Fast Unpaired Multi-view Clustering

Xingfeng Li[1,2,3], **Yuangang Pan**[2,3], **Yinghui Sun**[4*], **Quansen Sun**[1]
Ivor Tsang[2,3] and **Zhenwen Ren**[5*]
[1]Department of Computer Science, Nanjing University of Science and Technology
[2]Centre for Frontier AI Research, Agency for Science, Technology and Research, Singapore
[3]Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore
[4]School of Computer Science and Engineering, Southeast University
[5]School of National Defence Science and Technology, Southwest University of Science and Technology
{lixingfeng, yinghuisun, sunquansen, rzw}@njust.edu.cn, {yuangang.pan, ivor.tsang}@gmail.com

Abstract

Anchor based pair-wised multi-view clustering often assumes multi-view data are paired, and has demonstrated significant advancements in recent years. However, this presumption is easily violated, and data is commonly unpaired fully in practical applications due to the influence of data collection and storage processes. Addressing unpaired large-scale multi-view data through anchor learning remains a research gap. The absence of pairing in multi-view data disrupts the consistency and complementarity of multiple views, posing significant challenges in learning powerful and meaningful anchors and bipartite graphs from unpaired multi-view data. To tackle this challenge, this study proposes a novel **Fast Unpaired Multi-view Clustering (FUMC)** framework for fully unpaired large-scale multi-view data. Specifically, FUMC first designs an inverse local manifold learning paradigm to guide the learned anchors for effective pairing and balancing, ensuring alignment, fairness, and power in unpaired multi-view data. Meanwhile, a novel bipartite graph matching framework is developed to align unpaired bipartite graphs, creating a consistent bipartite graph from unpaired multi-view data. The efficacy, efficiency, and superiority of our FUMC are corroborated through extensive evaluations on numerous benchmark datasets with shallow and deep SOTA methods.

1 Introduction

In recent years, driven by rapid advancements in science and technology, significant volumes of data have been amassed from diverse sources or feature extractors to delineate a singular entity, giving rise to the construction of multi-view data [Cai *et al.*, 2024; Yao *et al.*, 2023; Zhang *et al.*, 2022; Li *et al.*, 2023a]. For instance, a singular news story can be preserved and disseminated through various formats such as video, audio, and text. Moreover, it can also be reported in disparate nations, spanning distinct languages like Chinese, English, Russian, and French. The recent surge in multi-view clustering (MVC) finds its impetus in its ability to harness the affinity between samples and views [Zhang *et al.*, 2023; Chen *et al.*, 2023b; Lu *et al.*, 2023; Sun *et al.*, 2023b]. This allows for the effective amalgamation of consistency and complementary attributes inherent in multi-view data, ultimately leading to the attainment of optimal cluster allocation outcomes [Cai *et al.*, 2022; Chen *et al.*, 2024; Sun *et al.*, 2023a; Li *et al.*, 2022].

The existing MVC critically hinges on an implicit assumption that each sample collected and stored from different views is perfectly aligned and paired, implying the same positional arrangement across all views. However, this assumption is facilely contradicted in practical applications, destroying the cross-view consistency and complementary of paired data and causing intricate challenges in dealing with unpaired multi-view data, named the Arbitrary View-unpaired Problem (AVP). Presently, only a few deep or shallow attempts have been made to address view-unpaired data [Yu *et al.*, 2021; Lin *et al.*, 2022; Huang *et al.*, 2020; Yang *et al.*, 2021; Yang *et al.*, 2023]. To our knowledge, existing most deep methods [Huang *et al.*, 2020; Yang *et al.*, 2021; Yang *et al.*, 2023] require partially paired data as the training data, which greatly limits practicality. Comparatively, the above shallow methods [Yu *et al.*, 2021; Lin *et al.*, 2022] apply to handle fully unpaired multi-view data. Regrettably, these shallow endeavors involve constructing sample n square affinity graphs, whose computations suffer from n cube computational complexity and n square storage complexity, both waning their applicability in large-scale data tasks.

As a potent approach for handling large-scale data, the anchor technique facilitates the sampling or learning of a limited set of m representative anchor points ($m << n$) from n original multi-view samples. Constructing a bipartite graph of dimensions $m * n$, these techniques remarkably alleviate the spatial and temporal complexities from $O(n^2)$ and $O(n^3)$ to $O(n)$, consequently significantly diminishing the computational and storage burdens. Existing anchor techniques emerge two primary classes: (1) static anchor learning encompasses methods such as random extraction, k-means, and VDA anchor selection policies [Kang *et al.*, 2021;

*Corresponding author.

Xia *et al.*, 2022]. (2) dynamic anchor learning simultaneously learns anchors and bipartite graphs dynamically, where also a consistent bipartite graph is learned to feed into final clustering [Wang *et al.*, 2021; Ji and Feng, 2023; Dong *et al.*, 2023]. Both for static and dynamic anchor learning, learning a consistent bipartite graph plays a vital role in the final clustering results. Unlike static learning, dynamic anchor learning adjusts anchors dynamically in line with the model's requirements during optimization. Existing dynamic anchor learning based methods mainly include two classes: View-consistent anchor learning [Chen *et al.*, 2022] and view-specific anchor learning [Wang *et al.*, 2022]. Although effective, they can only address paired multi-view data. While multi-view data tends to be unpaired in real applications, causing existing anchor based methods to fail. Taking two views of Fig. 1 as an example, both popular view-consistent anchor learning in Fig. 1 (b) and view-specific anchor learning in Fig. 1 (c) would fail to directly learn a consistent bipartite graph from the unpaired multi-view data. Because consistency and complementary of multiple views get disrupted.

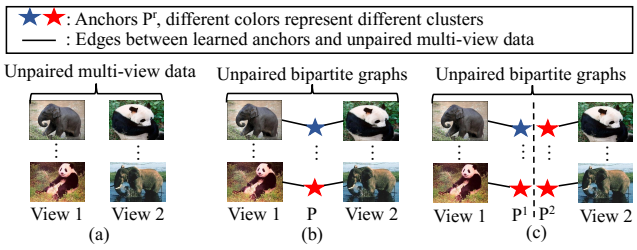

Figure 1: Various color stars indicate anchors of different clusters. Solid lines symbolize the correlation of anchors and unpaired multi-view data, where all the edges in each view construct a bipartite graph. (a) is unpaired multi-view data. (b) and (c) are view-consistency and view-specific anchor learning frameworks for unpaired multi-view data.

Intuitively, it is difficult to directly learn consistent anchors from unpaired data via Fig. 1 (b). So we use view-specific anchor learning of Fig. 1 (c) to learn a consistent bipartite graph from unpaired multi-view data for clustering. We observe two challenging issues required to be considered: (1) How to avoid learning unaligned or weak cross-view anchors from unpaired data? See Fig. 1 (c), blue anchor (blue star) of View 1 and red anchor (red star) of View 2 have unaligned order and unaligned number; Even worse, View 2 may suffer from blue anchor absence. (2) With the aligned and powerful anchors, how to learn a consistent bipartite graph from the unpaired multi-view data? To address these two challenging issues, we propose a novel **Fast Unpaired Multi-view Clustering (FUMC)** framework for large-scale multi-view clustering. Specifically, for a first tricky challenge, we dexterously designed a novel inverse local manifold learning paradigm to learn aligned and powerful view-specific anchors by predefining a prior similarity matrix according to the desired anchor order and anchor number. With the aligned and powerful anchors, we further design a bipartite graph matching framework, which enforces multiple view-specific bipartite graphs to align a powerful bipartite graph learned from best-view data. To do so, we could learn a consistent bipartite graph from unpaired multi-view data for fast clustering. Primary contributions are summarized as follows:

- We design a new inverse local manifold learning paradigm to enforce the learned anchors towards pairing and balancing, thereby ensuring alignment, fairness, and power of learned anchors in unpaired multi-view data.
- We design an ingenious bipartite graph matching framework to align unpaired bipartite graphs, obtaining a consistent bipartite graph from unpaired multi-view data.
- We propose a novel FUMC framework. To the best of our knowledge, this is the first attempt to handle unpaired multi-view data with the anchor technique. Extensive experiments on shallow and deep SOTA methods verify the superiority of our FUMC.

2 Related Work

2.1 Unpaired Multi-view Subspace Clustering

Lately, only a handful of studies have attempted to address the Arbitrary View-unpaired Problem for Unpaired Multi-view Subspace Clustering (UMSC) [Yu *et al.*, 2021; Lin *et al.*, 2022]. These endeavors involve constructing candidate affinity graphs for each view, expanding them into n-square dimensions, and subsequently aligning them using n-by-n matching matrices to learn an aligned consensus graph. This aligned graph is fed into spectral clustering to obtain final clustering results. Note that both n-by-n graphs and n-by-n matching matrices suffer from $O(n^3)$ time complexity and $O(n^2)$ space complexity, which waning their applicability in large-scale tasks. Compared to the above shallow methods, although some deep unpaired multi-view clustering methods are proposed [Huang *et al.*, 2020; Yang *et al.*, 2021; Yang *et al.*, 2023], they suffer from the limitation of partially paired data as the training data.

