

Article

Robust Tensor Learning for Multi-View Spectral Clustering

Deyan Xie ¹, Zibao Li ¹, Yingkun Sun ^{1,*} and Wei Song ^{2,*} 

¹ School of Science and Information Science, Qingdao Agricultural University, Qingdao 266109, China; dyxie@qau.edu.cn (D.X.); zibaoli@stu.qau.edu.cn (Z.L.)

² Institute of Microscale Optoelectronics, Shenzhen University, Shenzhen 518060, China

* Correspondence: sunyk678@qau.edu.cn (Y.S.); weisong1220@gmail.com (W.S.)

Abstract: Tensor-based multi-view spectral clustering methods are promising in practical clustering applications. However, most of the existing methods adopt the $\ell_{2,1}$ norm to depict the sparsity of the error matrix, and they usually ignore the global structure embedded in each single view, compromising the clustering performance. Here, we design a robust tensor learning method for multi-view spectral clustering (RTL-MSC), which employs the weighted tensor nuclear norm to regularize the essential tensor for exploiting the high-order correlations underlying multiple views and adopts the nuclear norm to constrain each frontal slice of the essential tensor as the block diagonal matrix. Simultaneously, a novel column-wise sparse norm, namely, $\ell_{2,p}$, is defined in RTL-MSC to measure the error tensor, making it sparser than the one derived by the $\ell_{2,1}$ norm. We design an effective optimization algorithm to solve the proposed model. Experiments on three widely used datasets demonstrate the superiority of our method.

Keywords: multi-view spectral clustering; tensor low rank; tensor singular value decomposition



Citation: Xie, D.; Li, Z.; Sun, Y.; Song, W. Robust Tensor Learning for Multi-View Spectral Clustering. *Electronics* **2024**, *13*, 2181. <https://doi.org/10.3390/electronics13112181>

Academic Editors: Yongyong Chen and Yongqiang Tang

Received: 14 April 2024

Revised: 17 May 2024

Accepted: 29 May 2024

Published: 3 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Spectral clustering (SC), which clusters data points using eigenvectors of matrices, is one of the most widely used methods for data clustering. SC and its variants have achieved empirical success in many applications [1–3]. Generally, these existing methods investigate how to improve clustering performance with a fixed affinity matrix [4,5] or how to construct a more robust affinity matrix so as to improve clustering performance [6,7]. When accessing multi-view data, many multi-view spectral clustering (MSC) methods have been developed by pursuing a robust affinity matrix [8–10] or designing an MSC algorithm with a fixed affinity matrix [11–14]. Here, we mainly focus on how to learn a more robust affinity matrix from multiple views.

To learn an affinity matrix from multiple views, Nie et al. [15] developed an MSC method with adaptive neighbors (MSCAN) by performing clustering and local structure learning simultaneously. Based on this, a graph learning method is proposed by introducing a rank constraint on the Laplacian matrix in two steps [16]. Although these methods achieve a promising clustering performance, they directly learn the similarity matrix from original multi-view data, which usually contain noise. To address this problem, the subspace-based methods are widely investigated. For example, Xie et al. [17] propose a method by joint latent representation and similarity learning, which learns the common similarity matrix from the adaptive consistent latent representation. By using the Frobenius norm to measure the reconstruction error, the latent representation of each view, and the consistence between the common similarity matrix and the latent representations, a multi-graph fusion method is developed for MSC [18]. Although these methods learn the class indicator matrix directly, they usually perform k-means on this obtained class indicator matrix to obtain the final predicted label. To deal with this issue, some unified clustering methods are designed. For example, Zhong et al. [19] provide an MSC method by simultaneous consensus graph learning and discretization, which learns the similarity matrix and the discrete cluster

label matrix in a unified framework. Tang et al. [20] made a design to directly capture the discrete clustering indicator matrix from the unified graph. Meanwhile, a multi-view kernel spectral clustering method [21] is proposed by formulating a weighted kernel canonical correlation analysis in a primal–dual optimization setting typical of least squares support vector machines, where the clustering scores are calculated directly.

Different from these traditional shallow methods, some deep MSC methods are also widely investigated. For example, an MSC network [22] is designed by incorporating the local invariance within every single view and the consistency across different views into a unified framework, where the local invariance is defined by a deep metric learning network. Zhao et al. [23] propose a deep MSC via an ensemble model to fuse the similarity graphs from different views by ensemble clustering. They adopt a graph auto-encoder to learn the common indicator matrix directly. Wang et al. [24] use a set of encoders to obtain the latent representation of each view, and then combine the local structure, global structure, and discriminative constraint into a unified framework to learn a shared affinity matrix. Although these deep MSC methods achieved promising clustering methods by exploring deep features, they are usually limited by ignoring the high-order correlations underlying multiple views.

With the aiming of exploring high-order correlations between multiple views, many tensor-based MSC methods are investigated. These methods can be roughly categorized into two types according to the construction of a tensor. One is to construct rank-1 tensors by calculating the outer products of all view-specific feature maps, which incorporate with m th-order tensors of interconnection weights to perform calculations of all views simultaneously, leading to tensor-based high-order couplings and information fusion [25]. The other way is to construct a third-order tensor by merging a set of the representation matrix from each single view. These works mainly focus on how to approximate the tensor low-rank constraint and how to learn the common affinity matrix. One way to depict the tensor low-rank property is by using the Tucker decomposition [26]. Based on this, Chen et al. [27] recovered the latent representation of each view, and adaptively learned the affinity matrix from these representations. Meanwhile, the tensor singular value decomposition (t-SVD)-based tensor nuclear norm (TNN) is also a widely used approximation of tensor low-rank constraint. Based on this, Wu et al. [28] propose an essential tensor learning (ETL) method for MSC (ETL-MSC). Inspired by the low-rank and sparse decomposition, they recovered a latent tensor from the input tensor that was stacked by the transition probability matrix (TPM) of each view. With the obtained tensor, the affinity matrix is computed by calculating the average of all frontal slices of this tensor. By introducing a probability constraint, Zhang et al. [29] improved the clustering performance of ETL-MSC. Although these t-SVD-TNN-based methods achieved a promising clustering performance, they are still limited by the fact that the TNN is only the sum of singular values, causing a poor rank approximation. To address this problem, the nonconvex Laplace function [30], t-SVD-based weighted TNN (WTNN) [31], tailored tensor low-rank norm [32], and weighted tensor Schatten- p norm [33]-based MSC methods are developed. Unlike these methods that pursue the tensor low-rank approximate, Chen et al. [34] committed to learn a common non-negative affinity matrix directly. Unlike this, Wang et al. [35] aimed to emphasize the types of errors that can vary behavior inconsistency in each view by introducing a group ℓ_1 norm to the error tensor.

Inspired by the fact that t-SVD WTNN shrinks different singular values with different weights, leading to a satisfying clustering performance, we propose a novel robust tensor learning method for MSC (RTL-MSC). The proposed RTL-MSC aims to learn the latent tensor from the input transition probability tensor. Then, the affinity matrix, which would be input to the Markov chain SC algorithm to obtain the final clustering result, is computed by each frontal slice of the derived latent tensor. Unlike the aforementioned tensor-based methods, we impose the low-rank constraint on each frontal slice of the latent tensor to emphasize the low-rank property of each latent TPM. Simultaneously, we adopt the $\ell_{2,p}$ norm to regularize the error tensor to depict the noise corresponding to each sample. The contributions of this work can be summarized as follows:

- With a novel integrated strategy, the weighted tensor nuclear norm-based tensor low-rank constraint, the matrix nuclear norm-based low-rank regularization, and the $\ell_{2,p}$ regularization are integrated into a unified framework, where the WTNN regularization depicts the information among different samples and different views, and the matrix nuclear norm regularization makes each frontal slice of the learned tensor approximately block-diagonal, capturing the geometry of each single view. Thus, the affinity matrix calculated from the latent tensor depicts the intrinsic clustering structure of the multi-view data.
- A column-wise sparse norm, namely, $\ell_{2,p}$ norm, is introduced to enhance the robustness of our model. Compared with the $\ell_{2,1}$ norm, the $\ell_{2,p}$ norm, which is invariant, continuous, and differentiable, can be better in restricting the sparsity property of noised samples.

