, in this section, we design convergence experiments to further verify the stability by recording the values of the stop criteria, where the stop criteria used here are Reconstruction Error (RE): $RE = \max_v \|\mathbf{X}^v - \mathbf{X}^v \mathbf{Z}^v - \mathbf{E}^v\|_\infty$ and Match Error (ME): $ME = \|\mathcal{Z} - \mathcal{G}\|_\infty$. Due to the page limits, we show the results of two data sets in the Fig. 7, the values of RE and ME rapidly tend to 0 within 10 steps and remain stable, which indicates the excellent convergence property of our ASR-ETR.

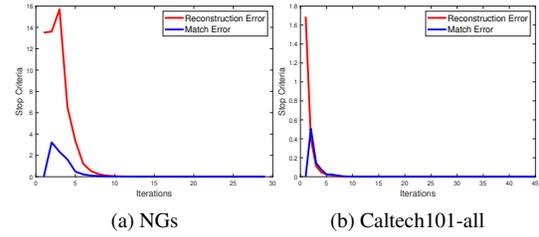


Figure 7: Convergence Analysis: The stop criteria (i.e., RE and ME) variation curves on two data sets.

5. Conclusion

In this article, we propose an Anchor Structure Regularization Induced Multi-view Subspace Clustering via Enhanced Tensor Rank Minimization (ASR-ETR). ASR-ETR adopts the anchor-representation strategy to construct an anchor-representation tensor, which greatly accelerates the tensor-related operations. An Anchor Structure Regularization is designed to keep the local geometric structure while enhancing the structural differences between classes. Furthermore, ASR-ETR exploits the similarity of the inter-view via Enhanced Tensor Rank minimization, which well explores the complementary information embedded in different views. An efficient scheme is employed to optimize the proposed model, which enjoys both low time complexity and linear space complexity. Extensive experiments on seven challenging data sets demonstrate the superiority of ASR-ETR.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (No. 2022JBZY019).

References

- [1] Xiaochun Cao, Changqing Zhang, Huazhu Fu, Si Liu, and Hua Zhang. Diversity-induced multi-view subspace clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 586–594, 2015. 1
- [2] Guoqing Chao, Shiliang Sun, and Jinbo Bi. A survey on multiview clustering. *IEEE Transactions on Artificial Intelligence*, 2(2):146–168, 2021. 1
- [3] Man-Sheng Chen, Chang-Dong Wang, and Jian-Huang Lai. Low-rank tensor based proximity learning for multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–1, 2022. 2
- [4] Peng Chen, Liang Liu, Zhengrui Ma, and Zhao Kang. Smoothed multi-view subspace clustering. In *International Conference on Neural Computing for Advanced Applications*, pages 128–140. Springer, 2021. 1
- [5] Yongyong Chen, Shuqin Wang, Chong Peng, Zhongyun Hua, and Yicong Zhou. Generalized nonconvex low-rank tensor approximation for multi-view subspace clustering. *IEEE Transactions on Image Processing*, 30:4022–4035, 2021. 2
- [6] Ehsan Elhamifar and René Vidal. Sparse subspace clustering: algorithm, theory, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2765–2781, 2013. 1
- [7] Hongchang Gao, Feiping Nie, Xuelong Li, and Heng Huang. Multi-view subspace clustering. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 4238–4246, 2015. 1
- [8] Quanxue Gao, Wei Xia, Zhizhen Wan, Deyan Xie, and Pu Zhang. Tensor-svd based graph learning for multi-view subspace clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3930–3937, 2020. 2
- [9] Donald Geman and Chengda Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, 4(7):932–946, 1995. 3
- [10] Jipeng Guo, Yanfeng Sun, Junbin Gao, Yongli Hu, and Bao-cai Yin. Logarithmic Schatten-p norm minimization for tensorial multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2022. 2, 3
- [11] Shudong Huang, Yixi Liu, Ivor W Tsang, Zenglin Xu, and Jiancheng Lv. Multi-view subspace clustering by joint measuring of consistency and diversity. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–12, 2022. 1
- [12] Zhao Kang, Zhiping Lin, Xiaofeng Zhu, and Wenbo Xu. Structured graph learning for scalable subspace clustering: From single view to multiview. *IEEE Transactions on Cybernetics*, 52(9):8976–8986, 2021. 2
- [13] 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 AAAI conference on artificial intelligence*, volume 34, pages 4412–4419, 2020. 4
- [14] Misha E Kilmer and Carla D Martin. Factorization strategies for third-order tensors. *Linear Algebra and its Applications*, 435(3):641–658, 2011. 2
- [15] Abhishek Kumar, Piyush Rai, and Hal Daume. Co-regularized multi-view spectral clustering. *Advances in Neural Information Processing Systems*, 24:1413–1421, 2011. 1
- [16] Lusi Li and Haibo He. Bipartite graph based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 34(7):3111–3125, 2020. 3
- [17] Ruihuang Li, Changqing Zhang, Huazhu Fu, Xi Peng, Tianyi Zhou, and Qinghua Hu. Reciprocal multi-layer subspace learning for multi-view clustering. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 8172–8180, 2019. 1
- [18] Xuelong Li, Han Zhang, Rong Wang, and Feiping Nie. Multiview clustering: A scalable and parameter-free bipartite graph fusion method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):330–344, 2020. 3, 5
- [19] Yeqing Li, Feiping Nie, Heng Huang, and Junzhou Huang. Large-scale multi-view spectral clustering via bipartite graph. In *Proceedings of the AAAI Conference on Artificial Intelligence*, page 2750–2756, 2015. 1
- [20] Zhouchen Lin, Risheng Liu, and Zhixun Su. Linearized alternating direction method with adaptive penalty for low-rank representation. *Advances in Neural Information Processing Systems*, 24, 2011. 4
- [21] Guangcan Liu, Zhouchen Lin, Shuicheng Yan, Ju Sun, Yong Yu, and Yi Ma. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):171–184, 2012. 1
- [22] Jiyuan Liu, Xinwang Liu, Yuexiang Yang, Li Liu, Siqi Wang, Weixuan Liang, and Jianguyong Shi. One-pass multi-view clustering for large-scale data. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 12344–12353, 2021. 1
- [23] Suyuan Liu, Siwei Wang, Pei Zhang, Kai Xu, Xinwang Liu, Changwang Zhang, and Feng Gao. Efficient one-pass multi-view subspace clustering with consensus anchors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 7576–7584, 2022. 2, 5
- [24] Yongli Liu, Xiaoqin Zhang, Guiying Tang, and Di Wang. Multi-view subspace clustering based on tensor Schatten-p norm. In *IEEE International Conference on Big Data (Big Data)*, pages 5048–5055. IEEE, 2019. 4
- [25] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, and Shuicheng Yan. Tensor robust principal component analysis: exact recovery of corrupted low-rank tensors via convex optimization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5249–5257, 2016. 1
- [26] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, and Shuicheng Yan. Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4):925–938, 2019. 1
- [27] Gui-Fu Lu, Qin-Ru Yu, Yong Wang, and Ganyi Tang. Hyperlaplacian regularized multi-view subspace clustering with low-rank tensor constraint. *Neural Networks*, 125:214–223, 2020. 3
- [28] Feiping Nie, Jing Li, Xuelong Li, et al. Self-weighted multi-view clustering with multiple graphs. In *Proceedings of the*

- Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 2564–2570, 2017. 1
- [29] Mengjing Sun, Pei Zhang, Siwei Wang, Sihang Zhou, Wenxuan Tu, Xinwang Liu, En Zhu, and Changjian Wang. Scalable multi-view subspace clustering with unified anchors. In *Proceedings of the 29th ACM International Conference on Multimedia*, pages 3528–3536, 2021. 2, 5
- [30] Xiaoli Sun, Rui Zhu, Ming Yang, Xiujun Zhang, and Yuanyan Tang. Sliced sparse gradient induced multi-view subspace clustering via tensorial arctangent rank minimization. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–14, 2022. 2
- [31] Pham Dinh Tao and LT Hoai An. Convex analysis approach to dc programming: theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1):289–355, 1997. 4
- [32] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007. 1, 5
- [33] Chu Wang, Marcello Pelillo, and Kaleem Siddiqi. Dominant set clustering and pooling for multi-view 3d object recognition. *arXiv preprint arXiv:1906.01592*, 2019. 1
- [34] Hao Wang, Yan Yang, and Bing Liu. Gmc: Graph-based multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 32(6):1116–1129, 2019. 5
- [35] Hao Wang, Yan Yang, Bing Liu, and Hamido Fujita. A study of graph-based system for multi-view clustering. *Knowledge-Based Systems*, 163:1009–1019, 2019. 1
- [36] Xiaobo Wang, Xiaojie Guo, Zhen Lei, Changqing Zhang, and Stan Z Li. Exclusivity-consistency regularized multi-view subspace clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 923–931, 2017. 1
- [37] Jianlong Wu, Zhouchen Lin, and Hongbin Zha. Essential tensor learning for multi-view spectral clustering. *IEEE Transactions on Image Processing*, 28(12):5910–5922, 2019. 5
- [38] Wei Xia, Quanxue Gao, Qianqian Wang, Xinbo Gao, Chris Ding, and Dacheng Tao. Tensorized bipartite graph learning for multi-view clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 5
- [39] Wei Xia, Xiangdong Zhang, Quanxue Gao, Xiaochuang Shu, Jungong Han, and Xinbo Gao. Multiview subspace clustering by an enhanced tensor nuclear norm. *IEEE Transactions on Cybernetics*, 52(9):8962–8975, 2021. 2
- [40] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, and Yanyun Qu. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, 126(11):1157–1179, 2018. 1, 2, 3, 5
- [41] Kun Zhan, Chaoxi Niu, Changlu Chen, Feiping Nie, Changqing Zhang, and Yi Yang. Graph structure fusion for multiview clustering. *IEEE Transactions on Knowledge and Data Engineering*, 31(10):1984–1993, 2018. 1
- [42] Kun Zhan, Changqing Zhang, Junpeng Guan, and Junsheng Wang. Graph learning for multiview clustering. *IEEE Transactions on Cybernetics*, 48(10):2887–2895, 2017. 5
- [43] Changqing Zhang, Huazhu Fu, Si Liu, Guangcan Liu, and Xiaochun Cao. Low-rank tensor constrained multiview subspace clustering. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 1582–1590, 2015. 1
- [44] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao. Latent multi-view subspace clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4333–4341, 2017. 1
- [45] Pan Zhou, Canyi Lu, Jiashi Feng, Zhouchen Lin, and Shuicheng Yan. Tensor low-rank representation for data recovery and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(5):1718–1732, 2019. 1