2.2 Anchor-based Multi-view Clustering

Recently, dynamic anchor learning-based multi-view clustering [You *et al.*, 2023; Ji and Feng, 2023] has been highly effective in efficiently managing extensive datasets and achieved promising progress. Their success critically hinges on an implicit assumption that each sample across views is perfectly aligned and paired, such that they could make full use of the consistent and complementary information of paired multi-view data. Further, due to $m \ll n$, they could alleviate time and space complexities significantly [Dong *et al.*, 2023]. Two widely used base paradigms are

$$\begin{aligned} &\min_{\mathbf{P},\mathbf{S}} \Phi(\mathbf{Y}^r, \mathbf{P}, \mathbf{S}) \\ &\min_{\mathbf{P}^r,\mathbf{S}^r,\mathbf{S}} \Psi(\mathbf{Y}^r, \mathbf{P}^r, \mathbf{S}^r) + \mathcal{R}(\mathbf{S}^r, \mathbf{S}) \end{aligned} \tag{1}$$

where $\mathbf{S} \in \mathbb{R}^{m*n}$ is a consistent bipartite graph, which builds a connection between view-consistency/view-specific anchor $\mathbf{P} \in \mathbb{R}^{k*m}/\mathbf{P}^r \in \mathbb{R}^{d^r*m}$ and paired multi-view data $\{\mathbf{Y}^r\}_{r=1}^v \in \mathbb{R}^{d^r*n}$. m and k are anchor number and dimension of shared subspace, respectively. d^r is the dimension

of r-th view. The final clustering result is obtained from $\mathbf{S}$ and its quality greatly depends on final performance. $\Phi(\cdot)$ and $\Psi(\cdot)$ denotes view-consistency and view-specific anchor graph learning frameworks, respectively. $\mathcal{R}$ is the regularization term. From Fig. 1, when samples across different views become unpaired/unaligned, causing inconsistent bipartite graph structures across views naturally. Thus, existing anchor-based multi-view clustering would fail since they cannot learn a consistent $\mathbf{S}$ from the unpaired multi-view data.

3 Formulation

3.1 FUMC Framework

From the analysis of Fig. 1, using anchor technique to perform unpaired multi-view clustering need to address two challenges. To address these two issues, we propose a novel general anchor graph learning framework, named **Fast Unpaired Multi-view Clustering (FUMC)**. Specifically, we first design an inverse local manifold learning paradigm, which aims to learn cross-view anchors with manifold protection like in Fig. 2 (b) from the high-dim manifold spaces in Fig. 2 (a). Then, we further design a bipartite graph matching framework, which enforces multiple view-specific bipartite graphs to align a powerful bipartite graph learned from best-view data.

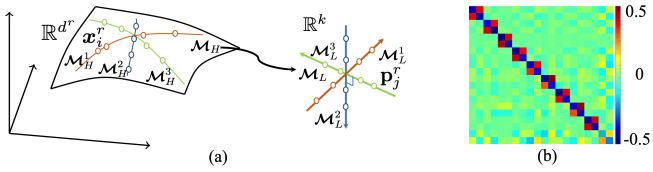

Figure 2: (a) The distribution $\mathcal{D}$ of high-dim data $\mathbf{x}_i^r \in \mathbb{R}^{d^r}$ is supported on a manifold $\boldsymbol{\mathcal{M}}_H$, with its classes residing on low-dimensional submanifolds $\boldsymbol{\mathcal{M}}_L^i$. We aim to learn the low-dim anchors lie on a union of maximally uncorrelated low-dim subspaces $\boldsymbol{\mathcal{M}}_L^i$ from the unpaired data, meanwhile, anchors also enjoy the properties P1 to P5. (b) shows the learned desired $(\mathbf{P}^2)^\intercal\mathbf{P}^3$ using our FUMC on the Mnist dataset with 10 classes.

Overall, FUMC aims to learn powerful and paired view-specific anchors from unpaired multi-view data, and then construct a consistent graph for clustering. To learn powerful and paired anchors, we first require our learned anchors to have five following properties:

P1: Intra-view Inter-cluster Discriminative: Anchor features in distinct clusters should manifest robust independence and relate to separate low-dimensional linear subspaces.

P2: Intra-view Intra-cluster Compressible: Anchor features within the same cluster should display significant correlation and be linked to a common low-dimensional linear subspace.

P3: Inter-view anchor order alignment: The order of anchors in the same clusters from different views is equal.

P4: Intra-view Maximally Diverse Anchors: Anchor features should be as large as possible as long as anchor features for each cluster stay uncorrelated from the other clusters.

P5: Inter-view anchor number alignment: The number of anchors in the same clusters from different views is equal.

To achieve this goal, we design the following inverse local manifold learning scheme as the following:

$$\min_{\mathbf{P}^r} \mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G}) \tag{2}$$

$\mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G}) = \sum_{r=1}^{v}\sum_{i=1}^{m}\sum_{j=1}^{m}(\|\mathbf{p}_i^b - \mathbf{p}_j^r\|_F^2)g_{ij} + \sum_{r=1}^{v}\|(\mathbf{P}^r)^\intercal\mathbf{P}^r - \mathbf{I}_m\|_F^2$, where the $\mathbf{G} \in \mathbb{R}^{m*m}$ is the prior matrix to guide the learned anchors to satisfy 5 properties (P1-P5). Each element g_{ij} of $\mathbf{G}$ encodes the similarity between i-th anchor $\mathbf{p}_i^b$ of b-th best view and j-th anchor $\mathbf{p}_j^r$ of r-th view in the manifold space. Thus, if we require $\mathbf{p}_i^b$ and $\mathbf{p}_j^r$ to be paired in the same cluster, g_{ij} should be greater than zero to drag these two anchors to be closer. Inversely, $g_{ij} = 0$ would separate $\mathbf{p}_i^b$ and $\mathbf{p}_j^r$ as possible. Inspired by this, we predefine a prior $\mathbf{G}$ to enforce cross-view anchors toward satisfying P1-P4 properties. Meanwhile, orthogonal constraints make anchor matrices satisfy the P5. $\mathbf{P}^b$ is an anchor matrix corresponding to the best-view data $\mathbf{X}^b$, which is learned beforehand by performing Silhouette Coefficient [Lin *et al.*, 2022] on $\{\mathbf{X}\}_{r=1}^v$. To do so, the learned cross-view anchors avoid order misalignment, number misalignment, or even anchor absence in some clusters as shown in Fig. 1 (c).

With the paired cross-view anchors at hand, we further construct bipartite graphs from these anchors, and then align them to generate a consistent graph for clustering by the following bipartite graph matching framework

$$\min_{\mathbf{S}^r, \mathbf{W}^r, \boldsymbol{\mathcal{Z}}} \mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}^r, \mathbf{G}) \tag{3}$$

$\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}^r, \mathbf{G}) = \mathtt{Tr}(\sum_{r=1}^{v}(\mathbf{S}^b)^\intercal\mathbf{L}_G\mathbf{W}^r\mathbf{S}^r) + \|\boldsymbol{\mathcal{Z}}\|_{ETNN}$, where $\boldsymbol{\mathcal{Z}} \in \mathbb{R}^{m*n*v} = \Psi(\mathbf{S}^b, \mathbf{W}^r\mathbf{S}^r, \cdots, \mathbf{W}^v\mathbf{S}^v)$ is a tensor stacked by a best-view bipartite graph $\mathbf{S}^b \in \mathbb{R}^{m*n}$ and multiple bipartite graphs $\{\mathbf{S}^r \in \mathbb{R}^{m*n}\}_{r=1,r\neq b}^v$ with rotation matrices $\{\mathbf{W}^r \in \mathbb{R}^{m*m}\}_{r=1,r\neq b}^v$. $\|\cdot\|_{ETNN}$ is a new Enhanced Tensor Nuclear Norm (ETNN) to exploit the better low-rank property of the core matrix whose entries are from the diagonal elements of a core tensor. Laplacian matrix $\mathbf{L}_G = \mathbf{T} - \mathbf{G}$, where the degree matrix $\mathbf{T}$ is a diagonal matrix with $T_{ii} = \sum_{j=1}^{m}\mathbf{G}_{ij}$. Overall, $\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}^r, \mathbf{G})$ could align $\mathbf{S}^r$ to $\mathbf{S}^b$ with rotation matrix $\mathbf{W}^r$ and $\mathbf{G}$. In this way, a consistent bipartite graph $\mathbf{S}^* = \mathbf{S}^b + \sum_{r=1,r\neq b}^{v}\mathbf{W}^r\mathbf{S}^r/v$ is learned to obtain the final clustering results. We seamlessly integrate Eq. (1), Eq. (2), and (3) to get the final objective function as

$$\begin{aligned}\min_{\mathbf{P}^r, \mathbf{Q}^r, \mathbf{S}^r, \mathbf{W}^r, \boldsymbol{\mathcal{Z}}} & \gamma\sum_{r=1}^{v}\|(\mathbf{Q}^r)^\intercal\mathbf{X}^r - \mathbf{P}^r\mathbf{S}^r\|_F^2 \\ & +\beta\mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G}) + \alpha\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}^r, \mathbf{G}) \\ \text{s.t.} & (\mathbf{S}^r)^\intercal\mathbf{1} = \mathbf{1}, \mathbf{S}^r \geq 0, (\mathbf{Q}^r)^\intercal\mathbf{Q}^r = \mathbf{I}_m, \\ & \boldsymbol{\mathcal{Z}} = \Psi(\mathbf{S}^b, \mathbf{W}^r\mathbf{S}^r, \cdots, \mathbf{W}^v\mathbf{S}^v)\end{aligned} \tag{4}$$

where γ, β, and α are the control parameters. $\mathbf{Q}^r \in \mathbb{R}^{d^r \times k}$ is the projection matrix for r-th veiw.

