Nonconvex Low-Rank Tensor Representation for Multi-View Subspace Clustering With Insufficient Observed Samples

Meng Ding, Jing-Hua Yang, Xi-Le Zhao, Jie Zhang, and Michael K. Ng, *Senior Member, IEEE*

***Abstract*—Multi-view subspace clustering (MVSC) separates the data with multiple views into multiple clusters, and each cluster corresponds to one certain subspace. Existing tensor-based MVSC methods construct self-representation subspace coefficient matrices of all views as a tensor, and introduce the tensor nuclear norm (TNN) to capture the complementary information hidden in different views. The key assumption is that the data samples of each subspace must be sufficient for subspace representation. This work proposes a nonconvex latent transformed low-rank tensor representation framework for MVSC. To deal with the insufficient sample problem, we study the latent low-rank representation in the multi-view case to supplement underlying observed samples. Moreover, we propose to use data-driven transformed TNN (TTNN), resulting from the intrinsic structure of multi-view samples, to preserve the consensus and complementary information in the transformed domain. Meanwhile, the proposed unified nonconvex low-rank tensor representation framework can better learn the high correlation among different views. To resolve the proposed nonconvex optimization model, we propose an effective algorithm under the framework of the alternating direction method of multipliers and theoretically prove that the iteration sequences converge to the critical point. Experiments on various datasets showcase outstanding performance.**

***Index Terms*—Multi-view subspace clustering, insufficient data sampling, nonconvex low-rank tensor representation, transformed tensor nuclear norm, convergence analysis.**

Received 9 December 2023; revised 1 January 2025; accepted 22 March 2025. Date of publication 26 March 2025; date of current version 1 May 2025. The work was supported in part by the National Natural Science Foundation of China under Grant 12201522, Grant 12371456, and Grant 12401605, in part by the Natural Science Foundation of Sichuan Province of China under Grant 2024NSFSC1389 and Grant 2024NSFSC1467, in part by Sichuan Science and Technology Program under Grant 2024NSFJQ0038, Grant 2023ZYD0007, and Grant 2024NSFSC0038, in part by Postdoctoral Fellowship Program of CPSF under Grant GZC20232198, in part by China Postdoctoral Science Foundation under Grant 2024M752661, in part by the National Key Research and Development Program of China under Grant 2024YFE0202900, in part by HKRGC GRF under Grant 17300021 and Grant C7004-21GF, in part by Joint NSFC-RGC under Grant N-HKU76921, and in part by Fundamental Research Funds for the Central Universities under Grant 2682024CX017. Recommended for acceptance by G. Wang *(Corresponding authors: Jing-Hua Yang; Michael K. Ng.)*

Meng Ding is with the School of Mathematics, Southwest Jiaotong University, Chengdu 611756, China (e-mail: dingmeng56@163.com).

Jing-Hua Yang is with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China (e-mail: yangjinghua110@126.com).

Xi-Le Zhao is with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: xlzhao122003@163.com).

Jie Zhang is with the Department of Mathematics, The University of Hong Kong, Hong Kong, SAR, China (e-mail: zj199607@connect.hku.hk).

Michael K. Ng is with the Department of Mathematics, Hong Kong Baptist University, Hong Kong, SAR, China (e-mail: michael-ng@hkbu.edu.hk).

This article has supplementary downloadable material available at https://doi.org/10.1109/TKDE.2025.3555043, provided by the authors.

Digital Object Identifier 10.1109/TKDE.2025.3555043

I. Introduction

MULTI-VIEW data [1], collected by diverse feature extractors, has multiple distinct features. One representative example is that an image can be described using various features, e.g., texture, edge, and color. Another example is that one can represent web pages using text or hyperlinks to transfer information. Each feature represents the particular information w.r.t. the corresponding view, different features can represent more complementary and discriminative information to each other, compared with the single-view data. Combining multiple features for data analysis has become a fundamental and popular tool for various tasks [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], e.g., clustering, image categorization, and segmentation. This work studies the multi-view clustering task, please refer to the comprehensive reviews [15], [16] and references therein.

Multi-view clustering (MVC) aims to classify multi-view data with similar structures into different groups by combining the complementary and consensus information embedded in multiple views. MVC has been widely studied in a number of works, e.g., [5], [10], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]. Among them, subspace clustering [28], [29], [30] attracts much attention and achieves promising performance. Subspace clustering assumes that the data samples lie in a union of low-dimensional subspaces. By exploring the correlation embedded in data, subspace clustering first learns the affinity matrix and then applies the spectral clustering technique to the learned matrix to obtain clustering results. Sparse subspace clustering (SSC) [29] found a sparse coefficient matrix of the data points using self-representation. Low-rank representation (LRR) [28] aimed to obtain the representation matrix with the low-rank structure. Inspired by the impressive performance, the subspace clustering has been extended from single-view data to the multi-view data case.

Multi-view subspace clustering (MVSC) methods construct a unified affinity matrix using the self-representation coefficients corresponding to multiple views to perform clustering. The work [9] performed the subspace clustering on the learned representations of all views simultaneously. The work [31]

combined the smoothness and diversity terms to enhance the complementary information between different views. Besides, the work [25] integrated subspace representations into a shared multi-view representation to encode the high correlation among all views. These methods can be regarded as matrix-based approaches and just capture the pairwise correlations among different views [10], [32].

Recently, some works stack the multiple learned representation coefficients into a third-order tensor [33], [34], [35], [36], [37] and then impose the structured constraint to fully exploit the inner correlation among each view. The work [10] first proposed the LRR-based multi-view subspace clustering strategy. They constructed a tensor using the learned coefficient matrices and imposed the Tucker tensor decomposition to explore the low-tensor-rank prior. Then the work [24] utilized the tensor nuclear norm (TNN)-based low-tensor-rank constraint [38] to capture the complementary information along all views. Along this line, a number of tricks were proposed to enhance the ability of the stacked tensor to explore the complementary information, such as graph learning [39], [40], kernel presentation [41], [42], and hyper-Laplacian regularizer [43]. The aforementioned LRR-based tensor low-rank methods can well capture the complementary and consensus information within different views and obtain impressive clustering performance. However, these LRR-based tensor methods are based on a basic assumption, in which the data samples in each individual view should be sufficient to obtain robust clustering performance [24], [30]. Therefore, it would depress the clustering performance when the data sampling is insufficient or/and noisy. Existing multi-view clustering methods do not focus on the insufficient observed data samples. The insufficient data problem has been studied in the single-view clustering [44], [45], [46], [47] and plays an important role in many applications, e.g., image fusion [48] and fashion compatibility prediction [49]. However, these methods [44], [45], [46], [47] are designed for the single-view data. Meanwhile, all TNN-based tensor multi-view clustering methods transform the representation tensor into the fixed Fourier domain to preserve the correlation among multiple views and different samples. Nonetheless, they neglect the intrinsic data structure and may not be suitable for various multi-view data in practice.

In this work, to tackle these issues, we propose a new latent transformed low-rank tensor representation framework for MVSC. Motivated by the work [30], we introduce the latent low-rank representation to the multi-view case. Under the framework of [30], the vth view data samples $\boldsymbol{X}^{(v)}$ can be approximately expressed as two terms, i.e., the principle component (denoted as $\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}$) and salient component (denoted as $\boldsymbol{L}^{(v)}\boldsymbol{X}^{(v)}$). Since the samples in different views belong to the same cluster, the subspace representations $\boldsymbol{Z}^{(v)}$ share the highly similar correlation with each other. Hence, we stack all representations together as a third-order tensor and impose the low-tensor-rank constraint to capture the high correlation and complementary information hidden in multiple views. To better capture the high correlation, we propose a unified nonconvex data-driven transformed low-rank tensor representation. Meanwhile, we impose the low-rank matrix prior to salient representations $\boldsymbol{L}^{(v)}$ to explore the relationship among samples within each individual view.

We summarize the main contributions as follows:

1) The LRR-based MVSC methods require sufficient data samples to represent the subspace information. While this assumption may be invalid, and the performance will be depressed in real scenarios. In this work, we introduce the latent low-rank representation into the multi-view clustering. Then we stack the multiple subspace representations as a tensor. For the stacked tensor and salient representations, we employ the low-rank tensor/matrix priors to the correlation between and within different view samples, respectively. For the first time to our knowledge, we attempt to solve the problem of insufficient observed data sampling in MVSC.

2) We propose a unified nonconvex transformed low-rank tensor representation framework to better capture the consensus and complementary information. TNN-based tensor methods preserve the representation coefficients in the fixed Fourier domain, while the proposed unified transformed low-rank tensor representation can adaptively preserve the coefficients in transformed according to the data's intrinsic structure. Moreover, we adopt a unified nonconvex surrogate to exploit the difference of singular values of the representation tensor for better learning the subspace representations for all views.

3) We resolve the formulated optimization problem using the alternating direction method of multipliers (ADMM) framework. Meanwhile, we give an approximate solver for the proximal problem related to the proposed transformed nonconvex tensor representation. Each subproblem admits the closed-form solution, hence our algorithm can be easily employed. Although the proposed optimization problem is nonconvex, we establish the theoretical convergence of our algorithm. Numerical experiments showcase the outstanding performance of the proposed method.

The rest of this paper is organized as follows. Section II presents the notations and preliminaries used in this work. Section III reviews related subspace clustering and MVSC works. Section IV proposes the nonconvex data-driven transformed MVSC model and ADMM-based algorithm. Section V performs various experiments to demonstrate the performance of our method. Section VI concludes this work.

II. Notations and Preliminaries

A. Notations

The symbol x (or X) represents the scalar, $\boldsymbol{x}$ denotes the vector, $\boldsymbol{X}$ represents the matrix, and $\mathcal{X}$ denotes the tensor. For a third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, $x_{i,j,k}$ denotes the (i,j,k)th element. The symbol $\mathcal{X}(:,:,k)$ or $\boldsymbol{X}^{(k)}$ denotes the kth frontal slice of $\mathcal{X}$. The jth column of a matrix $\boldsymbol{X}$ with size $n_1 \times n_2$ is denoted as $\boldsymbol{x}_j$. Given two tensors $\mathcal{X}$ and $\mathcal{Y}$ with same sizes $n_1 \times n_2 \times n_3$, the inner product is defined as $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{k=1}^{n_3} \langle \boldsymbol{X}^{(k)}, \boldsymbol{Y}^{(k)} \rangle$. The Frobenius norm and the ℓ_∞ norm of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ are respectively denoted as $\|\mathcal{X}\|_F = \sqrt{\sum_{i,j,k} x_{i,j,k}^2}$ and $\|\mathcal{X}\|_\infty = \max_{i,j,k} |x_{i,j,k}|$. The $\ell_{2,1}$ norm of a matrix $\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}$ is defined as $\|\boldsymbol{X}\|_{2,1} = \sum_{j=1}^{n_2} \|\boldsymbol{x}_j\|_2$.

The matrix nuclear norm is defined as $\|\boldsymbol{X}\|_* = \sum \sigma(\boldsymbol{X})$, where $\sigma(\boldsymbol{X})$ is the singular value.

B. Preliminaries About Transformed Tensor Nuclear Norm

Given a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the block circular matrix $\texttt{bcirc}(\mathcal{X})$ is defined as

$$\texttt{bcirc}(\mathcal{X}) = \begin{bmatrix} \boldsymbol{X}^{(1)} & \boldsymbol{X}^{(n_3)} & \dots & \boldsymbol{X}^{(2)} \\ \boldsymbol{X}^{(2)} & \boldsymbol{X}^{(1)} & \dots & \boldsymbol{X}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X}^{(n_3)} & \boldsymbol{X}^{(n_3-1)} & \dots & \boldsymbol{X}^{(1)} \end{bmatrix} \in \mathbb{R}^{n_1 n_3 \times n_2 n_3}.$$

The operators $\texttt{unfold}$ and $\texttt{fold}$ are defined as

$$\texttt{unfold}(\mathcal{X}) = \begin{bmatrix} \boldsymbol{X}^{(1)} \\ \boldsymbol{X}^{(2)} \\ \vdots \\ \boldsymbol{X}^{(n_3)} \end{bmatrix} \in \mathbb{R}^{n_1 n_3 \times n_2},$$

$$\texttt{fold}(\texttt{unfold}(\mathcal{X})) = \mathcal{X}.$$

From the definitions, $\texttt{unfold}$ stacks the frontal slices of $\mathcal{X}$ to a column-wise matrix, and $\texttt{fold}$ is the inverse operator.

The work [50] proposed a new definition of *t-product* for the tensor-tensor product induced by the invertible linear transform. To begin with, let $\mathcal{D} : \mathbb{R}^{n_1 \times n_2 \times n_3} \to \mathbb{R}^{n_1 \times n_2 \times n_3}$ be the linear transform, and $\bar{\mathcal{X}}$ be the tensor obtained by applying the transform to $\mathcal{X}$ along the third mode, then

$$\bar{\mathcal{X}} = \mathcal{D}(\mathcal{X}) = \mathcal{X} \times_3 \boldsymbol{D}, \tag{1}$$

where $\times_3$ denotes the mode-3 product [51], and $\boldsymbol{D}$ is the arbitrary invertible matrix. The inverse operator can be formulated as

$$\mathcal{D}^{-1}(\mathcal{X}) = \mathcal{X} \times_3 \boldsymbol{D}^{-1}.$$

For two tensors $\mathcal{A}$ of size $n_1 \times n_2 \times n_3$ and $\mathcal{B}$ of size $n_2 \times n_4 \times n_3$, we denote $\mathcal{C} = \mathcal{A} \odot \mathcal{B}$ as the frontal-slice-wise product [50], i.e., $\boldsymbol{C}^{(k)} = \boldsymbol{A}^{(k)} \boldsymbol{B}^{(k)}$, $k = 1, \dots, n_3$.

Note that the following definitions are all based on the linear transform $\mathcal{D}$.

Definition 1. (T-product [50]): The transform $\mathcal{D}$ based t-product of two tensors $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times n_4 \times n_3}$ is denoted as $\mathcal{C} = \mathcal{A} *_{\boldsymbol{D}} \mathcal{B}$, which is equivalent to $\mathcal{D}(\mathcal{C}) = \mathcal{D}(\mathcal{A}) \odot \mathcal{D}(\mathcal{B})$.