2. Background and Motivation

The Markov chain method is one of the classical clustering approaches, which takes a transition probability matrix as the input to perform a Markov chain-based SC algorithm. Based on this, a robust multi-view spectral clustering (RMSC) method [36] is proposed to recover the consistent TPM by low-rank and sparse decomposition. Let $\{\mathbf{X}^v\}_{v=1}^m$ be the multi-view data including m different features. $\mathbf{X}^v = \{\mathbf{x}_1^v, \mathbf{x}_2^v, \dots, \mathbf{x}_n^v\} \in \mathbb{R}^{d^v \times n}$ ($v = 1, 2, \dots, m$) denotes the features of n samples in the v -th view, and d^v is the feature dimension of the v -th view. RMSC first computes the similarity matrix $\mathbf{A}^v \in \mathbb{R}^{n \times n}$ by $A_{i,j}^v = \exp(-\|\mathbf{x}_i^v - \mathbf{x}_j^v\|_2^2 / \tau)$, and then calculates the TPM $\mathbf{P}^v$ by $\mathbf{P}^v = (\mathbf{D}^v)^{-1} \mathbf{A}^v$, where $\mathbf{D}^v$ is a diagonal matrix with the i -th element $D_{i,i}^v = \sum_{j=1}^n A_{i,j}^v$. Considering that the features in each view might be corrupted by noise, RMSC separates the transition probability matrices $\mathbf{P}^v$ into two parts by the low-rank and sparse decomposition: the common transition probability matrix $\mathbf{P}$, which is shared by all views, and the error matrices $\mathbf{E}^v$, $v = 1, \dots, m$, which encode the noise of the transition probability matrices. By assuming that the common TPM $\mathbf{P}$ tends to be low rank, while the error matrices are sparse, RMSC is formulated as follows:

$$\begin{aligned} \min_{\mathbf{P}, \mathbf{E}^v} & \|\mathbf{P}\|_* + \lambda \sum_{v=1}^m \|\mathbf{E}^v\|_1 \\ \text{s.t. } & \mathbf{P}^v = \mathbf{P} + \mathbf{E}^v, v = 1, 2, \dots, m \\ & \mathbf{P} \geq 0, \mathbf{P}\mathbf{1} = \mathbf{1} \end{aligned} \quad (1)$$

where λ is a non-negative parameter. RMSC explores the consistent information among multiple views while ignoring the unique information of each single view.

To emphasize this issue, a tensor-based TPM learning method, termed as ETL-MSC, is proposed by extending RMSC to a third-order tensor form. ETL-MSC stacks $\mathbf{P}^v$ as the v -th frontal slice of the tensor $\mathcal{P} \in \mathbb{R}^{n \times n \times m}$, and takes $\mathcal{P}$ as the input tensor. By assuming that the essential tensor satisfies tensor low-rank constraint and that the noises in transition probability vectors of noised samples are not sparse while the noisy samples should be sparse, ETL-MSC formulates the following problem:

$$\begin{aligned} \min_{\mathcal{C}, \mathcal{E}} & \|\hat{\mathcal{C}}\|_{\otimes} + \lambda \|\mathcal{E}\|_{2,1} \\ \text{s.t. } & \mathcal{P} = \mathcal{C} + \mathcal{E} \\ & \hat{\mathcal{C}} = \varphi(\mathcal{C}) \in \mathbb{R}^{n \times m \times n} \end{aligned} \quad (2)$$

where the function $\varphi(\mathcal{C})$ rotates the tensor $\mathcal{C} \in \mathbb{R}^{n \times n \times m}$ to obtain a new tensor $\hat{\mathcal{C}}$ with the size of $n \times m \times n$. Correspondingly, $\mathcal{C} = \varphi^{-1}(\hat{\mathcal{C}})$. $\|\hat{\mathcal{C}}\|_{\otimes} = \sum_{j=1}^n \|\hat{\mathcal{C}}_f^{(j)}\|_* = \sum_{j=1}^n \sum_{i=1}^{\min(n,m)} \sigma_i(\hat{\mathcal{C}}_f^{(j)})$ is the t-SVD-based TNN of $\hat{\mathcal{C}}$, $\hat{\mathcal{C}}_f = \text{fft}(\hat{\mathcal{C}}, [], 3)$ represents the discrete Fourier transform (DFT) along the third dimension of $\hat{\mathcal{C}}$, $\hat{\mathcal{C}}_f^{(j)}$ is the j -th frontal slice of $\hat{\mathcal{C}}_f$, and $\sigma_i(\hat{\mathcal{C}}_f^{(j)})$

denotes the i -th largest singular values of $\hat{\mathbf{C}}_f^{(j)}$. The $\ell_{2,1}$ -norm of $\mathcal{E}$ is defined as $\|\mathcal{E}\|_{2,1} \triangleq \|\text{unfold}(\mathcal{E})\|_{2,1}$, resulting in jointly consistent magnitudes for the column of each frontal slice of the error tensor, $\text{unfold}(\mathcal{E}) = [\mathbf{E}^{(1)}; \mathbf{E}^{(2)}; \dots; \mathbf{E}^{(m)}]$.

Two advantages are brought by the t-SVD-based TNN regularization and the rotation operation. First, ETL-MSC could preserve the relationship among all views. Second, the computation complexity can be reduced largely. However, the TNN minimization problem shrinks each singular value with the same parameter. That is, each singular value is considered to be equally important, which contradicts the following facts. As we know, there is a great difference between the non-zero singular values of an arbitrary image. Generally, the larger singular values are considered to be more important than the smaller ones because the larger ones are associated with the salient parts in the image. Thus, to preserve the salient parts, the larger singular values should be shrunk less. Unusually, for a noised image, the larger singular values may carry undesirable information in the existence of noise (such as illumination). In either case, different singular values are suggested to be shrunk differently. To overcome this limit, a t-SVD-based WTNN is proposed and used to formulate the following model to cluster multi-view data [37] by the following:

$$\begin{aligned} \min & \|\hat{\mathcal{C}}\|_{\mathbf{w}, \otimes} + \lambda \|\mathcal{E}\|_{2,1} \\ \text{s.t. } & \mathcal{P} = \mathcal{C} + \mathcal{E} \end{aligned} \quad (3)$$

where $\|\hat{\mathcal{C}}\|_{\mathbf{w}, \otimes} = \sum_{j=1}^n \|\hat{\mathbf{C}}_f^{(j)}\|_{\mathbf{w}, *}$ is the t-SVD-based WTNN of $\hat{\mathcal{C}}$. $\mathbf{w}$ is the weighted vector, with which different singular values could be shrunk differently.