3.2 Optimization of FUMC

To solve Eq. (4), the auxiliary variable $\boldsymbol{\mathcal{H}}$ is introduced to make Eq. (4) separable. Then, Eq. (4) is rewritten as the

following augmented Lagrangian function

$$\begin{aligned}\min_{\substack{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{H}},\mathbf{S}^r,\\ \mathbf{W}^r,\mathbf{P}^r,\mathbf{Q}^r}} & \gamma\sum_{r=1}^{v}\|(\mathbf{Q}^r)^\intercal\mathbf{X}^r-\mathbf{P}^r\mathbf{S}^r\|_F^2+\frac{\mu}{2}\|\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{H}}+\frac{\boldsymbol{\mathcal{Y}}}{\mu}\|_F^2\\ &+\sum_{r=1,r\neq b}^{v}[\beta\mathtt{Tr}(\mathbf{P}^r\mathbf{L}_G(\mathbf{P}^b)^\intercal)+\|\boldsymbol{\mathcal{H}}\|_{ETNN}\\ &+\alpha\mathtt{Tr}((\mathbf{S}^b)^\intercal\mathbf{L}_G\mathbf{W}^r\mathbf{S}^r)]\\ \text{s.t. } \mathbf{S}^r\geq 0,&(\mathbf{S}^r)^\intercal\mathbf{1}=\mathbf{1},(\mathbf{Q}^r)^\intercal\mathbf{Q}^r=\mathbf{I}_m,(\mathbf{P}^r)^\intercal\mathbf{P}^r=\mathbf{I}_m,\\ &\boldsymbol{\mathcal{Z}}=\Psi(\mathbf{S}^b,\mathbf{W}^r\mathbf{S}^r,\cdots,\mathbf{W}^v\mathbf{S}^v),\boldsymbol{\mathcal{Z}}=\boldsymbol{\mathcal{H}}\end{aligned} \tag{5}$$

For the sake of easy optimization, $\mathbf{P}^r$ in the second term of the $\mathcal{R}_{\Delta_P}$ can be analytically solved. Therefore, we put it as constraint. Eq. (5) could be separately solved by developing an alternating iterative algorithm as follows.

► **Step-1.1: Solving $\mathbf{S}^r$ with P, Q, W, $\mathbf{S}^b$, and $\boldsymbol{\mathcal{H}}$ fixed.** Then, $\mathbf{S}^r$-subproblem of Eq. (5) changes to

$$\min_{\mathbf{S}^b}\ \|\mathbf{S}^r-\hat{\mathbf{S}}^r\|_F^2 \text{ s.t. } \mathbf{S}^r\geq 0,(\mathbf{S}^r)^\intercal\mathbf{1}=\mathbf{1}, \tag{6}$$

where $\hat{\mathbf{S}}^r=\frac{(\mathbf{P}^r)^\intercal(\mathbf{Q}^r)^\intercal\mathbf{X}^r-0.5\beta\mathbf{L}_G^\intercal\mathbf{S}^b+0.5\mu\mathbf{H}^r-0.5\mathbf{Y}^r}{(1+0.5\mu)\mathbf{I}}$. The j-th column vector of $\hat{\mathbf{S}}^r$ is defined as $\hat{\mathbf{s}}_j^r$, whose i-th element is $\hat{z}_{i,j}^r$. For the r-th view, letting $\hat{\mathbf{s}}_j^r$ and $\hat{s}_{i,j}^r$ be the j-th column and ij-th element of $\hat{\mathbf{s}}^r$, respectively, updating $\mathbf{s}^r$ converts to the following column form.

$$\min_{\mathbf{s}_j}\|\mathbf{s}_j-\hat{\mathbf{s}}_j\|_F^2,\ \text{s.t. } \forall ij,\mathbf{s}_j\mathbf{1}=1,\mathbf{s}_{ij}\geq 0 \tag{7}$$

Following **Theorem 1**, Eq. (7) could be optimized.

Theorem 1. [Li *et al.*, 2023d] *Given arbitrary v vectors $\{\hat{\mathbf{s}_j}\}_{j=1}^v$, we obtain the following closed-form solution $\mathbf{s}_j^*$*

$$\mathbf{s}_j^*=\arg\min_{\mathbf{s}_j}\|\mathbf{s}_j-\hat{\mathbf{s}_j}\|_F^2,\ \text{s.t. } \mathbf{s}_j^\intercal\mathbf{1}=1,\mathbf{s}_j\geq 0 \tag{8}$$

► **Step-1.2: Solving $\mathbf{S}^b$ with P, Q, W, $\mathbf{S}^r$, and $\boldsymbol{\mathcal{H}}$ fixed.** Then, $\mathbf{S}^b$-subproblem of Eq. (5) changes to

$$\min_{\mathbf{S}^b}\ \|\mathbf{S}^b-\hat{\mathbf{S}}^b\|_F^2 \text{ s.t. } \mathbf{S}^b\geq 0,(\mathbf{S}^b)^\intercal\mathbf{1}=\mathbf{1}, \tag{9}$$

where $\hat{\mathbf{S}}^b=\frac{(\mathbf{P}^b)^\intercal(\mathbf{Q}^b)^\intercal\mathbf{X}^b-0.5\beta\mathbf{D}^b+0.5\mu\mathbf{H}^b-0.5\mathbf{Y}^b}{(1+0.5\mu)\mathbf{I}}$ and $\mathbf{D}^b=\mathbf{L}_G^\intercal\sum_{r=1}^v\mathbf{W}^r\mathbf{S}^r$. Similar to optimize $\mathbf{S}^r$, Eq. (9) could be solved via **Theorem 1.**

► **Update-2: Solving Q with S, P, and $\boldsymbol{\mathcal{W}}$ fixed.** In this case, $\mathbf{Q}$-subproblem of Eq. (5) can be written as

$$\max_{\mathbf{Q}^r}\mathtt{Tr}((\mathbf{Q}^r)^\intercal\mathbf{E}^r)\ \text{s.t.}\mathbf{Q}^r(\mathbf{Q}^r)^\intercal=\mathbf{I}_k, \tag{10}$$

where $\mathbf{E}^r=\mathbf{X}^r(\mathbf{S}^r)^\intercal(\mathbf{P}^r)^\intercal$. Eq. (10) can be solved via the SVD operator [Li *et al.*, 2024; Li *et al.*, 2023b] with complexity $\mathcal{O}(v\hat{d}(nm+k^2+km))$ for each iteration, where $\hat{d}=\sum_{p=1}^v d^p$.

► **Step-3-1 update $\mathbf{P}^r$:** Optimizing $\mathbf{P}^r$ with the irrelevant variables fixed is equivalent to the following optimization problem

$$\max_{\mathbf{P}^r}\mathtt{Tr}((\mathbf{P}^r)^\intercal\mathbf{C}^r)\ \text{s.t.}(\mathbf{P}^r)^\intercal\mathbf{P}^r=\mathbf{I}_m \tag{11}$$

where $\mathbf{C}^r=\gamma(\mathbf{Q}^r)^\intercal\mathbf{X}^r(\mathbf{S}^r)^\intercal-\beta\mathbf{P}^b\mathbf{L}_G$. The optimal solution of optimizing $\mathbf{P}^r$ can be effectively obtained via singular value decomposition (SVD) on $\mathbf{C}^r$.

► **Step-3-2 update $\mathbf{P}^b$:** Similar to optimize $\mathbf{P}^r$, the optimal $\mathbf{P}^b$ is obtained by perform SVD on the $\mathbf{C}^b=\gamma(\mathbf{Q}^b)^\intercal\mathbf{X}^b(\mathbf{S}^b)^\intercal-\beta\sum_{r=1,r\neq b}^v\mathbf{L}_G^\intercal\mathbf{P}^r$. The complexity of optimizing $\mathbf{P}$ is $\mathcal{O}(v(ndk+nmk+m^2k))$ per iteration, where $d_{all}=\sum_{r=1}^v d^l$.