We denote $\bar{\boldsymbol{X}} \in \mathbb{R}^{n_1 n_3 \times n_2 n_3}$ as a block diagonal matrix, whose kth diagonal block is the kth frontal slice $\bar{\boldsymbol{X}}^{(k)}$ of $\bar{\mathcal{X}} = \mathcal{D}(\mathcal{X})$, as follows:

$$\bar{\boldsymbol{X}} = \texttt{bdiag}(\bar{\mathcal{X}}) = \begin{bmatrix} \bar{\boldsymbol{X}}^{(1)} & & & \\ & \bar{\boldsymbol{X}}^{(2)} & & \\ & & \ddots & \\ & & & \bar{\boldsymbol{X}}^{(n_3)} \end{bmatrix},$$

where the operator $\texttt{bdiag}(\cdot)$ represents the frontal slices of $\bar{\mathcal{X}}$ as the diagonal blocks of a matrix. Then one can get that $\mathcal{D}(\mathcal{C}) = \mathcal{D}(\mathcal{A}) \odot \mathcal{D}(\mathcal{B})$ is equivalent to $\bar{\boldsymbol{C}} = \bar{\boldsymbol{A}} \bar{\boldsymbol{B}}$. This implies that the linear transformed t-product of two tensors in the original domain can be calculated using the matrix-matrix product in the transformed domain.

Definition 2. (Tensor transpose [50]): Let $\mathcal{X}$ be a third-order tensor with size $n_1 \times n_2 \times n_3$. Its tensor transpose, denoted as $\mathcal{X}^\top \in \mathbb{R}^{n_2 \times n_1 \times n_3}$, satisfies $\mathcal{D}(\mathcal{X}^\top)^{(k)} = (\mathcal{D}(\mathcal{X})^{(k)})^\top$, $k = 1, \dots, n_3$.

Definition 3. (Identify tensor [50]): Assuming $\mathcal{I} \in \mathbb{R}^{n \times n \times n_3}$ satisfied that the each frontal slice of $\bar{\mathcal{I}} = \mathcal{D}(\mathcal{I})$ is an identify matrix with size $n \times n$. Then $\mathcal{I} = \mathcal{D}^{-1}(\bar{\mathcal{I}})$ is called as the identify tensor.

Definition 4. (Orthogonal tensor [50]): If a tensor $\mathcal{Q} \in \mathbb{R}^{n \times n \times n_3}$ satisfies $\mathcal{Q}^\top *_{\boldsymbol{D}} \mathcal{Q} = \mathcal{Q} *_{\boldsymbol{D}} \mathcal{Q}^\top = \mathcal{I}$, then $\mathcal{Q}$ is called as an orthogonal tensor.

Definition 5. (F-diagonal tensor [50]): If each frontal slice of a tensor $\mathcal{Q} \in \mathbb{R}^{n \times n \times n_3}$ is a diagonal matrix, then $\mathcal{Q}$ is an F-diagonal tensor.

Based on the above definitions, we give the definition of transformed tensor singular value decomposition (TTSVD) and tensor tubal rank.

Definition 6. (TTSVD [50]): Given a third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, it can be decomposed as

$$\mathcal{X} = \mathcal{U} *_{\boldsymbol{D}} \mathcal{S} *_{\boldsymbol{D}} \mathcal{V}^\top,$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are two orthogonal tensors, and $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is F-diagonal.

Definition 7. (Tensor tubal rank [52]): Let $\mathcal{X} = \mathcal{U} *_{\boldsymbol{D}} \mathcal{S} *_{\boldsymbol{D}} \mathcal{V}^\top$ be the TTSVD of a tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. The tensor tubal rank $\text{rank}_t(\mathcal{X})$ is equivalent to the number of nonzero singular tubes of $\mathcal{S}$, that is $\text{rank}_t(\mathcal{X}) = \#\{i : \mathcal{S}(i,i,:) \neq 0\}$.

If a tensor $\mathcal{X}$ has the tubal rank r, one can get its skinny TTSVD as $\mathcal{X} = \mathcal{U}_\mathcal{X} *_{\boldsymbol{D}} \mathcal{S}_\mathcal{X} *_{\boldsymbol{D}} \mathcal{V}_\mathcal{X}^\top$, where $\mathcal{U}_\mathcal{X} = \mathcal{U}(:, 1:r, :)$, $\mathcal{S}_\mathcal{X} = \mathcal{S}(1:r, 1:r, :)$, and $\mathcal{V}_\mathcal{X} = \mathcal{V}(:, 1:r, :)$ in which $\mathcal{U}_\mathcal{X}^\top *_{\boldsymbol{D}} \mathcal{U}_\mathcal{X} = \mathcal{I}$ and $\mathcal{V}_\mathcal{X}^\top *_{\boldsymbol{D}} \mathcal{V}_\mathcal{X} = \mathcal{I}$.

Next, we give the convex surrogate of the nonconvex tensor tubal rank, named as the transformed tensor nuclear norm (TTNN). To proceed, we have the following assumption on the matrix $\boldsymbol{D}$ in (1), i.e.,

$$\boldsymbol{D}^\top \boldsymbol{D} = \boldsymbol{D} \boldsymbol{D}^\top = D \boldsymbol{I}_{n_3}, \tag{2}$$

where $D > 0$ is a constant, and $\boldsymbol{I}_{n_3} \in \mathbb{R}^{n_3 \times n_3}$ is an identify matrix.

Definition 8. (TTNN [52]): Let the linear transform $\mathcal{D}$ satisfy the assumption in (2). For tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its TTNN is defined as

$$\|\mathcal{X}\|_* = \frac{1}{D} \sum_{k=1}^{n_3} \|\bar{\boldsymbol{X}}^{(k)}\|_*.$$

C. Unified Nonconvex Transformed Low-Rank Tensor Approximation

Here, we give the proposed unified nonconvex transformed low-rank tensor approximation.

Definition 9 (Unified nonconvex surrogate): Let $\mathcal{X} = \mathcal{U} *_{\boldsymbol{D}} \mathcal{S} *_{\boldsymbol{D}} \mathcal{V}^\top$ be the TTSVD of a third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, then the unified nonconvex surrogate of the transformed tensor

nuclear norm is defined as

$$\Phi(\mathcal{X}) = \frac{1}{D}\sum_{k=1}^{n_3}\sum_{i=1}^{n}\phi(\sigma_{ik}(\mathcal{X})), \tag{3}$$

where $n = \min\{n_1, n_2\}$, $\sigma_{ik}(\mathcal{X})$ is the ith singular value of the kth frontal slice, and $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ is the concave, continuous, and monotonically increasing function.

Many nonconvex functions ϕ have been widely studied, e.g., ℓ_p-norm [53], minimax concave penalty (MCP) [54], smoothly clipped absolute deviation (SCAD) [55], Logarithm [56], and Laplace [57], more details can be found in [58], [59].

Next, we consider the following generalized proximal problem of $\Phi(\mathcal{X})$

$$\min_{\mathcal{X}} \ \lambda\Phi(\mathcal{X}) + \frac{1}{2}\|\mathcal{X} - \mathcal{Y}\|_F^2. \tag{4}$$

However, the most challenge is that for a general nonconvex function $\Phi(\mathcal{X})$, it may not exist the explicit solver on tensor singular values of $\mathcal{Y}$ for (4). In this paper, we utilize the iteratively reweighted scheme to solve Problem (4).

For the nonconvex ϕ defined in Def. (3), due to the concavity, one can obtain the antimonotone property [58], [59], i.e., when $x \geq y, u \geq v$, for any $u \in \partial\phi(x)$ and $v \in \partial\phi(y)$, where $\partial\phi(x)$ is the supergradient of $\phi(\cdot)$ at x. We denote $\sigma_{ik} = \sigma_{ik}(\mathcal{X})$ and $\sigma_{ikt} = \sigma_{ik}(\mathcal{X}_t)$ for simplicity. On the other hand, ϕ satisfies

$$\phi(\sigma_{ik}) \leq \phi(\sigma_{ikt}) + \omega_{ikt}(\sigma_{ik} - \sigma_{ikt}), \tag{5}$$

where $\omega_{ikt} \in \partial\phi(\sigma_{ikt})$. Since $\sigma_{1kt} \geq \sigma_{2kt} \geq \cdots \geq \sigma_{nkt} \geq 0, k = 1, 2, \ldots, n_3$, by the antimonotone property, we have

$$0 \leq \omega_{1kt} \leq \omega_{2kt} \leq \cdots \leq \omega_{nkt}.$$

Instead of $\Phi(\mathcal{X})$, we can minimize the right-hand side of (5). Then we solve the following nonconvex relaxed problem

$$\begin{aligned}\mathcal{X}^\star =& \ \operatorname*{argmin}_{\mathcal{X}} \ \lambda\frac{1}{D}\sum_{k=1}^{n_3}\sum_{i=1}^{n}(\phi(\sigma_{ikt}) + \omega_{ikt}(\sigma_{ik} - \sigma_{ikt})) \\ &+ \frac{1}{2}\|\mathcal{X} - \mathcal{Y}\|_F^2 \\ =& \ \operatorname*{argmin}_{\mathcal{X}} \ \lambda\|\mathcal{X}\|_{\omega,*} + \frac{1}{2}\|\mathcal{X} - \mathcal{Y}\|_F^2,\end{aligned} \tag{6}$$

where $\|\mathcal{X}\|_{\omega,*} = \frac{1}{D}\sum_{k=1}^{n_3}\sum_{i=1}^{n}\omega_{ikt}\sigma_{ik}$. Fortunately, (6) admits a closed-form solution, as shown in the following weighted transformed tensor singular value thresholding.

Theorem 1: For any $\lambda > 0$, $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3}$, and the weight $0 \leq \omega_{1k} \leq \omega_{2k} \leq \cdots \leq \omega_{nk}, k = 1, 2, \ldots, n_3$, the following problem

$$\operatorname*{argmin}_{\mathcal{X}} \ \lambda\|\mathcal{X}\|_{\omega,*} + \frac{1}{2}\|\mathcal{X} - \mathcal{Y}\|_F^2 \tag{7}$$

admits a globally optimal solution

$$\mathcal{X}^\star = \mathcal{U} *_{\boldsymbol{D}} \mathcal{T}_{\omega,\lambda}(\mathcal{S}) *_{\boldsymbol{D}} \mathcal{V}^\top, \tag{8}$$

where $\mathcal{U} *_{\boldsymbol{D}} \mathcal{S} *_{\boldsymbol{D}} \mathcal{V}^\top$ is TTSVD of $\mathcal{Y}$, and $\mathcal{T}_{\omega,\lambda}(\mathcal{S}) = \mathcal{D}^{-1}((\bar{\mathcal{S}} - \lambda\mathcal{W})_+)$, in which $\mathcal{W} \in \mathbb{R}^{n_1\times n_2\times n_3}$ is F-diagonal whose kth diagonal frontal slice is $\mathcal{W}(:,:,k) = \mathrm{Diag}(\omega_{1k}, \omega_{2k}, \ldots, \omega_{nk})$.

The proof is presented in Appendix A, available online.

III. Related Works

In this section, we briefly introduce related works about subspace clustering and multi-view subspace clustering.

A. Subspace Clustering

Subspace clustering methods construct the affinity matrix using the representation coefficient matrices and group data samples into different clusters. Assuming there exists a dataset $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]$, and data points are sampled from multiple subspaces. Under the framework of subspace clustering, each data sample can be expressed as the linear combination of all data points, i.e., $\boldsymbol{X} = \boldsymbol{X}\boldsymbol{Z} + \boldsymbol{E}$, where $\boldsymbol{Z}$ is the linear self-representation coefficient matrix, and $\boldsymbol{E}$ is the error matrix.

Subspace clustering aims to obtain the subspace representation by solving

$$\begin{aligned}&\min_{\boldsymbol{Z},\boldsymbol{E}} \ \mathcal{R}_1(\boldsymbol{Z}) + \lambda\mathcal{R}_2(\boldsymbol{E}), \\ &\text{s.t. } \boldsymbol{X} = \boldsymbol{X}\boldsymbol{Z} + \boldsymbol{E},\end{aligned} \tag{9}$$

where $\mathcal{R}_1(\boldsymbol{Z})$ and $\mathcal{R}_2(\boldsymbol{E})$ are the regularization functions of $\boldsymbol{Z}$ and $\boldsymbol{E}$, respectively. Some popular regularization terms were studied to exploit the different underlying structures of the coefficient matrix. For example, SSC [29] applied the ℓ_1 norm (i.e., $\|\boldsymbol{Z}\|_1$ and $\|\boldsymbol{E}\|_1$) to find the sparse representation and noise; LRR [28] utilized the trace norm of matrix $\|\boldsymbol{Z}\|_*$ to capture the low-rankness of the representation coefficient and $\ell_{2,1}$ to learn the noise structure; the work [20] combined the sparsity and low-rankness for revealing the coefficient matrix. When the observed data samples are insufficient, the clustering performance will be dropped. To resolve the problem of insufficient observed samples, the work [30] introduced the hidden data and proposed the following problem

$$\min_{\boldsymbol{Z}} \ \|\boldsymbol{Z}\|_*, \ \text{s.t. } \boldsymbol{X} = [\boldsymbol{X}, \boldsymbol{X}_h]\boldsymbol{Z}, \tag{10}$$

where $\boldsymbol{X}_h$ is the hidden data. It is not practical to recover $\boldsymbol{Z}^\star$ due to the absence of $\boldsymbol{X}_h$. As proved in [30], one can reformulate the constraint in (10) as follows:

$$\boldsymbol{X} = \boldsymbol{X}\boldsymbol{Z} + \boldsymbol{L}\boldsymbol{X}.$$

Under the low-rank assumption, [30] proposed the following model for noisy subspace clustering:

$$\min_{\boldsymbol{Z},\boldsymbol{L},\boldsymbol{E}} \ \|\boldsymbol{Z}\|_* + \|\boldsymbol{L}\|_* + \lambda\|\boldsymbol{E}\|_1, \ \text{s.t. } \boldsymbol{X} = \boldsymbol{X}\boldsymbol{Z} + \boldsymbol{L}\boldsymbol{X} + \boldsymbol{E}.$$

Many variants [44], [45], [46], [47] have been developed to solve the problem of insufficient observed samples. To reduce the computational complexity, the work [44] proposed the fixed-rank representation based on matrix factorization to explore the low-rank prior of the representation coefficient, and the work [46] utilized the Frobenius norm to replace the nuclear norm. Based on [30], the works [45], [47] developed a variety of regularization terms to characterize the priors of the representation coefficients, such as Laplacian regularization [45] and weighted distance penalty regularization [47].