After the essential tensor $\mathcal{C}$ is derived, the common TPM $\mathbf{P}^*$, which is input to the Markov Chain-based SC algorithm to obtain the final clustering result, is often computed by $\mathbf{P}^* = \sum_{v=1}^m \mathbf{C}^{(v)}$. In other words, each frontal slice of the recovered tensor $\mathcal{C}$ is considered to be the intrinsic TPM of each single view, which tends to be a block-diagonal matrix. Such information is generally ignored by the aforementioned tensor methods.

3. The Proposed Method

With the aim of learning a robust TPM in a third-order tensor form, we design a novel robust tensor learning method for multi-view spectral clustering (RTL-MSC). Similar to MGL-WTNN, we also separate the input tensor $\mathcal{P}$ into two parts: the intrinsic tensor $\mathcal{C}$ and the error tensor $\mathcal{E}$. For the intrinsic tensor $\mathcal{C}$, we adopt the t-SVD-based WTNN to regularize its rotated tensor to capture the high-order correlations underlying multiple views. Meanwhile, we introduce the low-rank constraint to each frontal slice of $\mathcal{C}$ to emphasize its block-diagonal structure. Moreover, we utilize a novel $\ell_{2,p}$ norm to emphasize the sparsity of an error tensor. Accordingly, the mathematical model is designed as follows:

$$\begin{aligned} \min_{\mathcal{C}, \mathcal{E}} & \|\hat{\mathcal{C}}\|_{\mathbf{w}, \otimes} + \alpha \sum_{v=1}^m \|\mathbf{E}^{(v)}\|_{2,p} + \beta \sum_{v=1}^m \|\mathbf{C}^{(v)}\|_* \\ \text{s.t. } & \mathcal{P} = \mathcal{C} + \mathcal{E} \\ & \hat{\mathcal{C}} = \varphi(\mathcal{C}) \in \mathbb{R}^{n \times m \times n} \end{aligned} \quad (4)$$

where $\alpha > 0$ and $\beta > 0$ are two balanced parameters. The function $\varphi(\mathcal{C})$ rotates the tensor $\mathcal{C} \in \mathbb{R}^{n \times n \times m}$ to a size of $n \times m \times n$. Correspondingly, $\mathcal{C} = \varphi^{-1}(\hat{\mathcal{C}})$, where $\varphi^{-1}(\bullet)$ represents the inverse operation of $\varphi(\bullet)$. $\|\bullet\|_{2,p}$ denotes the $\ell_{2,p}$ norm, with the definition as follows:

$$\|\mathbf{M}\|_{2,p} = \sum_{i=1}^n \|\mathbf{M}_{:,i}\|_2^p = \sum_{i=1}^n \left(\sum_{j=1}^n (\mathbf{M}_{j,i})^2 \right)^{\frac{p}{2}}. \text{ Here, } 0 < p \leq 1. \text{ Obviously, when } p = 1, \text{ the } \ell_{2,p}$$

norm is reduced to the $\ell_{2,1}$ norm. The solution obtained by the $\ell_{2,p}$ norm is sparser than the one obtained by the $\ell_{2,1}$ norm. Thus, the second term makes it more robust to noised samples. The third term in Equation (4) makes each frontal slice of the essential tensor a block-diagonal matrix, which will captures the intrinsic clustering structure of the multi-view data.

To solve Equation (4), the alternating direction method of multipliers (ADMM) is used. ADMM reformulates the constrained problem as a unconstrained one by constructing an augmented Lagrangian function. By introducing an intermediate tensor $\mathcal{T}$, the Lagrangian function is written as follows:

$$\begin{aligned} \mathcal{L}(\mathcal{C}, \mathcal{E}, \mathcal{T}, \mathcal{Y}, \mathcal{W}) = & \|\hat{\mathcal{T}}\|_{\mathbf{w}, \otimes} + \alpha \sum_{v=1}^m \|\mathbf{E}^{(v)}\|_{2,p} \\ & + \beta \sum_{v=1}^m \|C^{(v)}\|_* + \frac{\mu_1}{2} \|\mathcal{P} - \mathcal{C} - \mathcal{E} + \frac{\mathcal{Y}}{\mu_1}\|_F^2 + \frac{\mu_2}{2} \|\mathcal{T} - \mathcal{C} + \frac{\mathcal{W}}{\mu_2}\|_F^2 \end{aligned} \quad (5)$$

where $\mu_1 > 0$ and $\mu_2 > 0$ are two penalty parameters, and $\mathcal{Y}$ and $\mathcal{W}$ are two Lagrange multipliers. Then, each variable is updated by fixing other ones.

Solving $\mathcal{T}$, by fixing $\mathcal{C}$, $\mathcal{E}$, and $\mathcal{W}$, the sub-problem of $\mathcal{T}$ can be reformulated as follows:

$$\begin{aligned} \min_{\mathcal{T}} & \|\hat{\mathcal{T}}\|_{\mathbf{w}, \otimes} + \frac{\mu_2}{2} \|\mathcal{T} - \mathcal{C} + \frac{\mathcal{W}}{\mu_2}\|_F^2 \\ \Rightarrow \min_{\mathcal{T}} & \|\hat{\mathcal{T}}\|_{\mathbf{w}, \otimes} + \frac{\mu_2}{2} \|\tilde{\mathcal{T}} - \phi(\mathcal{C} - \frac{\mathcal{W}}{\mu_2})\|_F^2 \end{aligned} \quad (6)$$

Denoting the tensor singular value decomposition of $\phi(\mathcal{C} - \frac{\mathcal{W}}{\mu_2})$ as $\mathcal{U} * \mathcal{S} * \mathcal{V}^T$, $\mathcal{S}_f = \text{fft}(\mathcal{S}, [], 3)$, the optimal solution of Equation (6) is $\mathcal{T}^* = \mathcal{U} * \text{ifft}(\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f)) * \mathcal{V}^T$, where $\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f)$ is a tensor stacked by $\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f^{(1)})$, $\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f^{(2)})$, $\dots$, $\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f^{(n)})$, and $\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f^{(k)})$, $k = 1, 2, \dots, n$ is defined as $\Gamma_{\frac{\mathbf{w}}{\mu_1}}(\mathcal{S}_f^{(k)}) = \text{diag}(\theta_1, \theta_2, \dots, \theta_{\min(n,m)})$, and $\theta_i = \max(S_{f_{ii}}^{(k)} - \frac{\omega_i}{\mu_1}, 0)$ [37]. Following [37], the weighted vector $\mathbf{w}$ is adaptively computed by $\mathbf{w} = \sqrt{mn} \times (\mathbf{S}_f^{(k)} + \epsilon)^{-1}$, where ϵ is set to be very small to avoid the singular value of 0.