► **Step-4 update $\mathbf{W}$:** Fixing the irrelevant variables, and updating $\mathbf{W}$ as

$$\begin{aligned}\min_{\mathbf{W}^r}\ &\sum_{r=1}^{v}\frac{\mu}{2}\|\mathbf{W}^r\mathbf{S}^r-\mathbf{H}^r+\frac{\mathbf{Y}^r}{\mu}\|_F^2+\alpha\mathtt{Tr}[(\mathbf{S}^b)^\intercal\mathbf{L}_G\mathbf{W}^r\mathbf{S}^r]\\ &\text{s.t.}(\mathbf{W}^r)^\intercal\mathbf{W}^r=\mathbf{I}_m\end{aligned} \tag{12}$$

where $\mathbf{B}^r=\frac{\mu}{2}(\mathbf{H}^r-\mathbf{Y}^r)(\mathbf{S}^r)^\intercal-\alpha\mathbf{L}_G^\intercal\mathbf{S}^b(\mathbf{S}^r)^\intercal$. The optimal solution of optimizing $\mathbf{W}^r$ can be effectively obtained via singular value decomposition (SVD) on $\mathbf{B}^r$ with complexity $\mathcal{O}(vd(nm+km+k^2))$ per iteration.

► **Step-5 update $\boldsymbol{\mathcal{H}}$:** Ignoring the irrelevant items $\boldsymbol{\mathcal{H}}$, updating $\boldsymbol{\mathcal{H}}$ subproblem is

$$\min_{\boldsymbol{\mathcal{H}}}\ \|\boldsymbol{\mathcal{H}}\|_{ETNN}+\frac{\mu}{2}\|\boldsymbol{\mathcal{H}}-(\boldsymbol{\mathcal{Z}}+\frac{\boldsymbol{\mathcal{Y}}}{\mu})\|_F^2 \tag{13}$$

Which can be solved into two steps as follows: (1) minimizing the core matrix, and (2) minimizing t-TNN.

(1) Updating core matrix as

$$\min_{\mathfrak{P}(\boldsymbol{\mathcal{Z}})}\|\mathfrak{P}(\boldsymbol{\mathcal{Z}})\|_*+\frac{1}{2\lambda}\|\boldsymbol{\mathcal{F}}-(\boldsymbol{\mathcal{Z}}+\frac{\boldsymbol{\mathcal{Y}}}{\mu})\|_F^2 \tag{14}$$

where regularization parameter $\lambda=1/(max(m,v)n)^{\frac{1}{2}}$. And the tensor $\boldsymbol{\mathcal{Z}}$ is obtained from t-SVD on the temporary variable $\boldsymbol{\mathcal{F}}$, $\boldsymbol{\mathcal{F}}=\boldsymbol{\mathcal{U}}*\boldsymbol{\mathcal{Z}}*\boldsymbol{\mathcal{V}}$.

(2) Updating $\boldsymbol{\mathcal{H}}$ as

$$\min_{\boldsymbol{\mathcal{H}}}\ \|\boldsymbol{\mathcal{H}}\|_{ETNN}+\frac{\mu}{2}\|\boldsymbol{\mathcal{H}}-\boldsymbol{\mathcal{J}}\|_F^2 \tag{15}$$

With the learned low-rank core matrix $\mathfrak{P}(\boldsymbol{\mathcal{Z}})$, we can use t-product to reconstruct a tensor as $\boldsymbol{\mathcal{J}}=\boldsymbol{\mathcal{U}}*\mathfrak{P}(\boldsymbol{\mathcal{Z}})^{-1}*\boldsymbol{\mathcal{V}}$. The learned $\boldsymbol{\mathcal{J}}$ can further produce a closed-form solution [Li *et al.*, 2023c; Li *et al.*, 2023d].

Updating ADMM variables are written as

$$\begin{aligned}\boldsymbol{\mathcal{Y}}&=\boldsymbol{\mathcal{Y}}+\mu(\boldsymbol{\mathcal{Z}}-\boldsymbol{\mathcal{H}})\\ \mu&=min(\rho\mu,\mu_{max})\end{aligned} \tag{16}$$

where $\mu=1e^{-4}$ and $\mu_{max}=10^{10}$, and the complexity is $\mathcal{O}(n)$. The whole optimization procedure of Eq. (5) is outlined in Algorithm 1, where convergence criterion is checked by computing the objective value obj^t at the t-th iteration.

3.3 Complexity Analysis

Time complexity. In the optimizing Algorithm 1, **Step-1**, **Step-2**, **Step-3**, **Step-4**, **Step-5**, and ADMM variables costs $\mathcal{O}(v\hat{d}nm)$, $\mathcal{O}(vd(nm+k^2+km))$, $\mathcal{O}(v(ndk+m^2k+nmk))$, $\mathcal{O}(vd(nm+km+k^2))$, $\mathcal{O}\left(v^2nm+vnm\log(n)\right)$, and $\mathcal{O}(v)$ per iteration. After completing Algorithm 1, the final $\mathbf{S}^*$

Algorithm 1 The algorithm of FUMC

Require: Unpaired multi-view data $\{\mathbf{X}^r\}_{r=1}^v$, cluster number c, latent space dimension k, and parameters α, β, γ.
Initialize $\mathbf{Q}^r = \mathbf{I}_k$, and the others matrices as $\mathbf{0}$.
repeat
  Update $\mathbf{S}^r$ and $\mathbf{S}^b$ by using Eq. (6) and Eq. (9);
  Update $\mathbf{Q}^r$, $\mathbf{P}^r$, $\mathbf{P}^b$, $\mathbf{W}^r$, and $\boldsymbol{\mathcal{H}}$ via Eqs (10)-(13);
  Update ADMM variables via Eq. (16);
until convergence.
Obtain the left singular value matrix $\mathbf{U}_{\mathbf{S}^*}$ from $\mathbf{S}^*$.
Ensure: Clustering performance

costs $\mathcal{O}(nm^2)$ by performing SVD and the subsequent k-means. Totally, the calculation complexity of our FUMC is $\mathcal{O}(t(v\hat{d}nm + v\hat{d}(nm + k^2 + km) + v(n\hat{d}k + m^2k + nmk) + v^2nm + vnm\log(n) + nm^2))$. In our FUMC optimization, $m \ll n$ and $k \ll n$. Thus, time complexity of optimizing FUMC is approximatively $\mathcal{O}(n)$.

Space complexity. The space consumption of our FUMC mainly comes from $\mathbf{Q}^r \in \mathbb{R}^{d^p \times m}$, $\mathbf{X}^r \in \mathbb{R}^{n \times d^p}$, $\mathbf{P}^r \in \mathbb{R}^{m \times k}$, $\mathbf{S}^r \in \mathbb{R}^{n \times m}$, $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{m \times n \times v}$, and $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{m \times n \times v}$. Therefore, the total space complexity of FUMC costs $(n + k)m + 3mnv + \hat{d}(n+m)$. Similar to time complexity, $m \ll n$ and $k \ll n$ deduce that Algorithm 1 inherits linear space complexity to sample numbers.

4 Experiments

4.1 Experimental Setting

Dataset	Size	Dimensionality	Views	Classes
ORL*	400	$\{4096, 3304, 6750\}$	3	40
MSRCv1*	210	$\{24, 512, 256, 254\}$	4	7
BBCSport*	282	$\{2582, 2544.2456\}$	3	5
Mnist*	2000	$\{9, 9, 30\}$	3	10
YTF-10*	38654	$\{944, 576, 512, 640\}$	4	10
YTF-20*	63896	$\{944, 576, 512, 640\}$	4	20
YTF-100*	195537	$\{944, 576, 512, 640\}$	4	100
Scene-15*	4485	$\{20, 59\}$	2	15
ANIMAL*	10158	$\{4096, 4096\}$	2	50
LandUse-21*	2100	$\{59, 40\}$	2	21
RGBD*	1449	$\{2048, 300\}$	2	13
Synthetic*	200	$\{2, 2\}$	2	4

Table 1: Datasets for experiment evaluation, $*$ denotes the unpaired datasets or competitors in this paper.

Datasets. In the experiments, seven fully paired multi-view data are employed to compare with shallow methods, including ORL, BBCSport, MSRCv1, Mnist, YTF-10, YTF-20, and YTF-100 [Chen *et al.*, 2022; Lin *et al.*, 2022]. Four partially paired multi-view data with pair ratio ($\phi = 50\%$) to compare with deep methods, including Scene-15, Animal, LandUse-21, and RGBD [Yang *et al.*, 2023]. Their detailed statistical information is summarized in Table 1. We generate misaligned data by randomly shuffling sample arrangement of different views, where source codes are requested from author of [Lin *et al.*, 2022] to ensure the fairness of experiments. The implementation of all competitors is collected from their public homepages or requested from authors. For shallow methods, our experiments are conducted on a 32GB RAM and Intel Core i7 CPU, 2021 Mac mini computing platform with Matlab 2021b, while deep methods are on PyTorch 1.8.1, and an NVIDIA 3090 GPU with Ubuntu 18.04.