Although these methods achieved good performance in solving the problem of insufficient observed samples, they are designed to handle the data with single-view feature. Compared with the single-view data, the multi-view data contains multiple distinct features, each of which represents the particular information relative to the corresponding view, and different features can represent more complementary information to each other. Therefore, for the multi-view clustering, the main challenges are how to simultaneously handle the insufficient observed samples in multi-view case and fully explore the consensus and complementary information among multiple views.

B. Multi-View Subspace Clustering

Multi-view subspace clustering first learns the affinity matrix using the multiple representation matrices and then performs spectral clustering to get clustering results. The single-view subspace clustering shown in (9) can be intuitively developed for the multi-view case. To be specific, let $\boldsymbol{X}^{(v)}(v=1,\ldots,V)$ be the vth view's data matrix, and $\boldsymbol{Z}^{(v)}$ be the corresponding subspace representation, then the general model of MVSC can be formulated as

$$\min_{\boldsymbol{Z}^{(v)},\boldsymbol{E}^{(v)}} \mathcal{R}_1(\boldsymbol{Z}^{(1)},\ldots,\boldsymbol{Z}^{(V)})+\lambda\mathcal{R}_2\left(\boldsymbol{E}^{(1)},\ldots,\boldsymbol{E}^{(V)}\right),$$
$$\text{s.t. } \boldsymbol{X}^{(v)}=\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}+\boldsymbol{E}^{(v)},\ v=1,\ldots,V, \tag{11}$$

where $\mathcal{R}_1$ and $\mathcal{R}_2$ are the regularizers of $\boldsymbol{Z}^{(v)}$ and $\boldsymbol{E}^{(v)}$, respectively. Then the final affinity matrix for MVSC can be constructed using $\boldsymbol{S}=\sum_{v=1}^{V}(\boldsymbol{Z}^{(v)}+(\boldsymbol{Z}^{(v)})^{\top})/V$. The work [9] introduced an indicator matrix to ensure consistency among different views and performed the clustering for all subspace representations simultaneously. The work [60] constructed an underlying representation for all subspace representations to explore the complementary information among different views. Some other matrix-based methods, e.g., structured matrix factorization [19] and deep matrix factorization [23], were proposed to preserve the local information. However, these matrix-based methods only capture the pairwise correlations within different views [10], [32].

To fully explore the correlation between different views' representations, a number of MVSC methods from the tensor perspective, which stack the multiple representation matrices $(\{\boldsymbol{Z}_v\}_{v=1}^{V})$ as a third-order tensor, have been proposed using low-rank tensor representation. For example, the work [10] proposed a tensor multi-view clustering method based on Tucker tensor decomposition. They used the sum of the nuclear norm to explore the low-rankness of the stacked third-order representation tensor. Inspired by the powerful ability of TSVD, the work [24] applied TNN on the stacked representation tensor to exploit the correlation hidden in all views. The work [61] sought the underlying representation tensor under both sparse and low-rank constraints. Although the aforementioned LRR-based tensor clustering methods achieve impressive performance, there exist some limitations. First, the basic assumption of LRR is that the sampling points in each view should be sufficient for capturing the correlation between sample points [24], [30]. Second, the TNN-based methods keep the stacked self-representation tensor in the fixed Fourier domain—which may not be suitable for different multi-view data. Third, the existing MVSC methods based on the nuclear norm and TNN treat all singular values equally and ignore the prior information of each singular value. To solve these problems, we first introduce the latent low-rank presentation into MVSC to support the insufficient samples. We also propose a unified transformed low-rank tensor representation to better explore the high correlation and complementary information among all views and suit different kinds of multi-view data clustering problems.

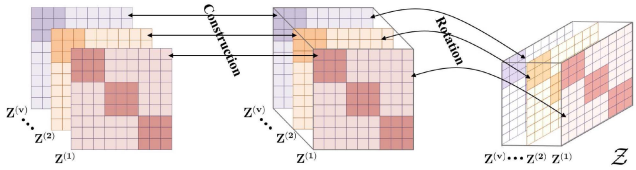


Fig. 1. Illustration of tensor $\mathcal{Z}$ construction by the operator Ψ.

IV. Proposed Model and Algorithm

In this section, we present the proposed nonconvex latent transformed low-rank tensor model for MVSC and the designed ADMM-based [62], [63] algorithm.

A. Proposed Model

To solve the insufficient sampling problem in MVSC, we utilize the latent low-rank representation strategy [30]. Specifically, denoting $\boldsymbol{X}^{(v)}\in\mathbb{R}^{M_v\times N}$ as the vth view data, we introduce the hidden data into each view and represent each observed data as follow:

$$\boldsymbol{X}^{(v)}=\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}+\boldsymbol{L}^{(v)}\boldsymbol{X}^{(v)}+\boldsymbol{E}^{(v)},$$

where $\boldsymbol{Z}^{(v)}\in\mathbb{R}^{N\times N}$ and $\boldsymbol{L}^{(v)}\in\mathbb{R}^{M_v\times M_v}$ are the subspace representation coefficient and salient representation, respectively, $\boldsymbol{E}^{(v)}\in\mathbb{R}^{M_v\times N}$ is the corresponding error matrix.

In order to fully capture the high correlation embedded in inter-views, we construct the tensor $\mathcal{Z}=\Psi(\boldsymbol{Z}^{(1)},\boldsymbol{Z}^{(2)},\ldots,\boldsymbol{Z}^{(V)})$ following [24], where the operator $\Psi(\cdot)$ merges the representation matrices $\{\boldsymbol{Z}^{(v)}\}_{v=1}^{V}$ to a 3D tensor, and then rotates its dimensionality to $N\times V\times N$; see Fig. 1. The inverse operator $\Psi^{-1}(\mathcal{Z})$ rotates $\mathcal{Z}$ to an $N\times N\times V$ tensor and $\Psi_{(v)}^{-1}(\mathcal{Z})=\boldsymbol{Z}^{(v)}$. We apply the noncovex approximation $\Phi(\cdot)$ on the tensor $\mathcal{Z}$ for better capturing the correlation among views and impose the nuclear norm on the salient representations $\{\boldsymbol{L}^{(v)}\}_{v=1}^{V}$ to exploit the relationship between samples of one view.

We proposed the following optimization formulation

$$\min_{\boldsymbol{Z}^{(v)},\boldsymbol{L}^{(v)},\boldsymbol{E}^{(v)}} \lambda_1\Phi(\mathcal{Z})+\sum_{v=1}^{V}\left\|\boldsymbol{L}^{(v)}\right\|_{*}+\lambda_2\|\boldsymbol{E}\|_{2,1}$$
$$\text{s.t.}\boldsymbol{X}^{(v)}=\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}+\boldsymbol{L}^{(v)}\boldsymbol{X}^{(v)}+\boldsymbol{E}^{(v)},\ v=1,\ldots,V,$$
$$\mathcal{Z}=\Psi\left(\boldsymbol{Z}^{(1)},\boldsymbol{Z}^{(2)},\ldots,\boldsymbol{Z}^{(V)}\right),\boldsymbol{E}=\left[\boldsymbol{E}^{(1)};\boldsymbol{E}^{(2)};\ldots;\boldsymbol{E}^{(V)}\right], \tag{12}$$

where $\mathcal{Z} \in \mathbb{R}^{N\times V\times N}$, λ_1 and λ_2 are two parameters, $\boldsymbol{E} \in \mathbb{R}^{\sum M_v\times N}$ is the vertical concatenation along the column of the error matrices, and the $\ell_{2,1}$-norm is used to make the columns of $\boldsymbol{E}^{(v)}$ to have consistent magnitude values.

Remark 1: Compared with existing MVSC methods, our method have the following advantages: First, we are the first attempt to resolve the insufficient sampling problem of existing LRR-based tensor clustering approaches. Second, we propose a unified nonconvex transfromed tensor low-rank approximation to explore the prior information of the singular values of the representation tensor to better characterize the high correlation among all views. Third, different from previous TNN-based clustering works, we preserve the representation coefficient structure in the transformed domain determined by the intrinsic property of multi-view data itself.

Remark 2: In our proposed model, we apply the nuclear norm on the representation $\boldsymbol{L}^{(v)}$ to explore the relationship among samples within each individual view. The multi-view data describes the data from multiple views, hence there may exist some correlation among different $\boldsymbol{L}^{(v)}$. However, since the sizes of $\boldsymbol{L}^{(v)}$ are different with each other, we handle the representation $\boldsymbol{L}^{(v)}$ independently. One interesting direction is to stack all $\{\boldsymbol{L}^{(v)}\}_{v=1}^{V}$ into an irregular tensor whose size is not fixed, and then apply the irregular tensor factorization [64], [65], [66] to exploit the correlation among all salient representations.

Remark 3: Note that some nonconvex low-rank tensor approximations have been proposed for MVSC problem, e.g, [21], [40]. However, all of these methods deal with the representation tensor in the Fourier domain for different multi-view data, not consider the structure of data sampling. Different from existing nonconvex methods, we preserve the representation coefficient structure in the transformed domain determined by the intrinsic property of multi-view data itself, which is more useful to explore the high correlation among different views. Meanwhile, we propose a unified nonconvex low-rank tensor surrogate and give a general solver for the related proximal problem. The existing nonconvex approximations can be merged into the proposed nonconvex low-rank tensor representation framework.

B. Proposed Algorithm

We solve the proposed (12) using ADMM. By introducing the auxiliary variables $\mathcal{Z} = \mathcal{F}$ and $\boldsymbol{L}^{(v)} = \boldsymbol{P}^{(v)}$, the augmented Lagrangian function becomes

$$\begin{aligned}
&\mathcal{L}\left(\boldsymbol{Z}^{(v)};\boldsymbol{L}^{(v)};\boldsymbol{E}^{(v)};\mathcal{F};\mathcal{G};\boldsymbol{P}^{(v)};\boldsymbol{Q}^{(v)};\boldsymbol{Y}^{(v)}\right)\\
&=\lambda_1\Phi(\mathcal{F})+\sum_{v=1}^{V}\left\|\boldsymbol{P}^{(v)}\right\|_*+\lambda_2\|\boldsymbol{E}\|_{2,1}\\
&\quad+\langle\mathcal{G},\mathcal{Z}-\mathcal{F}\rangle+\frac{\rho}{2}\|\mathcal{Z}-\mathcal{F}\|_F^2\\
&\quad+\sum_{v=1}^{V}\left(\left\langle\boldsymbol{Y}^{(v)},\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{L}^{(v)}\boldsymbol{X}^{(v)}-\boldsymbol{E}^{(v)}\right\rangle\right.\\
&\quad\left.+\frac{\mu}{2}\left\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}-\boldsymbol{L}^{(v)}\boldsymbol{X}^{(v)}-\boldsymbol{E}^{(v)}\right\|_F^2\right)\\
&\quad+\sum_{v=1}^{V}\left(\left\langle\boldsymbol{Q}^{(v)},\boldsymbol{L}^{(v)}-\boldsymbol{P}^{(v)}\right\rangle+\frac{\theta}{2}\left\|\boldsymbol{L}^{(v)}-\boldsymbol{P}^{(v)}\right\|_F^2\right),
\end{aligned}\tag{13}$$

where the tensor $\mathcal{G}$ and the matrices $\boldsymbol{Y}^{(v)}$ and $\boldsymbol{Q}^{(v)}$ are Lagrange multipliers, and ρ, μ, and θ are penalty parameters. Next we respectively update $\boldsymbol{Z}^{(v)}$, $\boldsymbol{L}^{(v)}$, $\boldsymbol{E}^{(v)}$, $\boldsymbol{P}^{(v)}$, and $\mathcal{F}$ by using the alternating minimization scheme.

Updating $\mathcal{F}$: We minimize the following problem:

$$\operatorname*{argmin}_{\mathcal{F}}\ \lambda_1\Phi(\mathcal{F})+\frac{\rho_t}{2}\left\|\mathcal{Z}_t-\mathcal{F}+\frac{\mathcal{G}_t}{\rho_t}\right\|_F^2.\tag{14}$$

According to (6), we replace (14) by the following optimization problem

$$\operatorname*{argmin}_{\mathcal{F}}\ \lambda_1\|\mathcal{F}\|_{\omega,*}+\frac{\rho_t}{2}\left\|\mathcal{Z}_t-\mathcal{F}+\frac{\mathcal{G}_t}{\rho_t}\right\|_F^2,\tag{15}$$

which has the following closed-form solution by Theorem 1

$$\mathcal{F}_{t+1}=\mathcal{U}*_{\boldsymbol{D}}\mathcal{T}_{\omega,\frac{\lambda_1}{\rho_t}}(\mathcal{S})*_{\boldsymbol{D}}\mathcal{V}^\top,\tag{16}$$

where $\mathcal{U}*_{\boldsymbol{D}}\mathcal{S}*_{\boldsymbol{D}}\mathcal{V}^\top$ is the TTSVD of $\mathcal{Z}_t+\mathcal{G}_t/\rho_t$, and $\mathcal{T}_{\omega,\frac{\lambda_1}{\rho}}(\mathcal{S})=\mathcal{L}^{-1}((\bar{\mathcal{S}}-\frac{\lambda_1}{\rho_t}\mathcal{W})_+)$, $\mathcal{W}\in\mathbb{R}^{N\times V\times N}$ is a F-diagonal tensor that satisfies $\mathcal{W}(:,:,n)=\mathrm{Diag}(\omega_{1n},\omega_{2n},\ldots,\omega_{In})$, $n=1,2,\ldots,N$, in which ω_{mn} is the mth singular value of $\bar{\mathcal{F}}_t(:,:,n)$, for $m=1,2,\ldots,\min\{N,V\}$. Then we can get each $\boldsymbol{F}_{t+1}^{(v)}=\Psi_{(v)}^{-1}(\mathcal{F}_{t+1})$.

Updating $\boldsymbol{P}^{(v)}$: The $\boldsymbol{P}^{(v)}$-subproblem is

$$\operatorname*{argmin}_{\boldsymbol{P}^{(v)}}\left\|\boldsymbol{P}^{(v)}\right\|_*+\frac{\theta_t}{2}\left\|\boldsymbol{L}_t^{(v)}-\boldsymbol{P}^{(v)}+\frac{\boldsymbol{Q}_t^{(v)}}{\theta_t}\right\|_F^2.\tag{17}$$

We can get the optimal solution by the singular value thresholding operator [67] as follows:

$$\boldsymbol{P}_{t+1}^{(v)}=\boldsymbol{U}\boldsymbol{T}_{1/\theta_t}(\boldsymbol{S})\boldsymbol{V}^\top,\tag{18}$$

where $\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top$ is the SVD of $\boldsymbol{L}_t^{(v)}+\boldsymbol{Q}_t^{(v)}/\theta_t$, and $\boldsymbol{T}_{1/\theta_t}(\boldsymbol{S})=\mathrm{Diag}\{\boldsymbol{S}_{m_vm_v}-1/\theta_t\}_+$, $m_v=1,2,\ldots,M_v$.