Solving $\mathcal{E}$, for each view, $\mathbf{E}^{(v)}$, $v = 1, 2, \dots, m$ is independent. By fixing $\mathcal{T}$, $\mathcal{C}$, and $\mathcal{Y}$, the sub-problem of $\mathbf{E}^{(v)}$ is written as follows:

$$\begin{aligned} \min_{\mathbf{E}^{(v)}} & \alpha \|\mathbf{E}^{(v)}\|_{2,p} + \frac{\mu_1}{2} \|\mathbf{P}^{(v)} - \mathbf{C}^{(v)} - \mathbf{E}^{(v)} + \frac{\mathbf{Y}^{(v)}}{\mu_1}\|_F^2 \\ \triangleq & \min_{\mathbf{E}} \frac{\alpha}{\mu_1} \|\mathbf{E}\|_{2,p} + \frac{1}{2} \|\mathbf{E} - \mathbf{Q}\|_F^2 \\ = & \min_{\mathbf{E}} \frac{\alpha}{\mu_1} \sum_{i=1}^n \|\mathbf{E}_{:,i}\|_2^p + \frac{1}{2} \sum_{i=1}^n \|\mathbf{E}_{:,i} - \mathbf{Q}_{:,i}\|_2^2 \\ = & \min_{\mathbf{E}} \sum_{i=1}^n (\frac{\alpha}{\mu_1} \|\mathbf{E}_{:,i}\|_2^p + \frac{1}{2} \|\mathbf{E}_{:,i} - \mathbf{Q}_{:,i}\|_2^2) \end{aligned} \quad (7)$$

where $\mathbf{Q} = [\mathbf{Q}^{(1)}; \mathbf{Q}^{(2)}; \dots; \mathbf{Q}^{(m)}]$, $\mathbf{Q}^{(i)}$ is the i -th frontal slice of the tensor $\mathcal{P} - \mathcal{C} + \frac{\mathcal{Y}}{\mu_1}$. Then, $\mathbf{E}$ is solved column by column. Thus, Equation (7) is rewritten as follows:

$$\min_{\mathbf{E}_{:,i}} \frac{\alpha}{\mu_1} \|\mathbf{E}_{:,i}\|_2^p + \frac{1}{2} \|\mathbf{E}_{:,i} - \mathbf{Q}_{:,i}\|_2^2 \quad (8)$$

here, $\mathbf{E}_{:,i}$ can be treated as a special matrix. Then, we can perform a thin SVD on it. As observed, $\mathbf{E}_{:,i}$ has exactly one singular value, given by the following:

$$\sigma(\mathbf{E}_{:,i}) = \sqrt{\mathbf{E}_{:,i}^T \mathbf{E}_{:,i}} = \|\mathbf{E}_{:,i}\|_2 \quad (9)$$

Thus, Equation (8) can be rewritten as follows:

$$\min_{\mathbf{E}_{:,i}} \frac{\alpha}{\mu_1} (\sigma(\mathbf{E}_{:,i}))^p + \frac{1}{2} \|\mathbf{E}_{:,i} - \mathbf{Q}_{:,i}\|_2^2 \quad (10)$$

According to the work in [38], the solution of Equation (10) is as follows:

$$\mathbf{E}_{:,i} = \mathbf{u}_i \sigma^*(\mathbf{E}_{:,i}) \mathbf{v}_i^T \quad (11)$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the left and right singular vectors of $\mathbf{Q}_{:,i}$, respectively. Additionally,

$$\sigma^*(\mathbf{E}_{:,i}) = \arg \min_{x \geq 0} \tau x^p + \frac{1}{2} (x - \sigma(\mathbf{Q}_{:,i}))^2 \quad (12)$$

where $\frac{\alpha}{\mu_1} \triangleq \tau$, $\sigma(\mathbf{Q}_{:,i}) = \|\mathbf{Q}_{:,i}\|_2$ [39], and Equation (12) can be solved by Lemma 1 in [33]. Correspondingly, $\mathbf{E}_{:,j}$ is derived.

Solving $\mathbf{C}^{(v)}$, by fixing $\mathcal{T}$, $\mathcal{E}$, $\mathcal{W}$, and $\mathcal{Y}$, Equation (5), w.r.t. $\mathcal{C}$, can be written as follows:

$$\arg \min_{\mathbf{C}^{(v)}} \beta \sum_{v=1}^m \|\mathbf{C}^{(v)}\|_* + \frac{\mu_1}{2} \left\| \mathcal{C} - (\mathcal{P} - \mathcal{E} + \frac{\mathcal{Y}}{\mu_1}) \right\|_F^2 + \frac{\mu_2}{2} \left\| \mathcal{C} - (\mathcal{T} + \frac{\mathcal{W}}{\mu_2}) \right\|_F^2 \quad (13)$$

For each $v, v = 1, 2, \dots, m$, $\mathbf{C}^{(v)}$ is independent, so Equation (13) can be rewritten as follows:

$$\begin{aligned} & \arg \min_{\mathbf{C}^{(v)}} \beta \left\| \mathbf{C}^{(v)} \right\|_* + \frac{\mu_1}{2} \left\| \mathbf{C}^{(v)} - \mathbf{M}_1^{(v)} \right\|_F^2 + \frac{\mu_2}{2} \left\| \mathbf{C}^{(v)} - \mathbf{M}_2^{(v)} \right\|_F^2 \\ & = \arg \min_{\mathbf{C}^{(v)}} \beta \left\| \mathbf{C}^{(v)} \right\|_* + \frac{\mu_1 + \mu_2}{2} \left\| \mathbf{C}^{(v)} - \frac{\mu_1 \mathbf{M}_1^{(v)} + \mu_2 \mathbf{M}_2^{(v)}}{\mu_1 + \mu_2} \right\|_F^2 \end{aligned} \quad (14)$$

where $\mathbf{M}_1^{(v)}$ is the v -th frontal slice of $\mathcal{M}_1 = \mathcal{P} - \mathcal{E} + \frac{\mathcal{Y}}{\mu_1}$ and $\mathbf{M}_2^{(v)}$ is the v -th frontal slice of $\mathcal{M}_2 = \mathcal{T} + \frac{\mathcal{W}}{\mu_2}$. According to [40,41], the solution is $\mathbf{C}^{(v)} = \mathbf{U} \Gamma_{\frac{\beta}{\mu_1 + \mu_2}}(\Sigma) \mathbf{V}^T$, where $\mathbf{U} \Sigma \mathbf{V}^T$ is

the singular value decomposition of $\frac{\mu_1 \mathbf{M}_1^{(v)} + \mu_2 \mathbf{M}_2^{(v)}}{\mu_1 + \mu_2}$, and $\Gamma_{\theta}(x) = \max(x - \theta, 0) + \min(x + \theta, 0)$ is the shrinkage operator.

Summarily, Equation (4) can be solved by Algorithm 1.

Algorithm 1 RTL-MSA Algorithm

Input: Multi-view data $\mathbf{X}^j, j = 1, 2, \dots, m, \alpha, \beta$, sample number n , and cluster number c

- 1: Calculate the similarity matrix $\mathbf{A}^j$ of the j -th each view by $A_{i,j}^j = \exp(-\|\mathbf{x}_i^v - \mathbf{x}_j^v\|_2^2 / \tau)$.
 - 2: Compute the degree matrix $\mathbf{D}^j$ with the (i, i) -th element as $D_{i,i}^j = \sum_{k=1}^n A_{i,k}^j, i = 1, 2, \dots, n$, and the other elements are all set to 0.
 - 3: Compute the TPM of the j -th view by $\mathbf{P}^j = (\mathbf{D}^j)^{-1} \mathbf{A}^j$.
 - 4: Construct the input tensor $\mathcal{P}$ by stacking $\mathbf{P}^j$ as its j -th frontal slice.
 - 5: Initialize $\mathcal{T} = \mathcal{C} = \mathcal{E} = \mathcal{Y} = \mathcal{W} = \mathbf{0}; \mu_1 = \mu_2 = 10^{-5}, \eta = 2, \mu_{\max} = 10^{10}, \epsilon = 10^{-7}$
 - 6: **repeat**
 - 7: Update $\mathcal{T}$ by solving Equation (6).
 - 8: Update $\mathcal{C}$ by solving Equation (14).
 - 9: Update the Lagrange multiplier $\mathcal{W}$ using $\mathcal{W} = \mathcal{W} + \mu_2(\mathcal{T} - \mathcal{C})$.
 - 10: Update $\mathcal{E}$ by solving Equation (8).
 - 11: Update the Lagrange multiplier $\mathcal{Y}$ using $\mathcal{Y} = \mathcal{Y} + \mu_1(\mathcal{P} - \mathcal{C} - \mathcal{E})$.
 - 12: Update parameters μ_1 and μ_2 by $\mu_1 = \mu_2 = \min(\eta \mu_1, \mu_{\max})$;
 - 13: Check the convergence conditions:
 - 14: $\|\mathcal{P} - \mathcal{C} - \mathcal{E}\|_{\infty} < \epsilon$.
 - 15: **until** converge
 - 16: Obtain the affinity matrix $\mathbf{P}^*$ as $\mathbf{P}^* = (\sum_{v=1}^m |\mathbf{C}^{(v)}| + |\mathbf{C}^{(v)T}|) / (2m)$;
 - 17: $\mathbf{P}^*$ and c are input to the Markov chain-based SC algorithm.
- Output:** The predicted label.
-