To the best of our knowledge, a small amount of effort has been attempted to address arbitrary view-unpaired problem. On four common metrics [Wang *et al.*, 2022; Wen *et al.*, 2022; Kang *et al.*, 2020], 14 **shallow** and 7 **deep** competitors are totally employed to evaluate the validity and efficiency of our FUMC to handle arbitrary view-unpaired data, i.e., fully unpaired and partially unpaired ones. Concretely, *for shallow methods with fully unpaired data:* Two unpaired multi-view clustering methods: T-UMC [Lin *et al.*, 2022], MVC-UM [Yu *et al.*, 2021]; And *12* ***paired*** *methods :* FSMSC [Chen *et al.*, 2023c], FMVACC [Wang *et al.*, 2022], OMSC[Chen *et al.*, 2022], FPMVS [Wang *et al.*, 2021] WTNNM [Gao *et al.*, 2020] t-SVD-MSC [Xie *et al.*, 2018], ETLMSC [Wu *et al.*, 2019], LTMSC [Zhang *et al.*, 2015], RMSC [Xia *et al.*, 2014], LMSC [Zhang *et al.*, 2020], AMGL [Nie *et al.*, 2016], SC,[Ng *et al.*, 2002]. *For deep methods with partially unpaired data:* Seven unpaired competitors includes: PVC [Huang *et al.*, 2020], MVCLN [Yang *et al.*, 2021], DSIMVC [Tang and Liu, 2022], MFLVC [Xu *et al.*, 2022], CVCL[Chen *et al.*, 2023a], DCP [Lin *et al.*, 2023], SURE [Yang *et al.*, 2023].

4.2 Experiment Results

The average clustering performance and standard deviations (Each experiment is repeated 20 times.) are reported on four fully unpaired datasets in Table 2 and Table 3. Further, Table 4 only compares four anchor based multi-view clustering methods on the paired and fully unpaired large-scale datasets since other shallow competitors suffer from n cube time and n square space complexities. The best and second-best averages are marked in bold and underlined, respectively. Fig. 4 reports learned consistent bipartite graphs from aligned and unaligned Synthetic datasets. According to reported tables and figures, we can observe that:

Effectiveness and superiority over shallow fully unpaired competitors. (1) Due to the carefully designed matching module, our FUMC and unpaired multi-view subspace clustering (T-UMC and MVC-UM) are superior to tensor based multi-view clustering (WTNNM, t-SVD-MSC, ETLMSC, and LTMSC), graph based multi-view clustering baselines (RMSC, LMSC, and AMGL), and single view SC by a large margin. Furthermore, compared to T-UMC and MVC-UM, our FUMC may contain less redundant information by learning the high-quality anchor features in the refined subspaces; Meanwhile, FUMC could also capture more compact high-order tensor low-rank structure by deeper exploring the low-rank core matrix of entries in core tensor than existing TNN. Thus, our FUMC consistently outperforms all these paired and unpaired clustering methods.
(2) Compared to shallow anchor based multi-view clustering (FSMSC, FMVACC, OMSC, FPMVS, and SMVSC), our FUMC significantly outperforms them when performing unpaired multi-view clustering. The main reason is that un-

paired data destroys the cross-view consistency and complementary of paired data, which greatly reduces the quality of anchors and bipartite graphs. Especially in Table 4, these anchor based multi-view clustering methods obtain comparatively higher performance on the paired datasets. However, for unpaired datasets, their performance drops dramatically except for our FUMC, which is consistent with Fig. 4. Fig. 4 indicates that existing anchor based multi-view clustering methods are almost impossible to learn a consistent bipartite graph from unpaired multi-view data except our FUMC. Based on (1) and (2), effectiveness and superiority of FUMC are demonstrated.

Efficiency. Table 5 reports the time cost of three unpaired multi-view clustering methods, MVC-UM, T-UMC, and our FUMC. As can be seen, the time cost of our FUMC is far below than existing MVC-UM and T-UMC, demonstrating the efficiency of FUMC.

Effectiveness and superiority over deep methods. (Most of them are partially paired methods) These partially paired deep methods must require partially paired data as the train data, so we set the paired ratio $\phi = 50\%$ for four deep comparison datasets. From Table 6, our FUMC consistently outperforms all the competitors, which demonstrates the effectiveness and superiority of our FMUC.

4.3 Ablation Analysis

To demonstrate the significance and effectiveness of cross-view graph matching term, we remove $\mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G})$ and $\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}, \mathbf{G})$ as Ablation1 and Ablation 2, respectively. Whether removing $\mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G})$ or $\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}, \mathbf{G})$, our model degrades largely as mentioned in Table 7. The main reasons are that $\mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G})$ ensures the learned anchors satisfying the 1-5 properties, i.e., intra-view inter-cluster discriminative, intra-view intra-cluster compressible, intra-view maximally diverse, inter-view anchor number and order alignment. $\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}, \mathbf{G})$ ensures the learned bipartite graphs to be rotated and matching the bipartite graph of the best view. Thus, both proposed $\mathcal{R}_{\Delta_P}(\mathbf{P}^b, \mathbf{P}^r, \mathbf{G})$ and $\mathcal{R}_{\Delta_S}(\mathbf{S}^b, \mathbf{S}^r, \mathbf{W}, \mathbf{G})$ are significant and effective.

Method	ORL*				BBCSport *			
	ACC	NMI	PUR	F-score	ACC	NMI	PUR	F-score
FUMC*	**96.70**$_{\pm 0.01}$	**99.18**$_{\pm 0.01}$	**97.62**$_{\pm 0.35}$	**96.53**$_{\pm 0.01}$	**89.01**$_{\pm 0.00}$	75.31$_{\pm 0.00}$	**89.01**$_{\pm 0.01}$	79.06$_{\pm 0.00}$
T-UMC*	91.22$_{\pm 0.27}$	96.78$_{\pm 0.15}$	93.82$_{\pm 0.30}$	92.12$_{\pm 0.31}$	82.67$_{\pm 0.01}$	**78.64**$_{\pm 0.01}$	89.57$_{\pm 0.00}$	**83.07**$_{\pm 0.01}$
MVC-UM*	56.56$_{\pm 0.12}$	74.72$_{\pm 0.11}$	60.38$_{\pm 0.11}$	48.62$_{\pm 0.14}$	43.88$_{\pm 0.03}$	16.01$_{\pm 0.03}$	47.13$_{\pm 0.03}$	74.39$_{\pm 0.04}$
FSMSC	22.00$_{\pm 0.00}$	44.58$_{\pm 0.00}$	23.25$_{\pm 0.00}$	6.39$_{\pm 0.00}$	18.32 $_{\pm 0.00}$	11.36$_{\pm 0.00}$	15.93$_{\pm 0.00}$	8.33$_{\pm 0.00}$
FMVACC	45.79$_{\pm 0.02}$	64.23$_{\pm 0.02}$	49.20$_{\pm 0.02}$	27.80$_{\pm 0.02}$	26.31$_{\pm 0.03}$	3.68$_{\pm 0.01}$	28.24$_{\pm 0.02}$	15.91$_{\pm 0.03}$
OMSC	20.00$_{\pm 0.00}$	43.35$_{\pm 0.00}$	20.50$_{\pm 0.00}$	5.73$_{\pm 0.00}$	12.88$_{\pm 0.00}$	6.36$_{\pm 0.00}$	13.49$_{\pm 0.00}$	7.92$_{\pm 0.00}$
FPMVS	19.75$_{\pm 0.00}$	42.23$_{\pm 0.00}$	20.50$_{\pm 0.00}$	6.25$_{\pm 0.00}$	13.64$_{\pm 0.00}$	7.92$_{\pm 0.00}$	13.61$_{\pm 0.00}$	8.92$_{\pm 0.00}$
WTNNM	44.15$_{\pm 0.02}$	63.58$_{\pm 0.01}$	47.10$_{\pm 0.01}$	28.76$_{\pm 0.02}$	40.21$_{\pm 0.00}$	14.85$_{\pm 0.00}$	55.32$_{\pm 0.00}$	28.49$_{\pm 0.00}$
t-SVD-MSC	68.65$_{\pm 0.14}$	81.27$_{\pm 0.06}$	73.05$_{\pm 0.06}$	57.58$_{\pm 0.12}$	47.59$_{\pm 0.01}$	17.00$_{\pm 0.01}$	47.59$_{\pm 0.00}$	35.11$_{\pm 0.01}$
ETLMSC	41.90$_{\pm 0.26}$	60.52$_{\pm 0.20}$	44.36$_{\pm 0.28}$	25.97$_{\pm 0.31}$	26.94$_{\pm 0.01}$	2.83$_{\pm 0.01}$	38.84$_{\pm 0.01}$	23.26$_{\pm 0.01}$
LTMSC	63.10$_{\pm 0.22}$	75.78$_{\pm 0.17}$	65.80$_{\pm 0.23}$	48.84$_{\pm 0.28}$	41.35$_{\pm 0.02}$	13.97$_{\pm 0.02}$	48.09$_{\pm 0.01}$	30.10$_{\pm 0.01}$
RMSC	32.00$_{\pm 0.21}$	53.28$_{\pm 0.13}$	34.10$_{\pm 0.20}$	14.31$_{\pm 0.16}$	37.45$_{\pm 0.03}$	10.97$_{\pm 0.02}$	50.28$_{\pm 0.03}$	29.44$_{\pm 0.02}$
LMSC	30.68$_{\pm 0.02}$	26.78$_{\pm 0.02}$	31.42$_{\pm 0.02}$	20.86$_{\pm 0.02}$	39.33$_{\pm 0.08}$	25.15$_{\pm 0.07}$	41.62$_{\pm 0.06}$	27.03$_{\pm 0.06}$
AMGL	40.45$_{\pm 0.28}$	61.90$_{\pm 0.15}$	44.00$_{\pm 0.23}$	23.55$_{\pm 0.21}$	34.11$_{\pm 0.02}$	5.16$_{\pm 0.01}$	39.01$_{\pm 0.02}$	28.81$_{\pm 0.02}$
SC	77.60$_{\pm 0.28}$	89.34$_{\pm 0.10}$	80.68$_{\pm 0.21}$	71.29$_{\pm 0.23}$	44.50$_{\pm 0.02}$	22.93$_{\pm 0.01}$	56.03$_{\pm 0.02}$	36.94$_{\pm 0.01}$

Table 2: Performance comparison 14 shallow methods, where best and second-best averages are marked in bold and underlined, respectively.