Updating $\boldsymbol{E}^{(v)}$: The $\boldsymbol{E}^{(v)}$-subproblem is

$$\begin{aligned}
\operatorname*{argmin}_{\boldsymbol{E}}\lambda_2\|\boldsymbol{E}\|_{2,1}+\sum_{v=1}^{V}\frac{\mu_t}{2}\left\|\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}_t^{(v)}-\right.\\
\left.\boldsymbol{L}_t^{(v)}\boldsymbol{X}^{(v)}-\boldsymbol{E}^{(v)}+\frac{\boldsymbol{Y}_t^{(v)}}{\mu_t}\right\|_F^2.
\end{aligned}\tag{19}$$

Denoting $\widetilde{\boldsymbol{E}}\in\mathbb{R}^{\sum M_v\times N}$ as the vertically concatenating the matrices $\boldsymbol{X}^{(v)}-\boldsymbol{X}^{(v)}\boldsymbol{Z}_t^{(v)}-\boldsymbol{L}_t^{(v)}\boldsymbol{X}^{(v)}+\boldsymbol{Y}_t^{(v)}/\mu_t$ together along with the column, then (19) is equivalent to

$$\operatorname*{argmin}_{\boldsymbol{E}}\lambda_2\|\boldsymbol{E}\|_{2,1}+\frac{\mu_t}{2}\left\|\boldsymbol{E}-\widetilde{\boldsymbol{E}}\right\|_F^2.\tag{20}$$

The above problem can be divided into independent several column-wise subproblems

$$\operatorname*{argmin}_{\boldsymbol{E}(:,i)}\lambda_2\|\boldsymbol{E}(:,i)\|_2+\frac{\mu_t}{2}\left\|\boldsymbol{E}(:,i)-\widetilde{\boldsymbol{E}}(:,i)\right\|_2^2,\tag{21}$$

which has a closed-form solution

$$\boldsymbol{E}(:,i)_{t+1} = \begin{cases} \frac{\|\widetilde{\boldsymbol{E}}(:,i)\|_2 - \lambda_2/\mu_t}{\|\widetilde{\boldsymbol{E}}(:,i)\|_2}\widetilde{\boldsymbol{E}}(:,i), \|\widetilde{\boldsymbol{E}}(:,i)\|_2 > \lambda_2/\mu_t, \\ \boldsymbol{0}, \qquad\qquad\qquad \text{otherwise}, \end{cases} \tag{22}$$

where $\widetilde{\boldsymbol{E}}(:,i)$ represents the ith column of the matrix $\widetilde{\boldsymbol{E}}$.

Updating $\boldsymbol{Z}^{(v)}$: The $\boldsymbol{Z}^{(v)}$-subproblem is

$$\underset{\boldsymbol{Z}^{(v)}}{\operatorname{argmin}} \frac{\rho_t}{2}\left\|\boldsymbol{Z}^{(v)} - \boldsymbol{F}_{t+1}^{(v)} + \frac{\boldsymbol{G}_t^{(v)}}{\rho_t}\right\|_F^2 + \frac{\mu_t}{2}\left\|\widetilde{\boldsymbol{Z}} - \boldsymbol{X}^{(v)}\boldsymbol{Z}^{(v)}\right\|_F^2,$$

where $\widetilde{\boldsymbol{Z}} = \boldsymbol{X}^{(v)} - \boldsymbol{L}_t^{(v)}\boldsymbol{X}^{(v)} - \boldsymbol{E}_{t+1}^{(v)} + \boldsymbol{Y}_t^{(v)}/\mu_t$. The closed-form solution is

$$\boldsymbol{Z}_{t+1}^{(v)} = \left(\rho_t\boldsymbol{I} + \mu_t(\boldsymbol{X}^{(v)})^\top\boldsymbol{X}^{(v)}\right)^{-1}\left(\mu_t(\boldsymbol{X}^{(v)})^\top\widetilde{\boldsymbol{Z}} + \rho_t\boldsymbol{F}_{t+1}^{(v)} - \boldsymbol{G}_t^{(v)}\right). \tag{23}$$

Updating $\boldsymbol{L}^{(v)}$: The $\boldsymbol{L}^{(v)}$-subproblem is

$$\underset{\boldsymbol{L}^{(v)}}{\operatorname{argmin}} \frac{\mu_t}{2}\left\|\widetilde{\boldsymbol{L}} - \boldsymbol{L}^{(v)}\boldsymbol{X}^{(v)}\right\|_F^2 + \frac{\theta_t}{2}\left\|\boldsymbol{L}^{(v)} - \boldsymbol{P}_{t+1}^{(v)} + \frac{\boldsymbol{Q}_t^{(v)}}{\theta_t}\right\|_F^2,$$

where $\widetilde{\boldsymbol{L}} = \boldsymbol{X}^{(v)} - \boldsymbol{X}^{(v)}\boldsymbol{Z}_{t+1}^{(v)} - \boldsymbol{E}_{t+1}^{(v)} + \boldsymbol{Y}_t^{(v)}/\mu_t$. Then we can get the solution as

$$\boldsymbol{L}_{t+1}^{(v)} = \left(\mu_t\widetilde{\boldsymbol{L}}(\boldsymbol{X}^{(v)})^\top + \theta_t\boldsymbol{P}_{t+1}^{(v)} - \boldsymbol{Q}_t^{(v)}\right)\left(\theta_t\boldsymbol{I} + \mu_t\boldsymbol{X}^{(v)}(\boldsymbol{X}^{(v)})^\top\right)^{-1}. \tag{24}$$

Updating $\boldsymbol{Y}^{(v)}$, $\mathcal{G}$, and $\boldsymbol{Q}^{(v)}$: The multipliers $\boldsymbol{Y}^{(v)}$, $\mathcal{G}$, and $\boldsymbol{Q}^{(v)}$ are updated as follows:

$$\mathcal{G}_{t+1} = \mathcal{G}_t + \rho_t(\mathcal{Z}_{t+1} - \mathcal{F}_{t+1}), \tag{25}$$

$$\boldsymbol{Q}_{t+1}^{(v)} = \boldsymbol{Q}_t^{(v)} + \theta_t\left(\boldsymbol{L}_{t+1}^{(v)} - \boldsymbol{P}_{t+1}^{(v)}\right), \tag{26}$$

$$\boldsymbol{Y}_{t+1}^{(v)} = \boldsymbol{Y}_t^{(v)} + \mu_t\left(\boldsymbol{X}^{(v)} - \boldsymbol{X}^{(v)}\boldsymbol{Z}_{t+1}^{(v)} - \boldsymbol{L}_{t+1}^{(v)}\boldsymbol{X}^{(v)} - \boldsymbol{E}_{t+1}^{(v)}\right). \tag{27}$$

We summarize our unified latent tensor low-rank representation (ULatTLRR) for MVSC in Algorithm 1.

C. Computational Complexity

To update $\mathcal{F} \in \mathbb{R}^{N\times V\times N}$, computing the transform of $\mathcal{F} \times_3 \boldsymbol{D}$ takes $\mathcal{O}(VN^3)$, computing N SVDs on $N \times V$ matrices costs $\mathcal{O}(V^2N^2)$, and computing the inverse transform of $\mathcal{F} \times_3 \boldsymbol{D}^\top$ will take $\mathcal{O}(VN^3)$. To update $\boldsymbol{P}^{(v)} \in \mathbb{R}^{M_v\times M_v}$, it mainly involves the SVD of $M_v \times M_v$ at the cost of $\mathcal{O}(M_v^3)$. To update $\boldsymbol{E}^{(v)} \in \mathbb{R}^{M_v\times N}$, it costs $\mathcal{O}(M_vN(M_v + N))$. Computing $\boldsymbol{Z}^{(v)} \in \mathbb{R}^{N\times N}$ and $\boldsymbol{L}^{(v)} \in \mathbb{R}^{M_v\times M_v}$ cost $\mathcal{O}(M_v^2N + M_vN^2 + N^3)$ and $\mathcal{O}(M_v^2N + M_vN^2 + M_v^3)$, respectively. Therefore, Algorithm 1 takes $\mathcal{O}(VN^3 + V^2N^2 + \sum_{v=1}^V M_v(M_v^2 + M_vN + N^2))$ at each iteration.

Algorithm 1: ADMM-Based Algorithm for Solving (12).

Input: Multi-view data $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots, \boldsymbol{X}^{(V)}$, the regularization parameters λ_1 and λ_2, and cluster number K.

Initialize: $\mathcal{Z} = \mathcal{F} = 0$, $\boldsymbol{L}^{(v)} = \boldsymbol{P}^{(v)} = 0$, $\boldsymbol{E}^{(v)} = 0$, $\boldsymbol{Y}^{(v)} = 0$, $\boldsymbol{Q}^{(v)} = 0, v = 1, 2, \ldots, V$, $\mathcal{G} = 0$, $\mu_0 = \rho_0 = \theta_0 = 10^{-4}$, $\mu_{\max} = \rho_{\max} = \theta_{\max} = 10^6$, $\eta = 2$, and $\varepsilon = 10^{-7}$.

While not converged **do**

1. Update $\mathcal{F}$ by (16).
2. **for** $v = 1, \ldots, V$ **do**
3. Update $\boldsymbol{P}^{(v)}$ by (18).
4. **end**
5. **for** $v = 1, \ldots, V$ **do**
6. Update $\boldsymbol{E}^{(v)}$ by (22).
7. **end**
8. **for** $v = 1, \ldots, V$ **do**
9. Update $\boldsymbol{Z}^{(v)}$ by (23).
10. **end**
11. **for** $v = 1, \ldots, V$ **do**
12. Update $\boldsymbol{L}^{(v)}$ by (24).
13. **end**
14. Update multipliers by (25), (26), and (27).
15. Update $\mu_{t+1} = \min(\eta\mu_t, \mu_{\max})$, $\rho_{t+1} = \min(\eta\rho_t, \rho_{\max})$, and $\theta_{t+1} = \min(\eta\theta_t, \theta_{\max})$.
16. $(\boldsymbol{Z}^{(1)}, \boldsymbol{Z}^{(2)}, \ldots, \boldsymbol{Z}^{(V)}) = \Psi^{-1}(\mathcal{Z})$.
17. Check the stopping criteria:

$$\max_v\left(\left\|\boldsymbol{X}^{(v)} - \boldsymbol{X}^{(v)}\boldsymbol{Z}_{t+1}^{(v)} - \boldsymbol{L}_{t+1}^{(v)}\boldsymbol{X}^{(v)} - \boldsymbol{E}_{t+1}^{(v)}\right\|_\infty, \left\|\boldsymbol{Z}_{t+1}^{(v)} - \boldsymbol{F}_{t+1}^{(v)}\right\|_\infty, \left\|\boldsymbol{L}_{t+1}^{(v)} - \boldsymbol{P}_{t+1}^{(v)}\right\|_\infty\right) < \varepsilon.$$

end while

18. Obtain the affinity matrix by $\boldsymbol{S} = (\sum_{v=1}^V |\boldsymbol{Z}_{t+1}^{(v)}| + |(\boldsymbol{Z}_{t+1}^{(v)})^\top|)/V$.

Output: Applying the spectral clustering method on $\boldsymbol{S}$ to get the clustering result.

D. Space Complexity

The proposed method involves several variables, including the observed data $\{\boldsymbol{X}^{(v)} \in \mathbb{R}^{M_v\times N}\}_{v=1}^V$, the representation coefficients $\{\boldsymbol{Z}^{(v)} \in \mathbb{R}^{N\times N}\}_{v=1}^V$, salient representations $\{\boldsymbol{L}^{(v)} \in \mathbb{R}^{M_v\times M_v}\}_{v=1}^V$, error matrices $\{\boldsymbol{E}^{(v)} \in \mathbb{R}^{M_v\times N}\}_{v=1}^V$, auxiliary variables $\mathcal{F} \in \mathbb{R}^{N\times V\times N}$, $\{\boldsymbol{P}^{(v)} \in \mathbb{R}^{M_v\times M_v}\}_{v=1}^V$, and Lagrange variables $\mathcal{G} \in \mathbb{R}^{N\times V\times N}$, $\{\boldsymbol{Y}^{(v)} \in \mathbb{R}^{M_v\times N}\}_{v=1}^V$, and $\{\boldsymbol{Q}^{(v)} \in \mathbb{R}^{M_v\times M_v}\}_{v=1}^V$. Therefore, the space complexity of the proposed method is $\mathcal{O}(N^2V + N\sum_{v=1}^V M_v + \sum_{v=1}^V M_v^2)$.

E. Convergence Analysis

Now we present the convergence of the proposed algorithm.