For Algorithm 1, the computation is mainly consumed by the optimization of $\mathcal{T}$, $\mathcal{C}$, and $\mathcal{E}$. To solve $\mathcal{T} \in \mathbb{R}^{n \times m \times n}$, the FFT and inverse FFT of an $n \times m \times n$ tensor along the third dimension and the SVD of each frontal slice of an $n \times m \times n$ tensor are involved, leading to a complexity of $\mathcal{O}(m^2n^2 + mn^2 \log(n))$ at each iteration. $\mathcal{C} \in \mathbb{R}^{n \times n \times m}$ is solved slice by slice. For each frontal slice $\mathcal{C}^{(v)}$, the SVD of an $n \times n$ matrix is calculated, leading to a complexity of $\mathcal{O}(n^3)$. Thus, the solving of $\mathcal{C}$ takes a complexity of $\mathcal{O}(mn^3)$ at each iteration. The $\ell_{2,p}$ norm minimization problem requires a complexity of $\mathcal{O}(mn^2)$ at each iteration. Consequently, by denoting the number of iterations as K , the proposed algorithm takes a complexity of $\mathcal{O}(Kmn^2(m + n + \log(n)))$ in total.

4. Experiments

To evaluate the performance of the proposed methods, we conduct experiments on three widely used datasets. All experiments are performed in Matlab R2021a.

4.1. Competitors

The proposed methods are compared with the spectral clustering method and six multi-view state-of-the-art clustering methods. (1) SC_{best} is achieved by performing the spectral clustering on each single view and obtaining the best result among all views. (2) RMSC learns a common low-rank transition probability matrix via low-rank and sparse matrix decomposition. (3) MSCAN incorporates the common similarity matrix learning and the predicted label matrix learning into a unified framework, which allocates weight for each view automatically without any additional parameters. (4) ETL-MSC recovers an essential tensor via a tensor RPCA, which utilizes the t-SVD-based TNN to approximate the tensor low rank. (5) ATPML-MSC [34] learns a consistent TPM directly under the tensor RPCA framework. The obtained TPM is non-negative and symmetric without any postprocessing. (6) MGL-WTNN [37] recovers a low-rank tensor by using the t-SVD-based WTNN to approximate the tensor low rank.

4.2. Datasets

We take Flower (<https://www.robots.ox.ac.uk/~vgg/data/flowers>), Coil20 (<https://www.cs.columbia.edu/CAVE/software/softlib>), and NUS-WIDE (<https://lms.comp.nus.edu.sg/wp-content/uploads/2019/research/nuswide/NUS-WIDE.html>) to evaluate the clustering performance. The Flower dataset consists of 1360 samples from 17 different flower categories. These images have large scale, pose, and light variations. Following [34], three different features, 1360d-color, 1360d-texture, and 1360d-shape, are extracted. Coil20 is an object dataset containing 1440 images from 20 object categories. We employ the 4096d intensity features and 3304d LBP features as the two views. NUS-WIDE is a real-world web image dataset from the National University of Singapore. We adopt six types of low-level features extracted from 2400 images of 12 concepts, including a 64-D color histogram, 144-D color correlogram, 75-D edge direction histogram, 128-D wavelet texture, 225-D block-wise color moments, and 500-D bag of words based on SIFT descriptions.

4.3. Experimental Process

To quantitatively evaluate the performance of all these clustering methods, we adopt five widely used metrics, accuracy (ACC), normalized mutual information (NMI), Purity, F-score, and average entropy (AVE). For ACC, NMI, Purity, and F-score, larger values represent better a clustering performance. For AVE, lower values indicate better a clustering performance.

For the compared methods, we download their codes and perform experiments following the relative papers. For RMSC, ETL-MSC, and MGL-WTNN, one parameter is involved and is selected from $[0.0001, 1]$. Two parameters in ATPML-MSC are changed in the range of $[0.0001, 0.5]$, respectively. For each method, we fix the parameters to the ones corresponding to the better accuracy and repeat experiments ten times to obtain the mean value and standard deviation.

4.4. Experimental Results and Analysis

The clustering results on all the three datasets are shown in Tables 1–3, where the best results are highlighted in bold and the second ones are underlined.

Table 1. Clustering results (mean $\pm$ standard deviation) on Flower.

Methods	ACC	NMI	Purity	F-score	AVE
SC _{best}	0.3737 $\pm$ 0.020	0.3975 $\pm$ 0.017	0.3916 $\pm$ 0.017	0.2637 $\pm$ 0.008	2.4604 $\pm$ 0.027
RMSC	0.3796 $\pm$ 0.015	0.4167 $\pm$ 0.008	0.3982 $\pm$ 0.015	0.2685 $\pm$ 0.008	2.3890 $\pm$ 0.032
MSCAN	0.4585 $\pm$ 0.001	0.5070 $\pm$ 0.001	0.4852 $\pm$ 0.002	0.3164 $\pm$ 0.002	2.1405 $\pm$ 0.009
ETL-MSC	0.9108 $\pm$ 0.034	0.9114 $\pm$ 0.016	0.9167 $\pm$ 0.028	0.8626 $\pm$ 0.033	0.3693 $\pm$ 0.068
ATPML-MSC	0.9772 $\pm$ 0.001	0.9661 $\pm$ 0.002	0.9772 $\pm$ 0.001	0.9554 $\pm$ 0.004	0.1395 $\pm$ 0.010
MGL-WTNN	<u>0.9811 $\pm$ 0.001</u>	<u>0.9703 $\pm$ 0.001</u>	<u>0.9811 $\pm$ 0.001</u>	<u>0.9628 $\pm$ 0.001</u>	<u>0.1215 $\pm$ 0.003</u>
RTL-WTNN	0.9874 $\pm$ 0.001	0.9800 $\pm$ 0.001	0.9874 $\pm$ 0.001	0.9753 $\pm$ 0.001	0.0818 $\pm$ 0.003

Table 2. Clustering results (mean $\pm$ standard deviation) on Coil20.