Method	Mnist*				MSRCv1*			
	ACC	NMI	PUR	F-score	ACC	NMI	PUR	F-score
UFMC*	**99.60**$_{\pm 0.00}$	**98.98**$_{\pm 0.00}$	**99.60**$_{\pm 0.00}$	**99.33**$_{\pm 0.00}$	**88.57**$_{\pm 0.00}$	**82.99**$_{\pm 0.00}$	**88.57**$_{\pm 0.00}$	**79.90**$_{\pm 0.00}$
T-UMC*	98.68$_{\pm 0.00}$	96.65$_{\pm 0.00}$	98.68$_{\pm 0.00}$	97.26$_{\pm 0.00}$	83.97$_{\pm 0.01}$	77.74$_{\pm 0.01}$	83.97$_{\pm 0.01}$	75.89$_{\pm 0.02}$
MVC-UM*	65.40$_{\pm 0.01}$	54.26$_{\pm 0.01}$	68.10$_{\pm 0.01}$	56.69$_{\pm 0.01}$	44.96$_{\pm 0.01}$	34.06$_{\pm 0.01}$	47.08$_{\pm 0.01}$	41.30$_{\pm 0.01}$
FSMSC	18.21 $_{\pm 0.00}$	5.44$_{\pm 0.00}$	21.17$_{\pm 0.00}$	13.08$_{\pm 0.00}$	28.57$_{\pm 0.00}$	7.73$_{\pm 0.00}$	28.57$_{\pm 0.00}$	16.79$_{\pm 0.00}$
FMVACC	29.95$_{\pm 0.03}$	15.19$_{\pm 0.03}$	36.90$_{\pm 0.03}$	20.02$_{\pm 0.03}$	37.52$_{\pm 0.02}$	25.85$_{\pm 0.02}$	39.72$_{\pm 0.02}$	25.93$_{\pm 0.01}$
OMSC	16.85$_{\pm 0.00}$	2.44$_{\pm 0.00}$	20.10$_{\pm 0.00}$	12.81$_{\pm 0.00}$	22.15$_{\pm 0.00}$	8.39$_{\pm 0.00}$	23.66$_{\pm 0.00}$	16.38$_{\pm 0.00}$
FPMVS	16.25$_{\pm 0.00}$	2.54$_{\pm 0.00}$	18.55$_{\pm 0.00}$	13.73$_{\pm 0.00}$	23.81$_{\pm 0.00}$	7.65$_{\pm 0.00}$	25.71$_{\pm 0.00}$	17.15$_{\pm 0.00}$
WTNNM	38.70$_{\pm 0.00}$	22.30$_{\pm 0.00}$	42.09$_{\pm 0.00}$	24.90$_{\pm 0.00}$	42.67$_{\pm 0.00}$	27.96$_{\pm 0.01}$	44.10$_{\pm 0.01}$	30.23$_{\pm 0.01}$
t-SVD-MSC	39.47$_{\pm 0.00}$	28.04$_{\pm 0.00}$	41.93$_{\pm 0.00}$	32.34$_{\pm 0.00}$	40.76$_{\pm 0.01}$	25.86$_{\pm 0.01}$	42.48$_{\pm 0.00}$	28.94$_{\pm 0.01}$
ETLMSC	12.82$_{\pm 0.01}$	0.84$_{\pm 0.00}$	16.89$_{\pm 0.00}$	10.47$_{\pm 0.00}$	22.95$_{\pm 0.01}$	5.09$_{\pm 0.01}$	23.77$_{\pm 0.01}$	14.28$_{\pm 0.01}$
LTMSC	37.54$_{\pm 0.00}$	25.98$_{\pm 0.00}$	40.14$_{\pm 0.00}$	29.68$_{\pm 0.00}$	34.57$_{\pm 0.02}$	17.89$_{\pm 0.02}$	36.29$_{\pm 0.01}$	22.11$_{\pm 0.01}$
RMSC	42.12$_{\pm 0.03}$	31.75$_{\pm 0.02}$	46.79$_{\pm 0.01}$	33.76$_{\pm 0.03}$	37.43$_{\pm 0.02}$	27.35$_{\pm 0.04}$	39.52$_{\pm 0.03}$	26.96$_{\pm 0.02}$
LMSC	50.78$_{\pm 0.00}$	48.77$_{\pm 0.00}$	63.06$_{\pm 0.00}$	43.24$_{\pm 0.00}$	39.33$_{\pm 0.08}$	25.15$_{\pm 0.07}$	41.62$_{\pm 0.06}$	27.03$_{\pm 0.06}$
AMGL	17.08$_{\pm 0.01}$	03.96$_{\pm 0.00}$	19.73$_{\pm 0.01}$	12.05$_{\pm 0.00}$	29.71$_{\pm 0.02}$	12.63$_{\pm 0.02}$	30.38$_{\pm 0.02}$	18.43$_{\pm 0.01}$
SC	87.76$_{\pm 0.00}$	76.62$_{\pm 0.00}$	87.76$_{\pm 0.00}$	79.46$_{\pm 0.00}$	61.57$_{\pm 0.01}$	56.29$_{\pm 0.01}$	66.62$_{\pm 0.01}$	53.13$_{\pm 0.01}$

Table 3: Performance comparison 14 shallow methods, where best and second-best averages are marked in bold and underlined, respectively.

Datasets (samples)	Metric	FPMVS TIP'22	OMSC KDD'22	FMVACC NIPS'22	FSMSC TIP'23	FUMC
YTF-100 (195537)	ACC	52.93$_{\pm 0.00}$	66.51$_{\pm 0.01}$	63.58$_{\pm 0.01}$	62.13$_{\pm 0.01}$	**66.68**$_{\pm 0.01}$
	NMI	75.32$_{\pm 0.00}$	83.37$_{\pm 0.01}$	82.09$_{\pm 0.01}$	82.13$_{\pm 0.01}$	**83.56**$_{\pm 0.01}$
	Purity	54.46$_{\pm 0.00}$	71.41$_{\pm 0.01}$	71.78$_{\pm 0.01}$	70.68$_{\pm 0.01}$	**74.61**$_{\pm 0.00}$
	Fscore	35.41$_{\pm 0.00}$	58.46$_{\pm 0.01}$	38.13$_{\pm 0.02}$	40.62$_{\pm 0.01}$	**58.76**$_{\pm 0.02}$
YTF-100* (195537)	ACC	16.39$_{\pm 0.00}$	17.63$_{\pm 0.00}$	41.15$_{\pm 0.04}$	10.92$_{\pm 0.02}$	**62.02**$_{\pm 0.00}$
	NMI	1.68$_{\pm 0.00}$	2.09$_{\pm 0.00}$	44.08$_{\pm 0.05}$	0.96$_{\pm 0.01}$	**70.75**$_{\pm 0.00}$
	Purity	18.11$_{\pm 0.00}$	20.27$_{\pm 0.00}$	44.69$_{\pm 0.05}$	16.05$_{\pm 0.02}$	**74.93**$_{\pm 0.00}$
	Fscore	54.78$_{\pm 0.00}$	12.80$_{\pm 0.00}$	18.59$_{\pm 0.03}$	8.92$_{\pm 0.03}$	**66.05**$_{\pm 0.01}$
YTF-20 (63896)	ACC	69.48$_{\pm 0.00}$	74.46$_{\pm 0.00}$	66.83$_{\pm 0.02}$	67.93$_{\pm 0.01}$	**74.70**$_{\pm 0.00}$
	NMI	77.90$_{\pm 0.00}$	**81.70**$_{\pm 0.00}$	78.27$_{\pm 0.01}$	79.21$_{\pm 0.02}$	79.67$_{\pm 0.01}$
	Purity	72.59$_{\pm 0.00}$	77.31$_{\pm 0.00}$	73.18$_{\pm 0.01}$	74.35$_{\pm 0.01}$	**79.91**$_{\pm 0.01}$
	Fscore	62.61$_{\pm 0.00}$	68.35$_{\pm 0.00}$	55.46$_{\pm 0.05}$	60.27$_{\pm 0.01}$	**69.09**$_{\pm 0.00}$
YTF-20* (63896)	ACC	8.04$_{\pm 0.00}$	7.94$_{\pm 0.00}$	10.42$_{\pm 0.00}$	15.36$_{\pm 0.01}$	**71.22**$_{\pm 0.01}$
	NMI	0.56$_{\pm 0.00}$	0.70$_{\pm 0.00}$	3.37$_{\pm 0.00}$	3.67$_{\pm 0.00}$	**72.24**$_{\pm 0.01}$
	Purity	10.54$_{\pm 0.00}$	11.02$_{\pm 0.00}$	13.05$_{\pm 0.01}$	18.27$_{\pm 0.01}$	**75.77**$_{\pm 0.01}$
	Fscore	7.21$_{\pm 0.00}$	5.87$_{\pm 0.00}$	7.14$_{\pm 0.00}$	9.11$_{\pm 0.00}$	**61.24**$_{\pm 0.01}$
YTF-10 (38654)	ACC	73.25$_{\pm 0.00}$	78.20$_{\pm 0.00}$	76.52$_{\pm 0.01}$	77.23$_{\pm 0.01}$	**79.71**$_{\pm 0.01}$
	NMI	77.40$_{\pm 0.00}$	**82.75**$_{\pm 0.00}$	62.70$_{\pm 0.01}$	72.13$_{\pm 0.02}$	81.59$_{\pm 0.01}$
	Purity	76.21$_{\pm 0.00}$	82.98$_{\pm 0.00}$	66.15$_{\pm 0.01}$	73.24$_{\pm 0.01}$	**83.78**$_{\pm 0.01}$
	Fscore	69.59$_{\pm 0.00}$	74.56$_{\pm 0.00}$	70.20$_{\pm 0.04}$	72.99$_{\pm 0.02}$	**75.85**$_{\pm 0.01}$
YTF-10* (38654)	ACC	13.52$_{\pm 0.00}$	12.07$_{\pm 0.00}$	13.31$_{\pm 0.01}$	20.13$_{\pm 0.01}$	**63.15**$_{\pm 0.06}$
	NMI	0.21$_{\pm 0.00}$	0.51$_{\pm 0.00}$	0.59$_{\pm 0.01}$	8.21$_{\pm 0.01}$	**54.98**$_{\pm 0.03}$
	Purity	16.21$_{\pm 0.00}$	16.71$_{\pm 0.00}$	16.76$_{\pm 0.01}$	25.41$_{\pm 0.02}$	**66.91**$_{\pm 0.05}$
	Fscore	13.69$_{\pm 0.00}$	11.37$_{\pm 0.00}$	11.90$_{\pm 0.04}$	12.43$_{\pm 0.01}$	**54.78**$_{\pm 0.04}$