Theorem 2: Assume that the sequence $\boldsymbol{H}_t = (\boldsymbol{Z}_t^{(v)}, \boldsymbol{L}_t^{(v)}, \boldsymbol{E}_t^{(v)}, \mathcal{F}_t, \mathcal{G}_t, \boldsymbol{P}_t^{(v)}, \boldsymbol{Q}_t^{(v)}, \boldsymbol{Y}_t^{(v)})$ is generated by Algorithm 1, then

TABLE I
STATISTICS OF THE EXPERIMENTAL DATASETS

Datasets	Classes	Size	Views	Type
Yale	15	165	Gabor, Intensity, LBP	Image
ORL	40	400	Gabor, Intensity, LBP	Image
Extended YaleB	10	600	Gabor, Intensity, LBP	Image
Coil-20	20	1440	Gabor, Intensity, LBP	Image
WebKB	4	203	Page, Anchor text, Title	Webpage
Scene-15	15	4485	PHOW, PRI-CoLBP, CENTRIST	Image

1) any accumulation point of $\{\boldsymbol{H}_t\}_{t=1}^{\infty}$ is the Karush-Kuhn-Tucker (KKT) point, i.e., the accumulation point $\bar{\boldsymbol{H}} = (\bar{\boldsymbol{Z}}^{(v)}, \bar{\boldsymbol{L}}^{(v)}, \bar{\boldsymbol{E}}^{(v)}, \bar{\mathcal{F}}, \bar{\mathcal{G}}, \bar{\boldsymbol{P}}^{(v)}, \bar{\boldsymbol{Q}}^{(v)}, \bar{\boldsymbol{Y}}^{(v)})$ satisfies the following KKT conditions of (13):

$$\bar{\boldsymbol{Z}}^{(v)} = \bar{\boldsymbol{F}}^{(v)},\ \bar{\boldsymbol{L}}^{(v)} = \bar{\boldsymbol{P}}^{(v)},\ (\boldsymbol{X}^{(v)})^{\top}\bar{\boldsymbol{Y}}^{(v)} = \bar{\boldsymbol{G}}^{(v)},$$

$$\bar{\boldsymbol{Y}}^{(v)}(\boldsymbol{X}^{(v)})^{\top} = \bar{\boldsymbol{Q}}^{(v)}, \bar{\mathcal{G}} \in \partial\|\bar{\mathcal{F}}\|_{\omega,*}, \bar{\boldsymbol{Y}}(:,i) \in \lambda_2\partial\|\bar{\boldsymbol{E}}(:,i)\|_{2,1},$$

$$\boldsymbol{X}^{(v)} = \boldsymbol{X}^{(v)}\bar{\boldsymbol{Z}}^{(v)} + \bar{\boldsymbol{L}}^{(v)}\boldsymbol{X}^{(v)} + \bar{\boldsymbol{E}}^{(v)}, \bar{\boldsymbol{Q}}^{(v)} \in \partial\lambda_1 \left\|\bar{\boldsymbol{P}}^{(v)}\right\|_*.$$

2) $\{\boldsymbol{Z}_t^{(v)}\}, \{\boldsymbol{L}_t^{(v)}\}, \{\boldsymbol{E}_t^{(v)}\}, \{\mathcal{F}_t\}$, and $\{\boldsymbol{P}_t^{(v)}\}$ are Cauchy sequences and thus converge to the critical point of (13).

The proof can be found in Appendix B, available online.

Remark 4: In terms of convergence, the properties of our algorithm are similar to those of the work [21]. Nevertheless, the proof of convergence results of our case are more challenging. First, we consider a more general nonconvex transformed low-rank tensor approximation, making the convergence proof challenging. Second, we apply the latent low-rank representation to the multi-view case to resolve the insufficient sampling problem—that introduce many of auxiliary variables (i.e., $\boldsymbol{L}^{(v)}$, $\boldsymbol{P}^{(v)}$, and $\boldsymbol{Q}^{(v)}$). Therefore, our proof covers more cases that could not be covered by the convergence in [21].

V. Numerical Experiments

In this section, we present simulations on various real-world datasets to showcase the performance of the proposed algorithm. All simulations are performed using MATLAB 2018b on a desktop with 3.4 GHz i7 CPU and 64 GB RAM.

A. Datasets

The statistics of the six experimental datasets is summarized in Table I, and the detailed descriptions of the used datasets are presented as follows.

- *Yale*[1] *dataset* contains 165 grayscale images of 15 different subjects. Each subject has 11 images, and the images are obtained under various facial expressions or configurations. Following [68], we select three views, i.e., Gabor view, Intensity view, and LBP view.
- *ORL*[2] *dataset* consists of 40 different subjects, each of subject has 10 face images collected distinct time changing, facial details, and expressions. We choose three different views (Gabor view, Intensity view, and LBP view).
- *Extended YaleB*[3] *dataset* contains 38 individuals' face images. Each individual has around 60 images with various lighting conditions. As works in [10], [24], we use the first 10 classes with total 600 face images in the experiment. The Extended YaleB dataset has three views, i.e., Gabor view, Intensity view, and LBP view.
- *Coil-20*[4] *dataset* includes 1440 images of 20 objects obtained from varying angles. Each object has 72 images, in which each image is normalized to 32×32 following the works in [10], [31]. This dataset contains three different views Gabor, Intensity, and LBP.
- *WebKB*[5] *dataset* is a webpage set obtained by several universities. This data contains 203 samples belonging to different classes. Each sample has three views, including the page's content, the anchor text, and the title [69].
- *Scene-15*[6] *dataset* contains 4485 images of 15 natural scene categories with various indoor and outdoor environments. Similar to [24], [27], [70], we choose three views: PHOW, PRI-CoLBP, and CENTRIST.

B. Baselines

We choose state-of-the-art baselines for comparison. Since the proposed method is based on the multi-view subspace clustering framework, thus we select several subspace clustering methods, including affine subspace clustering [70] and dictionary representation-based subspace clustering methods [10], [24], [26], [40], [71], [72]. Besides, for comprehensive evaluation, we choose another two kinds of methods, i.e., Markov chain methods [20], [27] and latent representation methods [60], [73], [74]. Details of all baselines are shown in Appendix C in the supplemental material, available online.

For the proposed method, we choose two linear transforms, i.e., Discrete Cosine Transform (DCT), and the data-driven transform linked with the multi-view data. To be specific, we first obtain an estimated representation $\mathcal{Z}$ by the proposed DCT based-method and then apply the left singular matrix of the SVD of the mode-3 unfolding matrix of $\mathcal{Z}$ as the learned linear transform for MVSC. Meanwhile, we utilize two nonconvex functions, including Minimax Concave Penalty (MCP) [54] and Smoothly Clipped Absolute Deviation (SCAD) [55]. Our methods are named as ULatTLRR$_{\text{MCP}}^{\text{DCT}}$, ULatTLRR$_{\text{SCAD}}^{\text{DCT}}$, ULatTLRR$_{\text{MCP}}^{\text{Data}}$, and ULatTLRR$_{\text{SCAD}}^{\text{Data}}$. For all baselines, we do our best to turn the parameters to achieve the best clustering performance following the relevant papers' suggestions.

C. Quantitative Metrics

Following works [10], [24], to numerically evaluate the clustering performance, we employ six popular metrics, including accuracy (ACC), normalized mutual information (NMI), adjusted rank index (AR), F-score, Precision, and Recall. For all

[1] http://cvc.yale.edu/projects/yalefaces/yalefaces.html
[2] http://www.uk.research.att.com/facedatabase.html
[3] http://cvc.cs.yale.edu/cvc/projects/yalefacesB/yalefacesB.html
[4] http://www.cs.columbia.edu/CAVE/software/softlib/
[5] https://linqs.org/datasets/
[6] https://www.kaggle.com/datasets/zaiyankhan/15scene-dataset

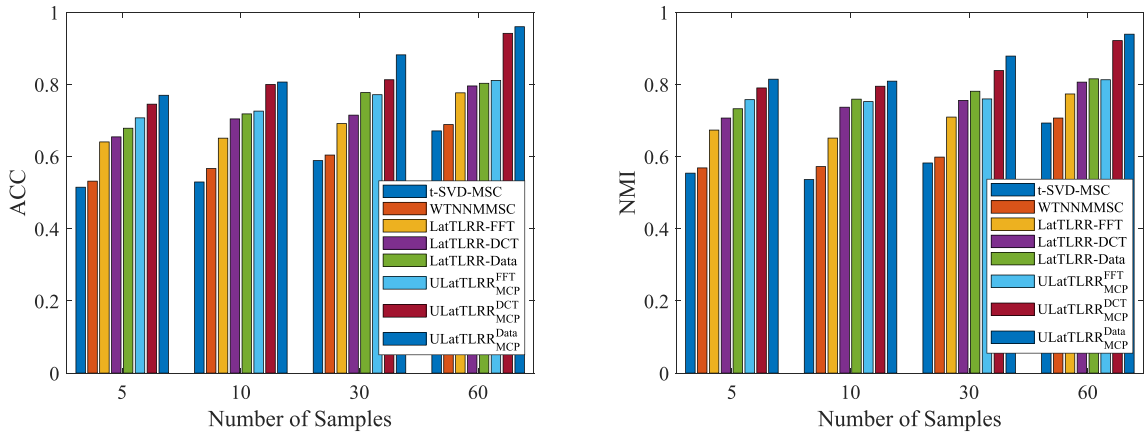


Fig. 2. Illustration of the effects of the latent representation, data-driven transform, and the unified nonconvex low-rank approximation.

of the metrics, the higher values indicate the better clustering quality. Note that all simulations are performed for 10 random trials, and we report the average results and standard deviations. We highlight the best values in bold.

D. Experimental Results

At first, we construct the experiments to showcase the motivations of our algorithm, i.e., the effects of the latent representation, data-driven transform, and the unified transformed nonconvex low-rank representation for the MUSC task. Here, we take the Extended YaleB as the texting dataset. Extended YaleB dataset has three views, each view contains 10 classes with 60 samples per class. To demonstrate the quality of our method in handling the insufficient sampling problem, we randomly choose S samples per class of all views, then each view has total of $10S$ observed data samples. In this experiment, we set $S = 5$, 10, and 30, respectively. To study the effects of data-driven transform and nonconvex low-rank approximation, we also set Fast Fourier Transform (FFT) as the transform and replace the nonconvex approximation in our model (12) with convex tensor nuclear norms induced different transforms. The convex methods are named as LatTLRR-FFT, LatTLRR-DCT, and LatTLRR-Data, respectively. We choose MCP function in our nonconvex methods, thus we denote our methods using FFT, DCT, and data-driven transform as ULatTLRR$^{FFT}_{MCP}$, ULatTLRR$^{DCT}_{MCP}$, and ULatTLRR$^{Data}_{MCP}$, respectively. We compare our methods with two TSVD-based tensor methods, i.e., t-SVD-MSC and WTNNMMSC.

Fig. 2 shows the clustering performance of different methods with different data samples in terms of ACC and NMI. From this figure, one can get the following insights:

- Compared t-SVD-MSC and LatTLRR-FFT, the only difference between these two methods is that LatTLRR-FFT considers the latent low-rank representation for multi-view data. When the number of the data sample is insufficient, i.e., $S = 5$, 10, and 30, the method LatTLRR-FFT achieves higher clustering values than t-SVD-MSC. This indicates that for multi-view clustering, with the help of latent low-rank representation, the proposed method can resolve the insufficient observed sample problem.
- Existing TSVD-based MVSC methods preserve the correlation in the Fourier domain, however, applying the FFT on $\mathcal{Z}$ requires that the tensor $\mathcal{Z}$ should have the periodic boundary, which may not hold in multi-view clustering. Hence, compared with LatTLRR-FFT, the methods with DCT and data-driven transform (i.e., LatTLRR-DCT and LatTLRR-Data) can obtain better clustering performance. We can also get the same result compared ULatTLRR$^{FFT}_{MCP}$ with ULatTLRR$^{DCT}_{MCP}$ and ULatTLRR$^{Data}_{MCP}$.
- From the results obtained by LatTLRR-FFT and ULatTLRR$^{FFT}_{MCP}$ (LatTLRR-DCT and ULatTLRR$^{DCT}_{MCP}$ or LatTLRR-Data and ULatTLRR$^{Data}_{MCP}$), one can see that the methods equipped with nonconvex MCP function can achieve higher values of ACC, NMI, and F-score. One reason is that the convex TNN-based methods treat all singular values equally—that ignore the prior information of the different singular values. Applying the nonconvex envelope to capture the priors of singular values can better explore the correlation among different views.

According to the aforementioned discussions, one can see that for all data sampling cases, only with the latent low-rank representation, the performance is already improved compared to t-SVD-MSC. With the linear transforms and nonconvex low-rank approximation, the performance can be further improved. This suggests the proposed method can resolve the insufficient sample problem in MVSC task and better preserve the correlation among different representation matrices in the transformed domain, thus improving the clustering performance.

Next, we report the clustering performance for all datasets.

Tables II and III list the clustering results in terms of six metrics on Yale and ORL datasets, respectively. One can see that the TNN-based tensor methods t-SVD-MSC and WTNNMMSC achieve the promising performance, e.g., 0.993 in ACC of t-SVD-MSC and 0.980 in ACC of WTNNMMSC for ORL dataset. The possible reasons are two-fold. First, the observed samples in the ORL dataset can provide extensive consensus and complementary information among different views. Therefore, many baselines achieve good performance in this dataset. Second, employing the tensor nuclear norm-based regularization on the representation tensor can fully capture the consensus

TABLE II
CLUSTERING PERFORMANCE ON YALE DATASET

Method	ACC	NMI	AR	F-score	Precision	Recall
RMSC [20]	0.566±0.000	0.607±0.000	0.383±0.000	0.422±0.000	0.412±0.000	0.432±0.000
CoMSC [71]	0.520±0.005	0.538±0.004	0.262±0.006	0.313±0.005	0.277±0.007	0.360±0.003
ETLMSC [27]	0.848±0.000	0.860±0.000	0.760±0.000	0.775±0.000	0.762±0.000	0.789±0.000
LMSC [60]	0.718±0.020	0.742±0.015	0.551±0.027	0.580±0.025	0.544±0.028	0.622±0.021
LTMSC [10]	0.738±0.002	0.763±0.003	0.598±0.004	0.624±0.003	0.602±0.003	0.647±0.004
MCLES [74]	0.740±0.006	0.761±0.006	0.587±0.009	0.614±0.008	0.584±0.011	0.647±0.007
OPMC [26]	0.500±0.059	0.566±0.036	0.323±0.045	0.370±0.040	0.336±0.050	0.412±0.031
JSTC [73]	0.612±0.000	0.615±0.000	0.449±0.000	0.418±0.000	0.395±0.000	0.433±0.000
ASR-ETR [72]	0.570±0.000	0.639±0.000	0.407±0.000	0.447±0.000	0.397±0.000	0.512±0.000
ARLRR-TU [70]	0.770±0.000	0.807±0.001	0.701±0.000	0.668±0.001	0.663±0.000	0.683±0.000
t-SVD-MSC [24]	0.955±0.021	0.957±0.015	0.920±0.028	0.925±0.027	0.915±0.031	0.935±0.022
WTNNMMSC [40]	0.968±0.017	0.973±0.011	0.949±0.021	0.952±0.020	0.944±0.026	0.960±0.014
ULatTLRR$^{DCT}_{MCP}$	0.999±0.002	0.999±0.002	0.999±0.003	0.999±0.002	0.999±0.003	0.998±0.003
ULatTLRR$^{DCT}_{SCAD}$	1.000±0.001	0.999±0.001	0.999±0.002	1.000±0.001	0.999±0.002	0.999±0.002
ULatTLRR$^{Data}_{MCP}$	0.999±0.002	0.999±0.003	0.999±0.004	0.999±0.002	0.998±0.005	0.998±0.005
ULatTLRR$^{Data}_{SCAD}$	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**