Methods	ACC	NMI	Purity	F-score	AVE
SC _{best}	0.6886 $\pm$ 0.026	0.8131 $\pm$ 0.010	0.7276 $\pm$ 0.019	0.6609 $\pm$ 0.025	0.8388 $\pm$ 0.048
RMSC	0.7026 $\pm$ 0.020	0.8022 $\pm$ 0.006	0.7101 $\pm$ 0.011	0.6628 $\pm$ 0.015	0.8668 $\pm$ 0.031
MSCAN	0.8061 $\pm$ 0.022	0.9299 $\pm$ 0.005	0.8481 $\pm$ 0.016	0.7441 $\pm$ 0.030	0.4266 $\pm$ 0.044
ETL-MSC	0.8730 $\pm$ 0.021	0.9204 $\pm$ 0.009	0.8811 $\pm$ 0.015	0.8510 $\pm$ 0.020	0.3518 $\pm$ 0.040
ATPML-MSC	<u>0.9513 $\pm$ 0.038</u>	<u>0.9795 $\pm$ 0.016</u>	<u>0.9616 $\pm$ 0.031</u>	<u>0.9481 $\pm$ 0.041</u>	<u>0.0990 $\pm$ 0.078</u>
MGL-WTNN	0.8853 $\pm$ 0.023	0.9300 $\pm$ 0.006	0.8919 $\pm$ 0.015	0.8645 $\pm$ 0.016	0.3098 $\pm$ 0.032
RTL-WTNN	0.9769 $\pm$ 0.028	0.9866 $\pm$ 0.011	0.9806 $\pm$ 0.020	0.9695 $\pm$ 0.027	0.0627 $\pm$ 0.053

Table 3. Clustering results (mean $\pm$ standard deviation) on NUS-WIDE.

Methods	ACC	NMI	Purity	F-score	AVE
SC _{best}	0.2338 $\pm$ 0.008	0.1095 $\pm$ 0.004	0.2457 $\pm$ 0.007	0.1361 $\pm$ 0.002	3.1950 $\pm$ 0.016
RMSC	0.2962 $\pm$ 0.004	0.1556 $\pm$ 0.004	0.3035 $\pm$ 0.004	0.1758 $\pm$ 0.003	3.029 $\pm$ 0.013
MSCAN	0.2493 $\pm$ 0.005	0.1972 $\pm$ 0.004	0.2643 $\pm$ 0.005	0.1739 $\pm$ 0.001	3.0376 $\pm$ 0.012
ETL-MSC	0.6906 $\pm$ 0.001	0.7007 $\pm$ 0.001	0.6927 $\pm$ 0.001	0.6243 $\pm$ 0.001	1.078 $\pm$ 0.004
ATPML-MSC	<u>0.9286 $\pm$ 0.001</u>	<u>0.8844 $\pm$ 0.001</u>	<u>0.9286 $\pm$ 0.001</u>	<u>0.8658 $\pm$ 0.002</u>	<u>0.4062 $\pm$ 0.005</u>
MGL-WTNN	0.8787 $\pm$ 0.001	0.7774 $\pm$ 0.001	0.8787 $\pm$ 0.001	0.7791 $\pm$ 0.001	0.7990 $\pm$ 0.003
RTL-WTNN	0.9441 $\pm$ 0.001	0.8976 $\pm$ 0.001	0.9441 $\pm$ 0.001	0.8933 $\pm$ 0.001	0.3650 $\pm$ 0.001

As can be seen from Tables 1–3, the proposed RTL-MSC method almost outperforms the other methods including the recently proposed ATPML-MSC and MGL-WTNN. For example, RTL-WTNN improves by about 2.56%, 1.90%, and 2.14% for ACC, Purity, and F-score over the most competitive method ATPML-MSC on the Coil20 dataset, respectively. This represents that the affinity matrix is more effective for a multi-view spectral clustering task. Compared with MGL-WTNN, RTL-MSC improves by 11.16%, 5.66%, and 8.87% in terms of ACC, NMI, and Purity on the Coil20 dataset. This means that the global geometric structure in each single view and the $\ell_{2,p}$ norm-based error constraint are useful for multi-view clustering. ATPML-MSC outperforms ETL-MSC on all these three datasets. This is mainly because the affinity matrix is learned in a one-step process with a probability constraint. MGL-WTNN outperforms ETL-MSC on all these three datasets for almost all metrics. This demonstrates that shrinking each singular value with an adaptive learned weight improves the clustering performance efficiently. On the NUS-WIDE dataset, the SC based on a single view, RMSC and MSCAN, obtains an unsatisfied clustering performance. The major reason may be that the intrinsic structure of each view and the noise are not well depicted by these three methods.

There are three parameters, α , β , and p , involved in Equation (4), which directly affect the performance of the proposed RTL-MSC. To investigate the influence of α and β , we set $p = 1$ and select α and β from $[0.01, 0.5]$ and $[0.0001, 0.01]$, respectively. We report the ACC and NMI versus different combinations of α and β on the Flower and Coil20 datasets in Figure 1. It is shown that when α changes in $[0.04, 0.5]$, both ACC and NMI fluctuate slightly via β , and RTL-WTNN achieves promising clustering results on the Flower and

Coil20 datasets. This means that the proposed RTL-WTNN is insensitive to α and β when they are changed in a suitable range.

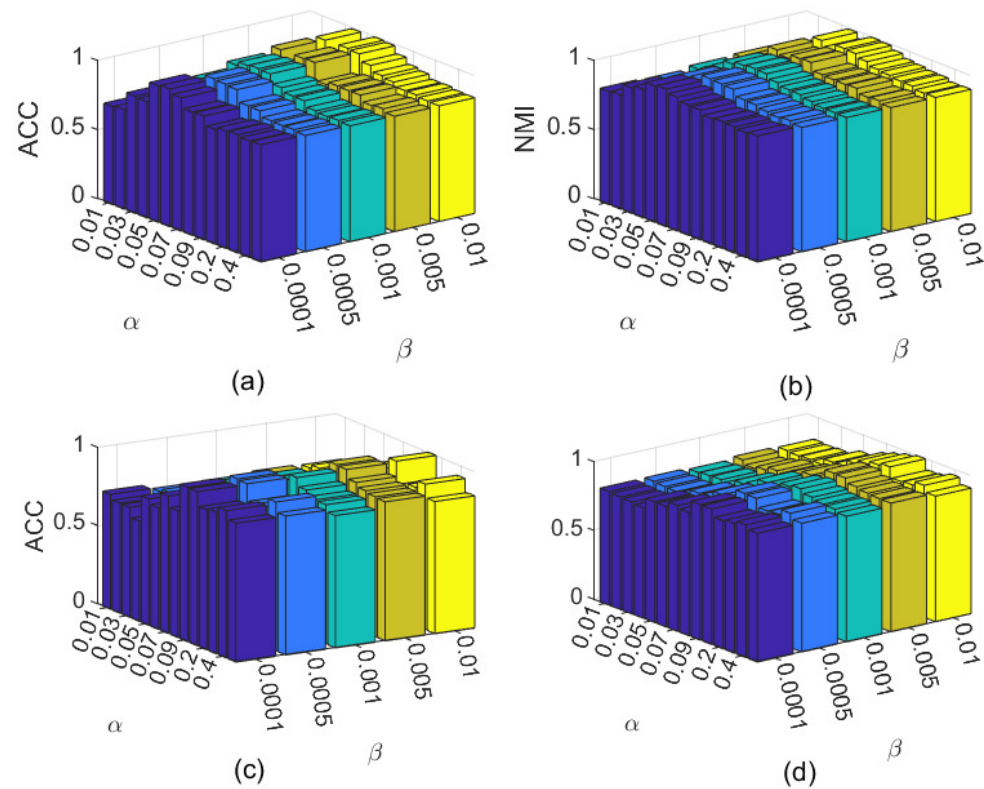


Figure 1. Parameter analysis on (a,b) Flower and (c,d) Coil20 datasets.