Table 4: The performance comparison anchor based shallow methods, where best results are marked in bold.

Method	ORL	MSRCv1	BBCSport	Mnist	YTF-10	YTF-20	YTF-100
MVC-UM	18.3	4.6	16.2	429.6	-	-	-
T-UMC	21.2	5.0	11.4	533.8	-	-	-
FUMC	2.8	0.7	0.7	3.2	99.9	174.3	2658.5

Table 5: Computational cost (In Seconds). '-' denotes out of CPU memory or storage memory.

Method	Scene-15		Animal		LandUse-21		RGBD	
	ACC	NMI	ACC	NMI	ACC	NMI	ACC	NMI
PVC	37.88	39.12	3.8	0.1	21.33	23.14	18.63	1.64
MVCLN	38.53	39.9	26.18	40.19	24.14	27.43	25.95	18.35
DSIMVC	12.27	3.88	7.41	6.26	6.11	1.79	28.50	1.79
MFLVC	22.36	10.52	6.94	5.09	12.05	5.03	25.33	4.97
CVCL	23.79	11.48	6.51	5.79	15.41	9.87	14.77	4.89
DCP	39.50	42.35	19.86	21.46	23.07	27	25.49	8.65
SURE	40.32	40.33	27.74	40.83	23.81	28.6	35.68	33.26
FUMC	**65.63**	**62.91**	**41.66**	**43.13**	**31.35**	**37.97**	**48.72**	**36.05**

Table 6: The performance comparison deep methods.

Dataset	Ablation 1	Ablation 2	ACC	NMI	PUR	F-score
ORL*	✓	✓	**96.70**	**99.18**	**97.62**	**96.70**
	✓		26.21	46.31	25.65	11.07
		✓	30.48	52.22	32.36	13.21
BBCSport*	✓	✓	**79.10**	**81.50**	**89.72**	**79.90**
	✓		20.76	11.22	29.641	15.27
		✓	58.36	33.49	63.47	44.68
Mnist*	✓	✓	**99.60**	**98.98**	**99.60**	**99.33**
	✓		39.45	48.57	39.45	29.08
		✓	75.28	66.59	80.22	66.05
MSRCv1*	✓	✓	**88.57**	**82.99**	**88.57**	**79.90**
	✓		27.83	18.14	30.47	10.91
		✓	39.09	22.79	40.21	24.75
YTF20*	✓	✓	**71.22**	**72.24**	**75.77**	**61.24**
	✓		16.95	20.68	25.48	8.63
		✓	32.96	22.16	32.61	15.93

Table 7: The performance comparison of ablation analysis.

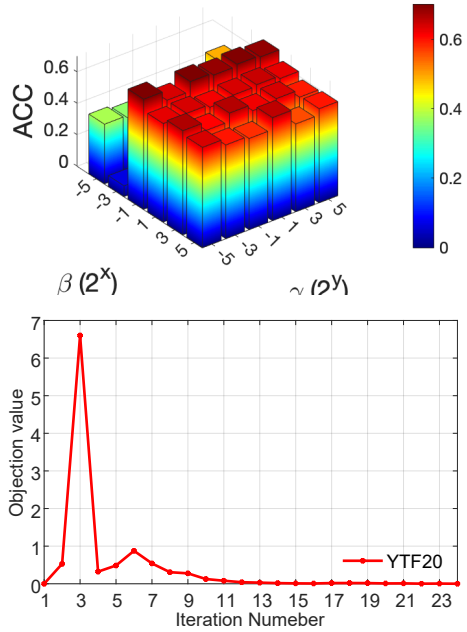

Figure 3: The parameter settings and convergence on YTF20*.

4.4 Parameter Settings and Convergence

Algorithm 1 involves three parameters to be set properly, parameters α β, γ. Through this paper, only $m = 2c$ anchors are learned and fixed for all datasets. The consensus dimension is also fixed as $k = 2c$. As shown in Fig. 3, our FUMC only provides the parameter settings and convergence of YTF20*.

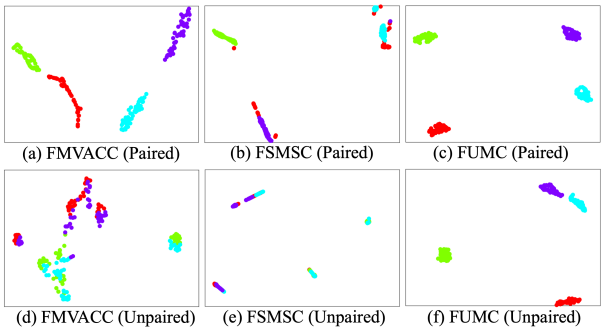

Figure 4: The learned consistent bipartite graphs from paired and unpaired Synthetic dataset over FSMSC, FMVACC, and our FUMC.

4.5 Conclusion

In this paper, we propose a novel FUMC framework to handle unpaired large-scale multi-view clustering. To our knowledge, FUMC presents the first study of addressing unpaired multi-view data through anchor learning. The carefully designed inverse local manifold learning and bipartite graph matching can respectively learn the powerful and aligned anchors and bipartite graphs from unpaired multi-view data. Comprehensive experiments and analysis have proved the superiority and efficiency of our method.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62372235), the A*STAR Career Development Fund (Grant No.C222812019), the Mianyang Science and Technology Program (Grant No. 2022ZYDF089), and the Base Strengthening Program of National Defense Science and Technology (Grant No. 2022-JCJQ-JJ-0292), the Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant No. KYCX23_0487), and the China Scholarship Council No. 202306840101.

References

[Cai *et al.*, 2022] Jinyu Cai, Jicong Fan, Wenzhong Guo, Shiping Wang, Yunhe Zhang, and Zhao Zhang. Efficient deep embedded subspace clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1–10, 2022.

[Cai *et al.*, 2024] Jinyu Cai, Yunhe Zhang, Shiping Wang, Jicong Fan, and Wenzhong Guo. Wasserstein embedding learning for deep clustering: A generative approach. *IEEE Transactions on Multimedia*, 2024.

[Chen *et al.*, 2022] Man-Sheng Chen, Chang-Dong Wang, Dong Huang, Jian-Huang Lai, and Philip S Yu. Efficient orthogonal multi-view subspace clustering. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 127–135, 2022.

[Chen *et al.*, 2023a] Jie Chen, Hua Mao, Wai Lok Woo, and Xi Peng. Deep multiview clustering by contrasting cluster assignments. *Proceedings of the IEEE International Conference on computer vision*, 2023.

[Chen *et al.*, 2023b] Zhaoliang Chen, Lele Fu, Jie Yao, Wenzhong Guo, Claudia Plant, and Shiping Wang. Learnable graph convolutional network and feature fusion for multi-view learning. *Information Fusion*, 95:109–119, 2023.

[Chen *et al.*, 2023c] Zhe Chen, Xiao-Jun Wu, Tianyang Xu, and Josef Kittler. Fast self-guided multi-view subspace clustering. *IEEE Transactions on Image Processing*, pages 1–1, 2023.