TABLE III
CLUSTERING PERFORMANCE ON ORL DATASET

Method	ACC	NMI	AR	F-score	Precision	Recall
RMSC [20]	0.715±0.000	0.837±0.000	0.604±0.000	0.613±0.000	0.583±0.000	0.646±0.000
CoMSC [71]	0.743±0.008	0.870±0.005	0.654±0.012	0.663±0.011	0.609±0.013	0.727±0.011
ETLMSC [27]	0.840±0.000	0.908±0.000	0.763±0.000	0.768±0.000	0.727±0.000	0.815±0.000
LMSC [60]	0.813±0.022	0.904±0.016	0.748±0.033	0.754±0.033	0.716±0.029	0.796±0.039
LTMSC [10]	0.826±0.010	0.917±0.006	0.769±0.015	0.775±0.015	0.733±0.018	0.821±0.011
MCLES [74]	0.841±0.012	0.917±0.005	0.778±0.013	0.784±0.013	0.751±0.013	0.820±0.014
OPMS [26]	0.554±0.020	0.744±0.013	0.412±0.018	0.427±0.016	0.378±0.015	0.490±0.021
JSTC [73]	0.652±0.000	0.824±0.000	0.621±0.000	0.475±0.000	0.526±0.000	0.539±0.000
ASR-ETR [72]	0.882±0.000	0.836±0.000	0.865±0.000	0.868±0.000	0.397±0.000	0.903±0.000
ARLRR-TU [70]	0.959±0.004	0.985±0.003	0.966±0.009	0.937±0.004	0.950±0.006	0.951±0.006
t-SVD-MSC [24]	0.993±0.007	0.998±0.002	0.992±0.007	0.992±0.007	0.986±0.012	0.998±0.002
WTNNMSC [40]	0.980±0.005	0.994±0.002	0.979±0.005	0.979±0.005	0.965±0.008	0.994±0.002
ULatTLRR$^{DCT}_{MCP}$	0.999±0.001	0.999±0.001	0.999±0.002	0.999±0.001	0.998±0.003	0.998±0.003
ULatTLRR$^{DCT}_{SCAD}$	0.996±0.011	0.999±0.004	0.999±0.003	0.996±0.009	0.996±0.012	0.992±0.020
ULatTLRR$^{Data}_{MCP}$	0.999±0.002	0.999±0.002	0.998±0.004	0.999±0.002	0.998±0.004	0.998±0.004
ULatTLRR$^{Data}_{SCAD}$	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**	**1.000±0.000**

TABLE IV
CLUSTERING PERFORMANCE ON EXTENDED YALEB DATASET

Method	ACC	NMI	AR	F-score	Precision	Recall
RMSC [20]	0.282±0.000	0.225±0.000	0.104±0.000	0.194±0.000	0.189±0.000	0.200±0.000
CoMSC [71]	0.608±0.001	0.581±0.001	0.348±0.001	0.421±0.001	0.367±0.001	0.494±0.001
ETLMSC [27]	0.456±0.005	0.432±0.010	0.304±0.009	0.373±0.008	0.370±0.008	0.377±0.008
LMSC [60]	0.458±0.020	0.408±0.024	0.184±0.010	0.279±0.009	0.235±0.009	0.343±0.022
LTMSC [10]	0.486±0.001	0.475±0.002	0.275±0.003	0.351±0.002	0.328±0.003	0.378±0.002
MCLES [74]	0.479±0.006	0.443±0.010	0.198±0.002	0.294±0.002	0.243±0.003	0.370±0.001
OPMC [26]	0.291±0.019	0.279±0.029	0.127±0.018	0.219±0.016	0.205±0.016	0.236±0.022
JSTC [73]	0.220±0.000	0.148±0.000	0.159±0.000	0.140±0.000	0.049±0.000	0.148±0.000
ASR-ETR [72]	0.512±0.000	0.510±0.000	0.349±0.000	0.415±0.000	0.403±0.000	0.427±0.000
ARLRR-TU [70]	0.739±0.000	0.748±0.003	0.670±0.009	0.646±0.004	0.620±0.006	0.658±0.006
t-SVD-MSC [24]	0.672±0.003	0.694±0.002	0.521±0.003	0.570±0.003	0.547±0.002	0.594±0.003
WTNNMMSC [40]	0.689±0.002	0.706±0.003	0.531±0.004	0.580±0.004	0.552±0.004	0.611±0.003
ULatTLRR$^{DCT}_{MCP}$	0.942±0.000	0.922±0.000	0.897±0.000	0.942±0.000	0.892±0.001	0.886±0.001
ULatTLRR$^{DCT}_{SCAD}$	0.940±0.000	0.912±0.000	0.894±0.000	0.940±0.000	0.886±0.000	0.878±0.000
ULatTLRR$^{Data}_{MCP}$	**0.960±0.000**	**0.939±0.000**	**0.926±0.000**	**0.960±0.000**	**0.923±0.001**	**0.920±0.001**
ULatTLRR$^{Data}_{SCAD}$	0.937±0.002	0.897±0.003	0.884±0.003	0.937±0.002	0.879±0.003	0.876±0.003

TABLE V
CLUSTERING PERFORMANCE ON COIL-20 DATASET

Method	ACC	NMI	AR	F-score	Precision	Recall
RMSC [20]	0.740±0.004	0.815±0.004	0.695±0.008	0.687±0.009	0.704±0.007	0.679±0.008
CoMSC [71]	0.647±0.008	0.746±0.004	0.543±0.007	0.568±0.006	0.513±0.008	0.637±0.004
ETLMSC [27]	0.833±0.005	0.885±0.003	0.798±0.006	0.808±0.005	0.803±0.007	0.813±0.004
LMSC [60]	0.745±0.014	0.826±0.007	0.668±0.013	0.685±0.012	0.644±0.019	0.733±0.016
LTMSC [10]	0.729±0.012	0.813±0.005	0.667±0.014	0.684±0.013	0.660±0.017	0.710±0.009
MCLES [74]	0.794±0.004	0.864±0.003	0.753±0.006	0.765±0.006	0.745±0.007	0.787±0.004
OPMS [26]	0.554±0.062	0.729±0.017	0.518±0.033	0.546±0.030	0.470±0.044	0.655±0.007
JSTC [73]	0.689±0.000	0.821±0.000	0.730±0.000	0.625±0.000	0.655±0.000	0.673±0.000
ASR-ETR [72]	0.903±0.000	0.933±0.000	0.873±0.000	0.879±0.000	0.845±0.000	0.916±0.000
ARLRR-TU [70]	0.977±0.024	0.976±0.009	0.965±0.018	0.960±0.029	0.961±0.025	0.963±0.006
t-SVD-MSC [24]	0.804±0.002	0.866±0.001	0.758±0.002	0.770±0.002	0.755±0.003	0.786±0.002
WTNNMSC [40]	0.826±0.004	0.875±0.002	0.772±0.001	0.783±0.001	0.767±0.009	0.801±0.008
ULatTLRR$^{DCT}_{MCP}$	0.957±0.006	0.967±0.002	0.939±0.002	0.957±0.006	0.932±0.004	0.925±0.007
ULatTLRR$^{DCT}_{SCAD}$	0.957±0.001	0.960±0.001	0.931±0.001	0.957±0.001	0.924±0.001	0.916±0.001
ULatTLRR$^{Data}_{MCP}$	0.977±0.007	0.977±0.005	**0.977±0.007**	0.959±0.010	0.956±0.012	0.963±0.009
ULatTLRR$^{Data}_{SCAD}$	**0.982±0.001**	**0.980±0.001**	0.969±0.001	**0.982±0.001**	**0.967±0.001**	**0.965±0.001**

TABLE VI
CLUSTERING PERFORMANCE ON WEBKB DATASET

Method	ACC	NMI	AR	F-score	Precision	Recall
RMSC [20]	0.538±0.016	0.227±0.027	0.483±0.023	0.604±0.028	0.403±0.020	0.246±0.033
CoMSC [71]	0.744±0.000	0.378±0.000	0.428±0.000	0.691±0.000	0.591±0.000	0.731±0.000
ETLMSC [27]	0.631±0.000	0.256±0.000	0.329±0.000	0.567±0.000	0.623±0.000	0.516±0.000
LMSC [60]	0.670±0.000	0.312±0.000	0.374±0.000	0.600±0.000	0.653±0.000	0.555±0.000
LTMSC [10]	0.537±0.001	0.153±0.000	0.179±0.001	0.470±0.001	0.522±0.000	0.428±0.001
MCLES [74]	0.650±0.013	0.233±0.016	0.301±0.018	0.563±0.014	0.590±0.017	0.539±0.028
OPMC [26]	0.652±0.039	0.391±0.045	0.424±0.055	0.621±0.046	0.713±0.039	0.553±0.067
JSTC [73]	0.621±0.000	0.255±0.000	0.484±0.000	0.585±0.000	0.271±0.000	0.530±0.000
ASR-ETR [72]	0.744±0.000	0.403±0.000	0.460±0.000	0.670±0.000	0.677±0.000	0.663±0.000
ARLRR-TU [70]	0.753±0.000	0.475±0.000	0.687±0.000	0.725±0.000	0.523±0.000	0.706±0.000
t-SVD-MSC [24]	0.714±0.000	0.409±0.000	0.450±0.000	0.665±0.000	0.670±0.000	0.661±0.000
WTNNMMSC [40]	0.719±0.000	0.416±0.000	0.456±0.000	0.669±0.000	0.672±0.000	0.667±0.000
ULatTLRR$^{DCT}_{MCP}$	0.759±0.000	0.438±0.001	0.719±0.000	0.793±0.000	0.666±0.000	0.782±0.000
ULatTLRR$^{DCT}_{SCAD}$	0.769±0.000	**0.481±0.001**	0.731±0.000	**0.823±0.000**	**0.740±0.000**	0.721±0.000
ULatTLRR$^{Data}_{MCP}$	**0.778±0.000**	0.447±0.007	**0.737±0.003**	0.803±0.000	0.673±0.004	**0.813±0.000**
ULatTLRR$^{Data}_{SCAD}$	0.773±0.000	0.458±0.000	0.733±0.000	0.798±0.000	0.670±0.000	0.809±0.000

TABLE VII
CLUSTERING PERFORMANCE ON SCENE-15 DATASET

Method	ACC	NMI	AR	F-score	Precision	Recall
RMSC [20]	0.425±0.004	0.434±0.001	0.254±0.000	0.307±0.000	0.302±0.000	0.312±0.000
CoMSC [71]	0.495±0.002	0.499±0.001	0.345±0.001	0.393±0.001	0.369±0.002	0.420±0.001
ETLMSC [27]	0.854±0.000	0.902±0.000	0.811±0.000	0.824±0.000	0.827±0.000	0.821±0.000
LMSC [60]	0.524±0.000	0.480±0.000	0.346±0.000	0.393±0.000	0.378±0.000	0.411±0.000
LTMSC [10]	0.568±0.000	0.571±0.000	0.424±0.000	0.465±0.000	0.453±0.000	0.478±0.000
MCLES [74]	–	–	–	–	–	–
OPMS [26]	0.500±0.062	0.532±0.017	0.344±0.033	0.392±0.030	0.372±0.044	0.414±0.007
JSTC [73]	0.498±0.000	0.543±0.000	0.429±0.000	0.359±0.000	0.342±0.000	0.391±0.000
ASR-ETR [72]	0.953±0.000	0.938±0.000	0.920±0.000	0.926±0.000	0.922±0.000	0.930±0.000
ARLRR-TU [70]	0.976±0.000	0.953±0.000	0.953±0.000	0.956±0.000	0.951±0.000	0.955±0.000
t-SVD-MSC [24]	0.809±0.002	0.825±0.001	0.749±0.002	0.767±0.002	0.734±0.003	0.804±0.002
WTNNMSC [40]	0.904±0.004	0.929±0.002	0.894±0.001	0.901±0.001	0.886±0.009	0.917±0.008
ULatTLRR$^{DCT}_{MCP}$	**0.984±0.000**	**0.969±0.000**	**0.971±0.000**	**0.984±0.000**	**0.974±0.000**	**0.968±0.000**
ULatTLRR$^{DCT}_{SCAD}$	0.981±0.000	0.962±0.000	0.963±0.000	0.981±0.000	0.966±0.000	0.960±0.000
ULatTLRR$^{Data}_{MCP}$	0.964±0.002	0.937±0.003	0.934±0.003	0.964±0.002	0.937±0.003	0.931±0.003
ULatTLRR$^{Data}_{SCAD}$	0.969±0.000	0.945±0.000	0.943±0.000	0.969±0.000	0.946±0.000	0.939±0.000

"–" denotes that the method exceeded the time limit.

and complementary information among all views, so that two methods (i.e., t-SVD-MSC and WTNNMSC) based on tensor nuclear norm achieve better performance compared with another baselines. However, our methods can obtain the nearly perfect clustering performance in terms of all six metrics, which means that for these two datasets, the proposed approaches still perform better than all baselines.

Tables IV and V present the clustering metrics by all methods on Extended YaleB and Coil-20 datasets, respectively. One can see that the proposed methods yield the best clustering performace over all compared methods. Note that for these two datasets, our methods achieve a significant improvement compared to the baselines. The proposed methods obtain the clear improvements around 25% and 15% in terms of ACC for Extended YaleB and Coil-20 datasets, respectively, over the second best WTNNMMSC.