Moreover, we fix α and β and change p in the range of $[0.9, 1]$ to investigate the influence of the $\ell_{2,p}$ norm. As shown in Figure 2, the highest values of ACC, NMI, and Purity is achieved at $p = 0.98$, which illustrates that the $\ell_{2,p}$ norm improves the clustering performance.

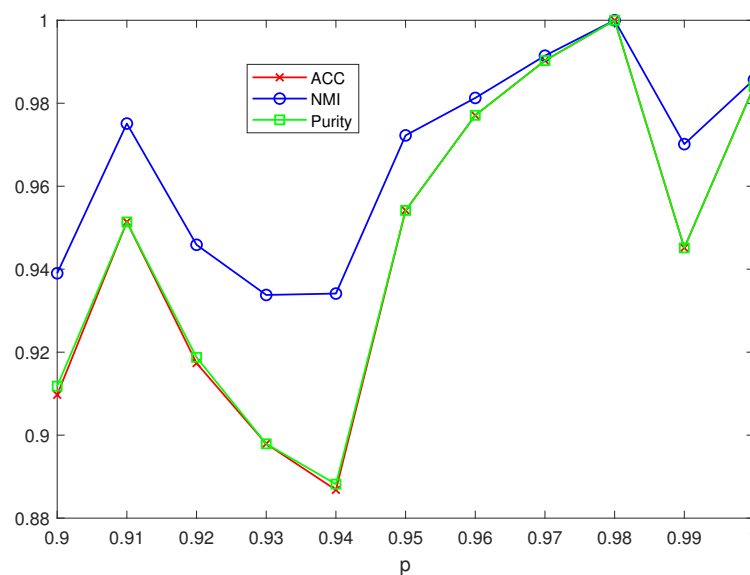


Figure 2. Clustering result change versus p on the Coil20 dataset.

4.5. Ablation Study

To investigate the effect of exploiting the low-rank structure of the each frontal slice of the essential tensor and the $\ell_{2,p}$ norm-based error constraint, we perform an ablation study.

First, we remove the third term in Equation (4) and mark the derived model as RTL-WNN-1. Then, we also perform experiments by using the $\ell_{2,1}$ norm to regularize the error tensor, which is termed as RTL-WTNN-2. The results are shown in Figure 3. It can be seen that RTL-WTNN outperforms RTL-WTNN-1 and RTL-WTNN-2 on all the three datasets.

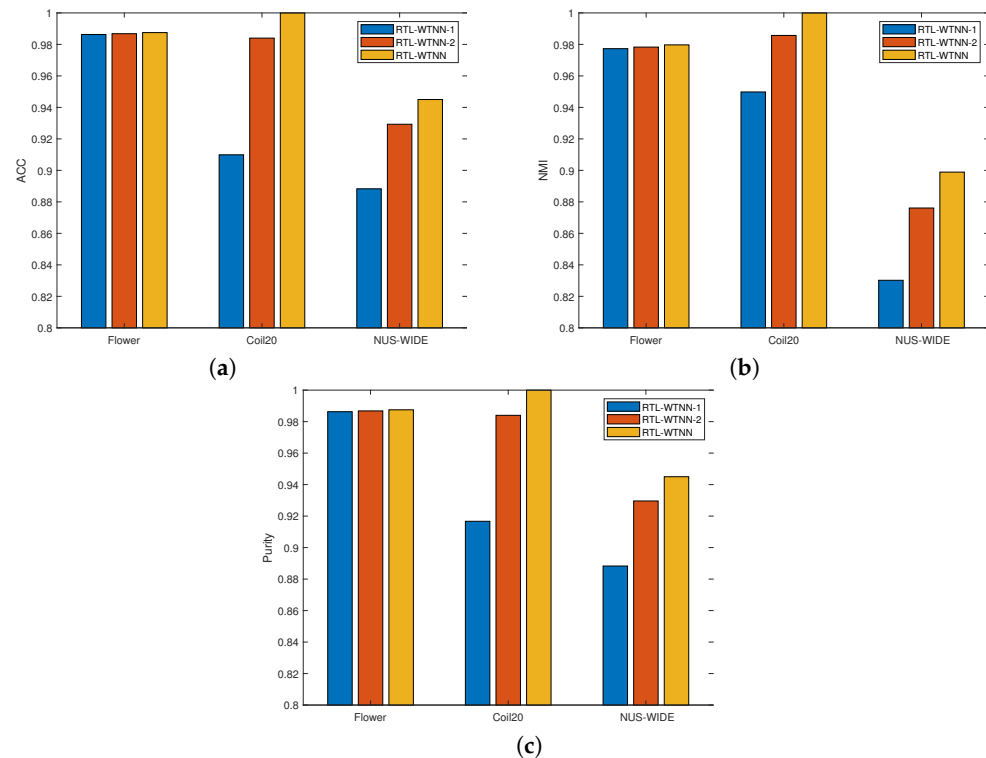


Figure 3. Ablation study on all the three datasets: (a) ACC, (b) NMI, and (c) Purity.

5. Conclusions

We develop a novel tensor learning method for multi-view spectral clustering, which can depict well the noise embedded in each view by using the $\ell_{2,p}$ norm to constrain the error tensor. Additionally, by adopting the tensor low-rank constraint and low-rank regularization of each frontal slice of the essential tensor, the common affinity matrix computed by the essential tensor can capture well the intrinsic clustering structure of the multi-view data. The experiments on three widely used datasets exhibit that the proposed RTL-MSC improves by about 1.02–11.16%, 1.39–12.02%, 1.02–8.87%, and 1.99–11.42% in terms of ACC, NMI, Purity, and F-score over the baseline MGL-WTNN, which illustrates the superiority of the proposed model. We plan to investigate a fast multi-view spectral clustering algorithm for clustering big data in our next work.