[Chen *et al.*, 2024] Zhaoliang Chen, Zhihao Wu, Luying Zhong, Claudia Plant, Shiping Wang, and Wenzhong Guo. Attributed multi-order graph convolutional network for heterogeneous graphs. *Neural Networks*, page 106225, 2024.

[Dong *et al.*, 2023] Zhibin Dong, Siwei Wang, Jiaqi Jin, Xinwang Liu, and En Zhu. Cross-view topology based consistent and complementary information for deep multi-view clustering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 19440–19451, 2023.

[Gao *et al.*, 2020] Quanxue Gao, Wei Xia, Zhizhen Wan, Deyan Xie, and Pu Zhang. Tensor-svd based graph learning for multi-view subspace clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3930–3937, 2020.

[Huang *et al.*, 2020] Zhenyu Huang, Peng Hu, Joey Tianyi Zhou, Jiancheng Lv, and Xi Peng. Partially view-aligned clustering. *Advances in Neural Information Processing Systems*, 33:2892–2902, 2020.

[Ji and Feng, 2023] Jintian Ji and Songhe Feng. Anchor structure regularization induced multi-view subspace clustering via enhanced tensor rank minimization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 19343–19352, 2023.

[Kang *et al.*, 2020] Zhao Kang, Wangtao Zhou, Zhitong Zhao, Junming Shao, Meng Han, and Zenglin Xu. Large-scale multi-view subspace clustering in linear time. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 4412–4419, 2020.

[Kang *et al.*, 2021] Zhao Kang, Zhiping Lin, Xiaofeng Zhu, and Wenbo Xu. Structured graph learning for scalable subspace clustering: From single view to multiview. *IEEE Transactions on Cybernetics*, 52(9):8976–8986, 2021.

[Li *et al.*, 2022] Xingfeng Li, Yinghui Sun, Quansen Sun, and Zhenwen Ren. Consensus cluster center guided latent multi-kernel clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2022.

[Li *et al.*, 2023a] Haoran Li, Yulan Guo, Zhenwen Ren, F Richard Yu, Jiali You, and Xiaojian You. Explicit local coupling global structure clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.

[Li *et al.*, 2023b] Liang Li, Junpu Zhang, Siwei Wang, Xinwang Liu, Kenli Li, and Keqin Li. Multi-view bipartite graph clustering with coupled noisy feature filter. *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[Li *et al.*, 2023c] Xingfeng Li, Zhenwen Ren, Quansen Sun, and Zhi Xu. Auto-weighted tensor schatten p-norm for robust multi-view graph clustering. *Pattern Recognition*, 134:109083, 2023.

[Li *et al.*, 2023d] Xingfeng Li, Yinghui Sun, Quansen Sun, Zhenwen Ren, and Yuan Sun. Cross-view graph matching guided anchor alignment for incomplete multi-view clustering. *Information Fusion*, page 101941, 2023.

[Li *et al.*, 2024] Liang Li, Yuangang Pan, Jie Liu, Yue Liu, Xinwang Liu, Kenli Li, Ivor W Tsang, and Keqin Li. Bgae: Auto-encoding multi-view bipartite graph clustering. *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[Lin *et al.*, 2022] Jia-Qi Lin, Man-Sheng Chen, Chang-Dong Wang, and Haizhang Zhang. A tensor approach for uncoupled multiview clustering. *IEEE Transactions on Cybernetics*, 2022.

[Lin *et al.*, 2023] Yijie Lin, Yuanbiao Gou, Xiaotian Liu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. Dual contrastive prediction for incomplete multi-view representation learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4447–4461, 2023.

[Lu *et al.*, 2023] Zhoumin Lu, Shiping Wang, Genggeng Liu, and Feiping Nie. Robust weighted co-clustering

with global and local discrimination. *Pattern Recognition*, 138:109405, 2023.

[Ng *et al.*, 2002] Andrew Y Ng, Michael I Jordan, and Yair Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pages 849–856, 2002.

[Nie *et al.*, 2016] Feiping Nie, Jing Li, Xuelong Li, et al. Parameter-free auto-weighted multiple graph learning: A framework for multiview clustering and semi-supervised classification. In *The Twenty-Fifth International Joint Conference on Artificial Intelligence*, pages 1881–1887, 2016.

[Sun *et al.*, 2023a] Yuan Sun, Zhenwen Ren, Peng Hu, Dezhong Peng, and Xu Wang. Hierarchical consensus hashing for cross-modal retrieval. *IEEE Transactions on Multimedia*, 2023.

[Sun *et al.*, 2023b] Yuan Sun, Xu Wang, Dezhong Peng, Zhenwen Ren, and Xiaobo Shen. Hierarchical hashing learning for image set classification. *IEEE Transactions on Image Processing*, 32:1732–1744, 2023.

[Tang and Liu, 2022] Huayi Tang and Yong Liu. Deep safe incomplete multi-view clustering: Theorem and algorithm. In *International Conference on Machine Learning*, pages 21090–21110. PMLR, 2022.

[Wang *et al.*, 2021] Siwei Wang, Xinwang Liu, Xinzhong Zhu, Pei Zhang, Yi Zhang, Feng Gao, and En Zhu. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Transactions on Image Processing*, 31:556–568, 2021.

[Wang *et al.*, 2022] Siwei Wang, Xinwang Liu, Suyuan Liu, Jiaqi Jin, Wenxuan Tu, Xinzhong Zhu, and En Zhu. Align then fusion: Generalized large-scale multi-view clustering with anchor matching correspondences. *Advances in Neural Information Processing Systems*, 35:5882–5895, 2022.

[Wen *et al.*, 2022] Jie Wen, Zheng Zhang, Lunke Fei, Bob Zhang, Yong Xu, Zhao Zhang, and Jinxing Li. A survey on incomplete multi-view clustering. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022.

[Wu *et al.*, 2019] Jianlong Wu, Zhouchen Lin, and Hongbin Zha. Essential tensor learning for multi-view spectral clustering. *IEEE Transactions on Image Processing*, 28(12):5910–5922, 2019.

[Xia *et al.*, 2014] Rongkai Xia, Yan Pan, Lei Du, and Jian Yin. Robust multi-view spectral clustering via low-rank and sparse decomposition. In *Proceedings of the twenty-eighth AAAI conference on artificial intelligence*, pages 2149–2155, 2014.

[Xia *et al.*, 2022] Wei Xia, Quanxue Gao, Qianqian Wang, and Xinbo Gao. Tensor completion-based incomplete multiview clustering. *IEEE Transactions on Cybernetics*, 52(12):13635–13644, 2022.

[Xie *et al.*, 2018] Yuan Xie, Dacheng Tao, Wensheng Zhang, Yan Liu, Lei Zhang, and Yanyun Qu. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *International Journal of Computer Vision*, 126(11):1157–1179, 2018.

[Xu *et al.*, 2022] Jie Xu, Huayi Tang, Yazhou Ren, Liang Peng, Xiaofeng Zhu, and Lifang He. Multi-level feature learning for contrastive multi-view clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16051–16060, 2022.

[Yang *et al.*, 2021] Mouxing Yang, Yunfan Li, Zhenyu Huang, Zitao Liu, Peng Hu, and Xi Peng. Partially view-aligned representation learning with noise-robust contrastive loss. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1134–1143, 2021.

[Yang *et al.*, 2023] Mouxing Yang, Yunfan Li, Peng Hu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. Robust multi-view clustering with incomplete information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):1055–1069, 2023.

[Yao *et al.*, 2023] Yinghua Yao, Yuangang Pan, Jing Li, Ivor W Tsang, and Xin Yao. Sanitized clustering against confounding bias. *Machine Learning*, pages 1–20, 2023.

[You *et al.*, 2023] Jiali You, Zhenwen Ren, Xiaojian You, Haoran Li, and Yuancheng Yao. Priori anchor labels supervised scalable multi-view bipartite graph clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 10972–10979, 2023.

[Yu *et al.*, 2021] Hong Yu, Jia Tang, Guoyin Wang, and Xinbo Gao. A novel multi-view clustering method for unknown mapping relationships between cross-view samples. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 2075–2083, 2021.

[Zhang *et al.*, 2015] Changqing Zhang, Huazhu Fu, Si Liu, Guangcan Liu, and Xiaochun Cao. Low-rank tensor constrained multiview subspace clustering. In *Proceedings of the IEEE international conference on computer vision*, pages 1582–1590, 2015.

[Zhang *et al.*, 2020] Changqing Zhang, Huazhu Fu, Qinghua Hu, Xiaochun Cao, Yuan Xie, Dacheng Tao, and Dong Xu. Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(1):86–99, 2020.

[Zhang *et al.*, 2022] Chao Zhang, Huaxiong Li, Chunlin Chen, Yuhua Qian, and Xianzhong Zhou. Enhanced group sparse regularized nonconvex regression for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(5):2438–2452, 2022.

[Zhang *et al.*, 2023] Chao Zhang, Huaxiong Li, Wei Lv, Zizheng Huang, Yang Gao, and Chunlin Chen. Enhanced tensor low-rank and sparse representation recovery for incomplete multi-view clustering. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(9):11174–11182, Jun. 2023.