Tables VI and VII give the clustering results on Webkb and Scene-15 datasets, respectively. Again, the proposed methods outperform all baselines in most metrics. One can see that both methods, including our approaches and all baselines, perform not very well for the Webkb dataset. The reason may be that

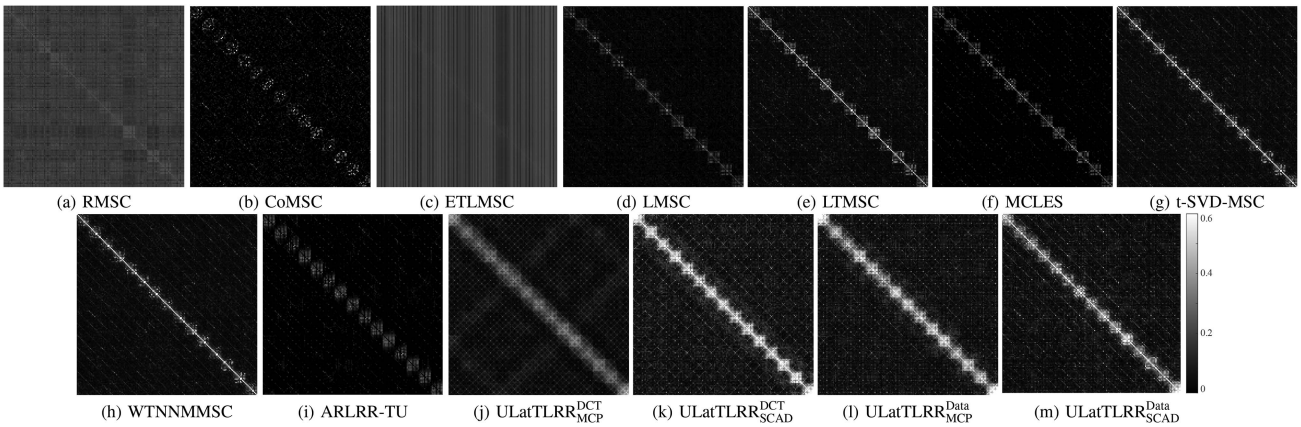


Fig. 3. The affinity matrices by all methods on Yale dataset.

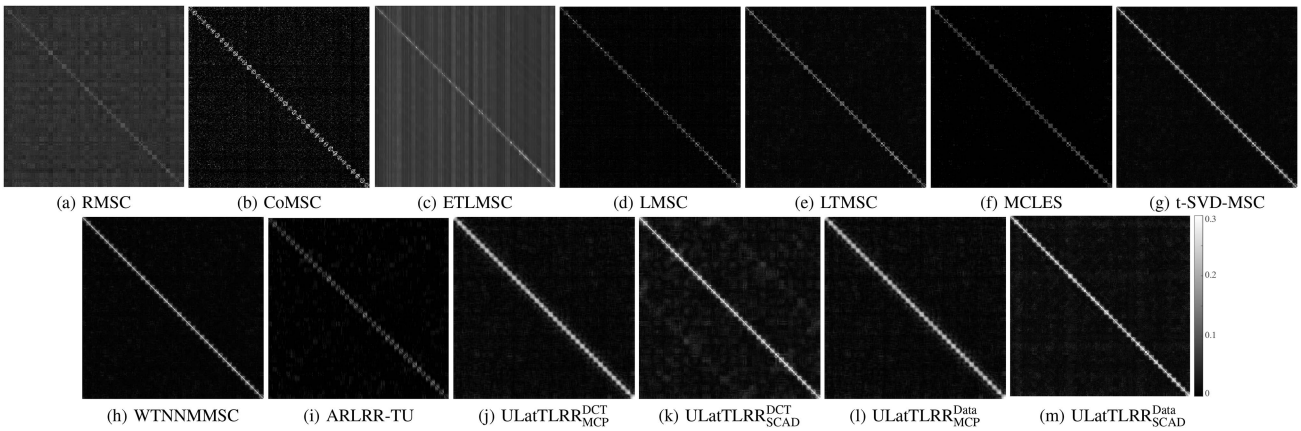


Fig. 4. The affinity matrices by all methods on ORL dataset.

Webkb is a web page dataset, it is hard to explore the high correlation among different views for all methods. One potential way to improve the clustering performance is to design the suitable regularization according to the data structure of Webkb for better capturing the high correlation.

As shown in Tables II—VII, one can see that our methods with the data-adaptive transform can achieve the higher results than the methods with DCT transform in most cases. It indicates that the data-adaptive transform based methods can preserve the high correlation and complementary information among all views in the transformed domain w.r.t. the intrinsic data structure, better than the methods with fixed transforms.

Figs. 3 and 4 compare the affinity matrices obtained by all methods on Yale and ORL datasets, respectively. Since the methods OPMC, JSTC, and ASR-ETR output vectors of clustering indexes, so we omit them. For the Yale dataset, one can see that the affinity matrices by our methods admit the clearer block-diagonal structure than the baselines. This shows that our methods can better explore the correlation among the samples belonging to the same object and distinguish the difference between samples of different objects. For the ORL dataset, one can see that the structure of affinity matrices by t-SVD-MSC and WTNNMMSC are similar to ours, and thus the performance of t-SVD-MSC, WTNNMMSC, and our method are comparable to each other—that is also demonstrated by the metrics shown in Table III.

In Fig. 5, we present the view-specific affinity matrices and the final matrices of t-SVD-MSC, WTNNMMSC, and our method for the Yale dataset. One can see that: 1) for both view-specific or final affinity matrices, our methods can achieve a clearer block-diagonal structure than t-SVD-MSC and WTNN-MMSC; 2) The block-diagonal structure of the view-specific affinity matrices is different from each other, which means that different views play a specific role in clustering the data samples.

E. Advantages of Our Methods Over Single-View Clustering Methods

As mentioned before, the insufficient observed sample problem has been studied in the single-view subspace clustering [30], [44], [45], [46], [47], which are designed to handle the single-view data. When directly applying these methods to multi-view data, it is essential to handle each view separately. Compared with the single-view data, the multi-view data contains multiple different features, and each feature represents the

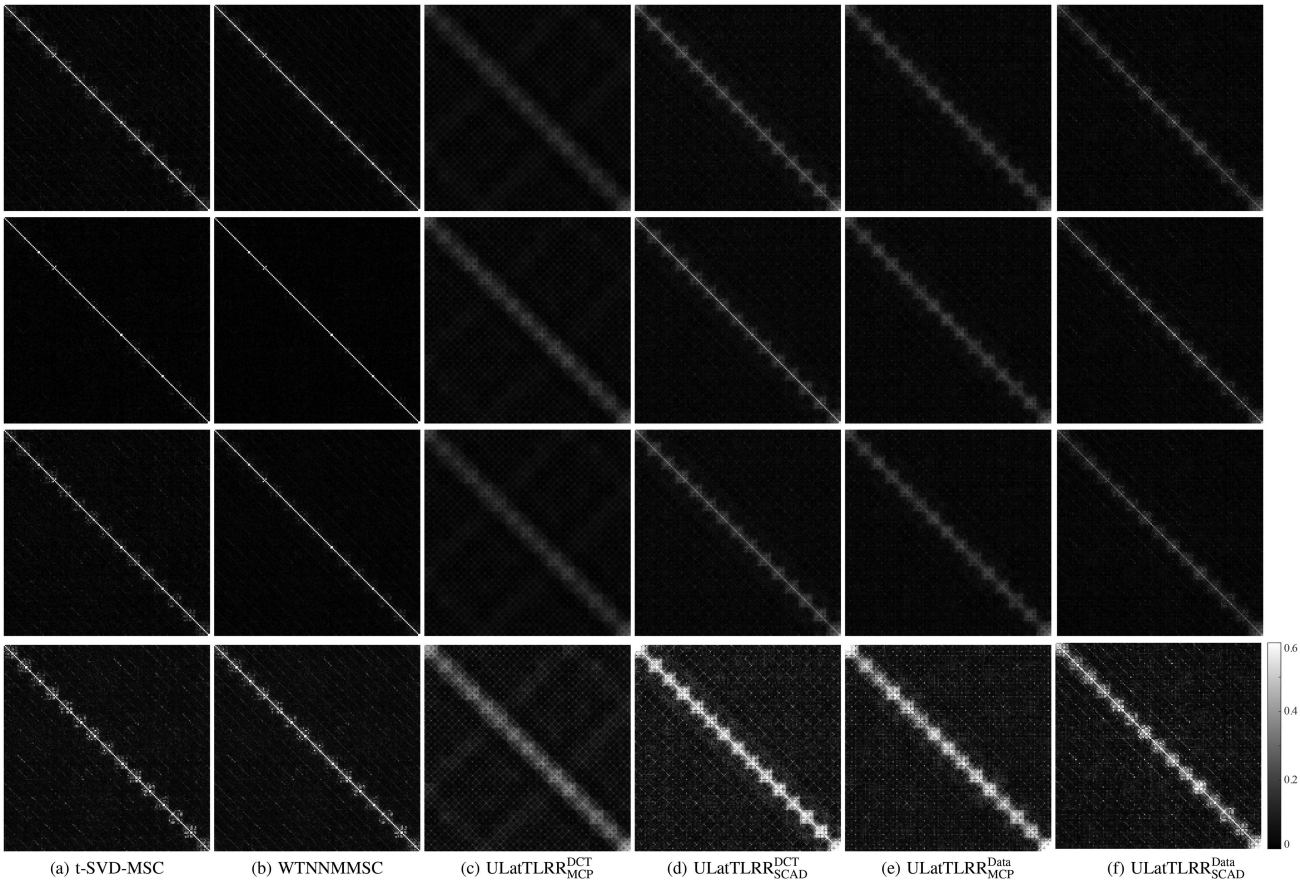


Fig. 5. The affinity matrices on each view $\boldsymbol{S}^{(v)} = (|\boldsymbol{Z}^{(v)}_{t+1}| + |(\boldsymbol{Z}^{(v)}_{t+1})^\top|)/2$ and the final version $\boldsymbol{S} = (\sum_{v=1}^{V} |\boldsymbol{Z}^{(v)}_{t+1}| + |(\boldsymbol{Z}^{(v)}_{t+1})^\top|)/V$ by all methods (from top to bottom: $\boldsymbol{S}^{(1)}$, $\boldsymbol{S}^{(2)}$, $\boldsymbol{S}^{(3)}$, and $\boldsymbol{S}$) on Yale dataset.

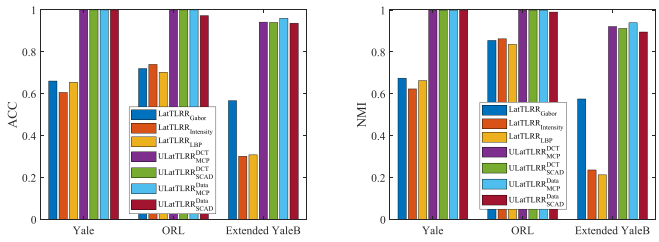


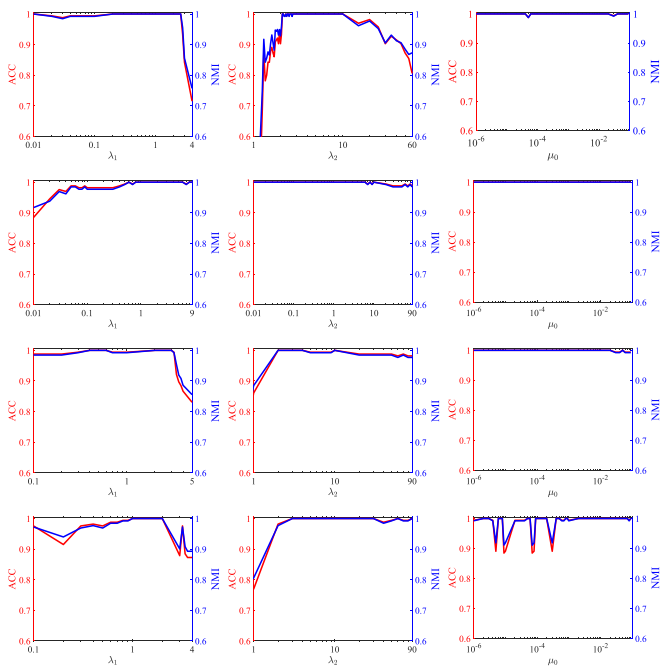
Fig. 6. Comparison between LatLRR with the single view and our methods in terms of ACC and NMI. LatLRR$_{\text{Gabor}}$, LatLRR$_{\text{Intensity}}$, and LatLRR$_{\text{LBP}}$ denote the applications of LatLRR on the Gabor, Intensity, and LBP view data, respectively.

particular information. Directly applying single-view methods on the multi-view data ignores consensus and complementary information among multiple views, leading to unsatisfactory performance.

Taking the method LatLRR [30], one of SOTA methods handling the single-view insufficient sample problem, as an example, Fig. 6 shows the results of LatLRR using each single view on Yale, ORL, Extended YaleB datasets. One can see that although LatLRR works to a certain extent, the performance of LatLRR using different views vary significantly. For example, on ORL dataset, LatLRR achieves best performance in Intensity view, while there is a serious degeneration on Extended YaleB dataset. On the contrary, the proposed method directly used all views and achieves almost perfect performance, which demonstrates that our model can simultaneously resolve the insufficient data problem and capture the consensus and complementary information among multiple views.

F. Effects of Parameter Settings

In the proposed algorithm, there are two regularization parameters λ_1 and λ_2 and three initial penalty parameters μ_0, ρ_0, and θ_0 need to be turned. In experiments, we set $\mu_0 = \rho_0 = \theta_0$. Fig. 7 shows the clustering performance in terms of ACC and NMI w.r.t. different choices of λ_1, λ_2, and μ_0 (i.e., ρ_0 and θ_0) on Yale dataset. One can see that our methods achieve nearly perfect high metrics in a relatively wide range of three parameters, which indicates that our methods are stable while changing λ_1, λ_2, and μ_0. In the experiments, we empirically set $\mu_0 = \rho_0 = \theta_0 = 10^{-4}$, and select λ_1 and λ_2 from the values in $\{0.1, 0.3, 0.5, 0.8, 1, 3, 5, 8\}$.



(a) Parameter λ_1 (b) Parameter λ_2 (c) Parameter μ_0

Fig. 7. The clustering performance in terms of ACC and NMI for the turning parameters λ_1, λ_2, and μ_0 on the Yale dataset. From top to bottom: ULatTLRR$^{\text{DCT}}_{\text{MCP}}$, ULatTLRR$^{\text{DCT}}_{\text{SCAD}}$, ULatTLRR$^{\text{Data}}_{\text{MCP}}$, and ULatTLRR$^{\text{Data}}_{\text{SCAD}}$.

VI. Conclusion

In this work, we develop a nonconvex latent transformed low-rank tensor model for MVSC. To solve the insufficient data sampling problem in existing LRR-based tensor clustering methods, we introduce the latent low-rank representation into the multi-view case. By exploiting the prior information of the singular values of the stacked representation tensor and the intrinsic structure of data, we develop a unified nonconvex transformed low-rank tensor representation for better exploring the high correlation among different views. An ADMM-based algorithm is proposed with guaranteed convergence [75], [76], [77], [78], [79]. Extensive numerical experiments show that our proposed method outperforms the existing contemporary multi-view clustering approaches. As the future research works, it is interesting to come up with algorithms that can reduce the running time and apply the irregular tensor factorization [64], [65], [66] on the irregular tensor stacked by $\{\boldsymbol{L}^{(v)}\}_{v=1}^{V}$ to exploit the correlation among all salient representations.