Author Contributions: Software, Z.L.; Writing—original draft, D.X.; Writing—review & editing, Y.S.; Supervision, W.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China grant number 62175159, National Key Research and Development Program of China grant number 2023YFF0715300, Natural Science Foundation of Shandong Province grant number ZR202102180986, Natural Science Foundation of Guangdong Province grant number 2023A1515012888, Qingdao Agricultural University grant number 665/1120051, and Qingchuang Talents Induction Program of Shandong Higher Education Institution grant number 018-1622001. The APC was funded by Natural Science Foundation of Shandong Province grant number ZR202102180986, and Natural Science Foundation of Shandong Province grant number ZR202102180986.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Shi, J.; Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 888–905.
- Ng, A.; Jordan, M.; Weiss, Y. On spectral clustering: Analysis and an algorithm. *Adv. Neural Inf. Process. Syst.* **2001**, *14*, 849–856.
- Fan, J.; Tu, Y.; Zhang, Z.; Zhao, M.; Zhang, H. A simple approach to automated spectral clustering. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 9907–9921.
- Yan, Y.; Shen, C.; Wang, H. Efficient semidefinite spectral clustering via Lagrange duality. *IEEE Trans. Image Process.* **2014**, *23*, 3522–3534. [[CrossRef](#)] [[PubMed](#)]
- Meila, M.; Shi, J. Learning segmentation by random walks. *Adv. Neural Inf. Process. Syst.* **2000**, *13*, 837–843.
- Roweis, S.T.; Saul, L.K. Nonlinear dimensionality reduction by locally linear embedding. *Science* **2000**, *290*, 2323–2326. [[CrossRef](#)] [[PubMed](#)]
- Elhamifar, E.; Vidal, R. Sparse subspace clustering: Algorithm, theory, and applications. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2765–2781. [[CrossRef](#)] [[PubMed](#)]
- Nie, F.; Cai, G.; Li, X. Multi-view clustering and semi-supervised classification with adaptive neighbours. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
- Zhou, S.; Liu, X.; Liu, J.; Guo, X.; Zhao, Y.; Zhu, E.; Zhai, Y.; Yin, J.; Gao, W. Multi-view spectral clustering with optimal neighborhood Laplacian matrix. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 6965–6972.
- Alzate, C.; Suykens, J.A. Multiway spectral clustering with out-of-sample extensions through weighted kernel PCA. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *32*, 335–347. [[CrossRef](#)] [[PubMed](#)]
- Xia, T.; Tao, D.; Mei, T.; Zhang, Y. Multiview spectral embedding. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* **2010**, *40*, 1438–1446.
- Lu, C.; Yan, S.; Lin, Z. Convex sparse spectral clustering: Single-view to multi-view. *IEEE Trans. Image Process.* **2016**, *25*, 2833–2843. [[CrossRef](#)]
- El Hajjar, S.; Dornaika, F.; Abdallah, F. Multi-view spectral clustering via constrained nonnegative embedding. *Inf. Fusion* **2022**, *78*, 209–217. [[CrossRef](#)]
- Zong, L.; Zhang, X.; Liu, X.; Yu, H. Weighted multi-view spectral clustering based on spectral perturbation. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
- Nie, F.; Li, J.; Li, X. Parameter-free auto-weighted multiple graph learning: A framework for multiview clustering and semi-supervised classification. In Proceedings of the IJCAI, New York, NY, USA, 9–15 July 2016; Volume 9, pp. 1881–1887.
- Zhan, K.; Zhang, C.; Guan, J.; Wang, J. Graph learning for multiview clustering. *IEEE Trans. Cybern.* **2017**, *48*, 2887–2895. [[CrossRef](#)] [[PubMed](#)]
- Xie, D.; Zhang, X.; Gao, Q.; Han, J.; Xiao, S.; Gao, X. Multiview clustering by joint latent representation and similarity learning. *IEEE Trans. Cybern.* **2019**, *50*, 4848–4854. [[CrossRef](#)] [[PubMed](#)]
- Kang, Z.; Shi, G.; Huang, S.; Chen, W.; Pu, X.; Zhou, J.T.; Xu, Z. Multi-graph fusion for multi-view spectral clustering. *Knowl.-Based Syst.* **2020**, *189*, 105102. [[CrossRef](#)]
- Zhong, G.; Shu, T.; Huang, G.; Yan, X. Multi-view spectral clustering by simultaneous consensus graph learning and discretization. *Knowl.-Based Syst.* **2022**, *235*, 107632. [[CrossRef](#)]
- Tang, C.; Li, Z.; Wang, J.; Liu, X.; Zhang, W.; Zhu, E. Unified one-step multi-view spectral clustering. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 6449–6460. [[CrossRef](#)]
- Houthuys, L.; Langone, R.; Suykens, J.A. Multi-view kernel spectral clustering. *Inf. Fusion* **2018**, *44*, 46–56. [[CrossRef](#)]
- Huang, Z.; Zhou, J.T.; Peng, X.; Zhang, C.; Zhu, H.; Lv, J. Multi-view Spectral Clustering Network. In Proceedings of the IJCAI, Macao, 10–16 August 2019; Volume 2, p. 4.
- Zhao, M.; Yang, W.; Nie, F. Deep multi-view spectral clustering via ensemble. *Pattern Recognit.* **2023**, *144*, 109836. [[CrossRef](#)]
- Wang, Q.; Cheng, J.; Gao, Q.; Zhao, G.; Jiao, L. Deep multi-view subspace clustering with unified and discriminative learning. *IEEE Trans. Multimed.* **2020**, *23*, 3483–3493. [[CrossRef](#)]
- Tao, Q.; Tonin, F.; Patrinos, P.; Suykens, J.A. Tensor-based multi-view spectral clustering via shared latent space. *Information Fusion* **2024**, *108*, 102405. [[CrossRef](#)]
- Goldfarb, D.; Qin, Z. Robust low-rank tensor recovery: Models and algorithms. *SIAM J. Matrix Anal. Appl.* **2014**, *35*, 225–253. [[CrossRef](#)]
- Chen, Y.; Xiao, X.; Peng, C.; Lu, G.; Zhou, Y. Low-rank tensor graph learning for multi-view subspace clustering. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *32*, 92–104. [[CrossRef](#)]
- Wu, J.; Lin, Z.; Zha, H. Essential tensor learning for multi-view spectral clustering. *IEEE Trans. Image Process.* **2019**, *28*, 5910–5922. [[CrossRef](#)] [[PubMed](#)]

29. Zhang, Y.; Yang, W.; Liu, B.; Ke, G.; Pan, Y.; Yin, J. Multi-view spectral clustering via tensor-SVD decomposition. In Proceedings of the 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), Boston, MA, USA, 6–8 November 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 493–497.
30. Wang, S.; Chen, Y.; Cen, Y.; Zhang, L.; Wang, H.; Voronin, V. Nonconvex low-rank and sparse tensor representation for multi-view subspace clustering. *Appl. Intell.* **2022**, *52*, 14651–14664. [[CrossRef](#)]
31. Gao, Q.; Xia, W.; Wan, Z.; Xie, D.; Zhang, P. Tensor-SVD based graph learning for multi-view subspace clustering. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 3930–3937.
32. Jia, Y.; Liu, H.; Hou, J.; Kwong, S.; Zhang, Q. Multi-view spectral clustering tailored tensor low-rank representation. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 4784–4797. [[CrossRef](#)]
33. Gao, Q.; Zhang, P.; Xia, W.; Xie, D.; Gao, X.; Tao, D. Enhanced tensor RPCA and its application. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 2133–2140. [[CrossRef](#)] [[PubMed](#)]
34. Chen, Y.; Xiao, X.; Hua, Z.; Zhou, Y. Adaptive transition probability matrix learning for multiview spectral clustering. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 4712–4726. [[CrossRef](#)]
35. Wang, S.; Chen, Y.; Jin, Y.; Cen, Y.; Li, Y.; Zhang, L. Error-robust low-rank tensor approximation for multi-view clustering. *Knowl.-Based Syst.* **2021**, *215*, 106745. [[CrossRef](#)]
36. Xia, R.; Pan, Y.; Du, L.; Yin, J. Robust multi-view spectral clustering via low-rank and sparse decomposition. In Proceedings of the AAAI Conference on Artificial Intelligence, Québec City, QC, Canada, 27–31 July 2014; Volume 28.
37. Xie, D.; Gao, Q.; Deng, S.; Yang, X.; Gao, X. Multiple graphs learning with a new weighted tensor nuclear norm. *Neural Netw.* **2021**, *133*, 57–68. [[CrossRef](#)]
38. Yang, M.; Luo, Q.; Li, W.; Xiao, M. Multiview clustering of images with tensor rank minimization via nonconvex approach. *SIAM J. Imaging Sci.* **2020**, *13*, 2361–2392. [[CrossRef](#)]
39. Horn, R.A.; Johnson, C.R. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 2012.
40. Liu, G.; Lin, Z.; Yu, Y. Robust subspace segmentation by low-rank representation. In Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 21–24 June 2010; pp. 663–670.
41. Cai, J.F.; Candès, E.J.; Shen, Z. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **2010**, *20*, 1956–1982. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.