References

[1] C. Xu, D. Tao, and C. Xu, "A survey on multi-view learning," 2013, *arXiv:1304.5634*.

[2] K. Chaudhuri, S. M. Kakade, K. Livescu, and K. Sridharan, "Multi-view clustering via canonical correlation analysis," in *Proc. Int. Conf. Mach. Learn.*, 2009, pp. 129–136.

[3] C. Tang, Z. Li, J. Wang, X. Liu, W. Zhang, and E. Zhu, "Unified one-step multi-view spectral clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 6, pp. 6449–6460, Jun. 2023.

[4] M. Ozay, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Fusion of image segmentation algorithms using consensus clustering," in *Proc. IEEE Int. Conf. Image Process.*, 2013, pp. 4049–4053.

[5] H. Wang, F. Nie, and H. Huang, "Multi-view clustering and feature learning via structured sparsity," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 352–360.

[6] A. Djelouah, J.-S. Franco, E. Boyer, F. L. Clerc, and P. Pérez, "Multi-view object segmentation in space and time," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2013, pp. 2640–2647.

[7] J.-H. Yang, C. Chen, H.-N. Dai, M. Ding, Z.-B. Wu, and Z. Zheng, "Robust corrupted data recovery and clustering via generalized transformed tensor low-rank representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 7, pp. 8839–8853, Jul. 2024.

[8] P. Zhang et al., "Consensus one-step multi-view subspace clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 10, pp. 4676–4689, Oct. 2022.

[9] H. Gao, F. Nie, X. Li, and H. Huang, "Multi-view subspace clustering," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 4238–4246.

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

[11] L. Liu, F. Nie, A. Wiliem, Z. Li, T. Zhang, and B. C. Lovell, "Multi-modal joint clustering with application for unsupervised attribute discovery," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4345–4356, Sep. 2018.

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

[13] R. Chen, Y. Tang, Y. Xie, W. Feng, and W. Zhang, "Semisupervised progressive representation learning for deep multiview clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 10, pp. 14341–14355, Oct. 2024.

[14] X. Tian et al., "Variational distillation for multi-view learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 7, pp. 4551–4566, Jul. 2024.

[15] G. Chao, S. Sun, and J. Bi, "A survey on multiview clustering," *IEEE Trans. Artif. Intell.*, vol. 2, no. 2, pp. 146–168, Apr. 2021.

[16] Y. Yang and H. Wang, "Multi-view clustering: A survey," *Big Data Mining Analytics*, vol. 1, no. 2, pp. 83–107, 2018.

[17] R. Li, C. Zhang, H. Fu, X. Peng, J. T. Zhou, and Q. Hu, "Reciprocal multi-layer subspace learning for multi-view clustering," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 8171–8179.

[18] X. Zhang, X. Zhang, H. Liu, and X. Liu, "Multi-task multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 12, pp. 3324–3338, Dec. 2016.

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

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

[21] M. Yang, Q. Luo, W. Li, and M. Xiao, "Multiview clustering of images with tensor rank minimization via nonconvex approach," *SIAM J. Imag. Sci.*, vol. 13, no. 4, pp. 2361–2392, 2020.

[22] Z. Yu et al., "Incremental semi-supervised clustering ensemble for high dimensional data clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 3, pp. 701–714, Mar. 2016.

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

[24] Y. Xie, D. Tao, W. Zhang, Y. Liu, L. Zhang, and Y. Qu, "On unifying multi-view self-representations for clustering by tensormulti-rank minimization," *Int. J. Comput. Vis.*, vol. 126, pp. 1157–1179, 2018.

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

[26] J. Liu et al., "One-pass multi-view clustering for large-scale data," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 12324–12333.

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

[28] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.

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

[30] G. Liu and S. Yan, "Latent low-rank representation for subspace segmentation and feature extraction," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 1615–1622.

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

[32] Y. Chen, S. Wang, C. Peng, Z. Hua, and Y. Zhou, "Generalized nonconvex low-rank tensor approximation for multi-view subspace clustering," *IEEE Trans. Image Process.*, vol. 30, pp. 4022–4035, 2021.
[33] Y. Tang, Y. Xie, C. Zhang, and W. Zhang, "Constrained tensor representation learning for multi-view semi-supervised subspace clustering," *IEEE Trans. Multimedia*, vol. 24, pp. 3920–3933, 2022.
[34] Y. Chen, X. Xiao, Z. Hua, and Y. Zhou, "Adaptive transition probability matrix learning for multiview spectral clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4712–4726, Sep. 2022.
[35] S. Wang, Y. Chen, Z. Lin, Y. Cen, and Q. Cao, "Robustness meets low-rankness: Unified entropy and tensor learning for multi-view subspace clustering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 11, pp. 6302–6316, Nov. 2023.
[36] M.-S. Chen, C.-D. Wang, and J.-H. Lai, "Low-rank tensor based proximity learning for multi-view clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 5, pp. 5076–5090, May 2023.
[37] S. Wang, Z. Lin, Q. Cao, Y. Cen, and Y. Chen, "Bi-nuclear tensor schatten-p norm minimization for multi-view subspace clustering," *IEEE Trans. Image Process.*, vol. 32, pp. 4059–4072, 2023.
[38] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, Apr. 2020.
[39] Y. Chen, X. Xiao, C. Peng, G. Lu, and Y. Zhou, "Low-rank tensor graph learning for multi-view subspace clustering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 1, pp. 92–104, Jan. 2022.
[40] Q. Gao, W. Xia, Z. Wan, D. Xie, and P. Zhang, "Tensor-SVD based graph learning for multi-view subspace clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 3930–3937.
[41] Y. Tang, Y. Xie, X. Yang, J. Niu, and W. Zhang, "Tensor multi-elastic kernel self-paced learning for time series clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 3, pp. 1223–1237, Mar. 2021.
[42] Y. Xie et al., "Robust kernelized multiview self-representation for subspace clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 868–881, Feb. 2021.
[43] Y. Xie, W. Zhang, Y. Qu, L. Dai, and D. Tao, "Hyper-laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning," *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 572–586, Feb. 2020.
[44] R. Liu, Z. Lin, F. De la Torre, and Z. Su, "Fixed-rank representation for unsupervised visual learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 598–605.
[45] Z. Zhang, S. Yan, and M. Zhao, "Similarity preserving low-rank representation for enhanced data representation and effective subspace learning," *Neural Netw.*, vol. 53, pp. 81–94, 2014.
[46] S. Yu and Y. Wu, "Subspace clustering based on latent low rank representation with frobenius norm minimization," *Neurcomputing*, vol. 275, pp. 2479–2489, 2018.
[47] Z. Fu, Y. Zhao, D. Chang, Y. Wang, and J. Wen, "Latent low-rank representation with weighted distance penalty for clustering," *IEEE Trans. Cybern.*, vol. 53, no. 11, pp. 6870–6882, Nov. 2023.
[48] H. Li, X.-J. Wu, and J. Kittler, "MDLatLRR: A novel decomposition method for infrared and visible image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 4733–4746, 2020.
[49] P. Jing, J. Zhang, L. Nie, S. Ye, J. Liu, and Y. Su, "Tripartite graph regularized latent low-rank representation for fashion compatibility prediction," *IEEE Trans. Multimedia*, vol. 24, pp. 1277–1287, 2022.
[50] M. Kilmer and C. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, 2011.
[51] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
[52] C. Lu, X. Peng, and Y. Wei, "Low-rank tensor completion with a new tensor nuclear norm induced by invertible linear transforms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5989–5997.
[53] I. E. Frank and J. H. Friedman, "A statistical view of some chemometrics regression tools," *Technometrics*, vol. 35, no. 2, pp. 109–135, 1993.
[54] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, vol. 38, no. 2, pp. 894–942, 2010.
[55] J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Amer. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.
[56] J. H. Friedman, "Fast sparse regression and classification," *Int. J. Forecasting*, vol. 28, no. 3, pp. 722–738, 2012.
[57] J. Trzasko and A. Manduca, "Highly undersampled magnetic resonance image reconstruction via homotopic ℓ_0 -minimization," *IEEE Trans. Med. Imag.*, vol. 28, no. 1, pp. 106–121, Jan. 2009.
[58] C. Lu, J. Tang, S. Yan, and Z. Lin, "Generalized nonconvex nonsmooth low-rank minimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1–8.
[59] C. Lu, J. Tang, S. Yan, and Z. Lin, "Nonconvex nonsmooth low rank minimization via iteratively reweighted nuclear norm," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 829–839, Feb. 2016.
[60] C. Zhang, Q. Hu, H. Fu, P. Zhu, and X. Cao, "Latent multi-view subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4333–4341.
[61] M. Yin, J. Gao, S. Xie, and Y. Guo, "Multiview subspace clustering via tensorial t-product representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 3, pp. 851–864, Mar. 2019.
[62] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, pp. 1–122, 2011.
[63] T. Lin, S. Ma, and S. Zhang, "On the global linear convergence of the ADMM with multiblock variables," *SIAM J. Optim.*, vol. 25, no. 3, pp. 1478–1497, 2015.
[64] H. A. L. Kiers, J. M. F. Ten Berge, and R. Bro, "PARAFAC2-Part I. A direct fitting algorithm for the PARAFAC2 model," *J. Chemometrics*, vol. 13, no. 3/4, pp. 275–294, 1999.
[65] Y. Ren, J. Lou, L. Xiong, J. C. Ho, X. Jiang, and S. V. Bhavani, "MULTIPAR: Supervised irregular tensor factorization with multi-task learning for computational phenotyping," in *Proc. Mach. Learn. Health Symp.*, 2023, pp. 498–511.
[66] H. Zhang et al., "Learnable transform-assisted tensor decomposition for spatio-irregular multidimensional data recovery," *ACM Trans. Knowl. Discov. Data*, vol. 19, no. 1, 2024, Art. no. 12.
[67] J. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.
[68] S. Luo, C. Zhang, W. Zhang, and X. Cao, "Consistent and specific multi-view subspace clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 3730–3737.
[69] Q. Kun, S. E. Abhadiomhen, and Z. Liu, "Multiview subspace clustering via low-rank correlation analysis," *IET Comput. Vis.*, 2022. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cvi2.12155
[70] Y. Tang, Y. Xie, and W. Zhang, "Affine subspace robust low-rank self-representation: From matrix to tensor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 9357–9373, Aug. 2023.
[71] J. Liu, X. Liu, Y. Yang, X. Guo, M. Kloft, and L. He, "Multiview subspace clustering via co-training robust data representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5177–5189, Oct. 2022.
[72] J. Ji and S. Feng, "Anchor structure regularization induced multi-view subspace clustering via enhanced tensor rank minimization," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2023, pp. 19286–19295.
[73] Y. Tang, Y. Xie, C. Zhang, Z. Zhang, and W. Zhang, "One-step multiview subspace segmentation via joint skinny tensor learning and latent clustering," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9179–9193, Sep. 2022.
[74] M. Chen, L. Huang, C. Wang, and D. Huang, "Multi-view clustering in latent embedding space," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 3513–3520.
[75] K. Chen, H. Dong, and K.-S. Chan, "Reduced rank regression via adaptive nuclear norm penalization," *Biometrika*, vol. 100, no. 4, pp. 901–920, 2013.
[76] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2013, *arXiv:1009.5055v3*.
[77] A. S. Lewis and H. S. Sendov, "Nonsmooth analysis of singular values. Part I: Theory," *Set-Valued Anal.*, vol. 13, pp. 213–241, 2005.
[78] D. G. Luenberger, *Optimization by Vector Space Methods*. New York, NY, USA: Wiley, 1997.
[79] G. A. Watson, "Characterization of the subdifferential of some matrix norms," *Linear Algebra Appl.*, vol. 170, pp. 33–45, 1992.

Meng Ding received the PhD degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2021. From 2019 to 2020, he was an exchange PhD Student with Oregon State University, Corvallis, Oregon, supported by the China Scholarship Council. He is currently an associate professor with the School of Mathematics, Southwest Jiaotong University, Chengdu. His current research interests include image processing, tensor analysis, and machine learning.

Jing-Hua Yang received the PhD degree from the Macau University of Science and Technology, in 2023. She is currently an assistant professor with the School of Information Science and Technology, Southwest Jiaotong University. Her current research interests include data mining, image processing, tensor analysis, and federated learning.

Xi-Le Zhao received the MS and PhD degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009 and 2012, respectively. He worked as a post-doc with Prof. Michael K. Ng with Hong Kong Baptist University from 2013 to 2014. He worked as a visiting scholar with Prof. Jose Bioucas Dias with the University of Lisbon from 2016 to 2017. He is currently a professor with the School of Mathematical Sciences, UESTC. His research interests include scientific computing, image processing, and machine learning. More information can be found on his homepage https://zhaoxile.github.io/.

Jie Zhang received the PhD degree from the Department of Mathematics, The University of Hong Kong, Hong Kong, China, in 2024. She is currently a postdoc researcher with The University of Hong Kong. Her research interests include image processing, optimization, optimal transport, data science and machine learning.

Michael K. Ng (Senior Member, IEEE) received the BSc and MPhil degrees from The University of Hong Kong, Hong Kong, in 1990 and 1992, respectively, and the PhD degree from The Chinese University of Hong Kong, Hong Kong, in 1995. From 1995 to 1997, he was a research fellow with Computer Sciences Laboratory, The Australian National University, Canberra, ACT, Australia. He was an assistant professor/associate professor with The University of Hong Kong from 1997 to 2005. He was a professor/chair professor (2005–2019) with the Department of Mathematics, Hong Kong Baptist University, Hong Kong, chair professor (2019–2023) with the Department of Mathematics, The University of Hong Kong. He is currently a chair professor in mathematics and chair professor in Data Science with Hong Kong Baptist University. His research interests include applied and computational mathematics, machine learning and artificial intelligence, and data science. He serves as an editorial board member of several international journals. He was selected for the 2017 Class of Fellows of the Society for Industrial and Applied Mathematics. He received the Feng Kang Prize for his significant contributions to scientific computing.