Scalable Incomplete Multi-View Clustering with Structure Alignment

Yi Wen
wenyi21@nudt.edu.com
National University of Defense Technology
Changsha, China

Siwei Wang*
wangsiwei13@nudt.edu.com
Intelligent Game and Decision Lab
Beijing, China

Ke Liang
liangke200694@126.com
National University of Defense Technology
Changsha, China

Weixuan Liang
weixuanliang@nudt.edu.cn
National University of Defense Technology
Changsha, China

Xinhang Wan
wanxinhang@nudt.edu.cn
National University of Defense Technology
Changsha, China

Xinwang Liu*
xinwangliu@nudt.edu.cn
National University of Defense Technology
Changsha, China

Suyuan Liu
suyuanliu@nudt.edu.cn
National University of Defense Technology
Changsha, China

Jiyuan Liu
liujiyuan13@nudt.edu.cn
National University of Defense Technology
Changsha, China

En Zhu
enzhu@nudt.edu.com
National University of Defense Technology
Changsha, China

ABSTRACT

The success of existing multi-view clustering (MVC) relies on the assumption that all views are complete. However, samples are usually partially available due to data corruption or sensor malfunction, which raises the research of incomplete multi-view clustering (IMVC). Although several anchor-based IMVC methods have been proposed to process the large-scale incomplete data, they still suffer from the following drawbacks: i) Most existing approaches neglect the inter-view discrepancy and enforce cross-view representation to be consistent, which would corrupt the representation capability of the model; ii) Due to the samples disparity between different views, the learned anchor might be misaligned, which we referred as the Anchor-Unaligned Problem for Incomplete data (AUP-ID). Such the AUP-ID would cause inaccurate graph fusion and degrades clustering performance. To tackle these issues, we propose a novel incomplete anchor graph learning framework termed Scalable Incomplete Multi-View Clustering with Structure Alignment (SIMVC-SA). Specially, we construct the view-specific anchor graph to capture the complementary information from different views. In order to solve the AUP-ID, we propose a novel structure alignment module to refine the cross-view anchor correspondence. Meanwhile, the anchor graph construction and alignment are jointly optimized in our unified framework to enhance clustering quality.

*Corresponding author

Through anchor graph construction instead of full graphs, the time and space complexity of the proposed SIMVC-SA is proven to be linearly correlated with the number of samples. Extensive experiments on seven incomplete benchmark datasets demonstrate the effectiveness and efficiency of our proposed method. Our code is publicly available at https://github.com/wy1019/SIMVC-SA.

CCS CONCEPTS

• **Computing methodologies** → **Cluster analysis**; • **Theory of computation** → **Unsupervised learning and clustering**.

KEYWORDS

anchor graph, incomplete multi-view clustering, large-scale clustering, multi-view clustering

1 INTRODUCTION

Multi-view data, which consists of features extracted from objects by different sensors, have been massively generated in recent years. Multi-modal information [21, 25, 27, 35, 43, 45, 67] about the data can commonly be utilized to enhance the expressive ability of the model. For instance, the same news can be described from different views, *i.e.,* textual reports [28] and visual pictures [48, 56, 72, 85]. As an essential paradigm of multi-view learning, multi-view clustering (MVC) has drawn substantial attention because of its promising capability to reveal the intrinsic data structure [2, 22]. In general, multi-view clustering achieves remarkable performance by learning

a consensus representation by exploring consistency among diverse views [17, 53, 83]. For instance, Zhan et al. [79] optimizes the final consensus graph by imposing low-rank constraints and minimizing the discrepancies of individual graphs. Zhang et al. [81] reconstruct samples in the latent space to achieve a more precise and reliable subspace representation.

Although numerous methods have been proposed to enhance MVC in diverse ways, most of them assume that all data are fully available [4, 18, 19, 51, 52, 57, 58, 80, 82]. However, samples are often partially available in real scenarios due to data corruption or sensor malfunction. For instance, in software traffic detection, people can't use all detected software, which leads to the incompleteness of the samples in the corresponding view. The different sample absence between views destroys the original cross-view alignment information and enlarges the difficulty of exploring consensus and complementary information, causing incomplete multi-view clustering (IMVC) a challenging problem. To tackle these issues, several IMVC methods have been proposed in previous literature. For instance, Li et al. [23] learn a common potential representation from incomplete samples by non-negative matrix decomposition and ℓ_1 regularization terms. Wen et al. [71] propose a new regularization term to preserve the local geometric structure and fuse the individual incomplete graph. Although remarkable success has been made, the high time complexity hinders their application in large-scale scenarios[41]. One pioneer work, Liu et al. [36] efficiently reduce the algorithm complexity by utilizing the anchor graph to capture the clustering structure with incomplete views.

Although widely applied in large-scale applications, the existing anchor-based IMVC methods still suffer from the following drawbacks: Firstly, most approaches neglect the inter-view discrepancy and enforce cross-view representation to be consistent, which would corrupt the representation capability of the model. Secondly, as is shown in Fig. 1, the sample distribution of different views might be biased due to the incomplete multi-view data, which leads to the potential misalignment between cross-view anchors, which we referred as the Anchor-Unaligned Problem for Incomplete Data (AUP-ID). Such AUP-ID would result in inaccurate graph fusion and suboptimal clustering performance. This issue for complete data has been demonstrated in [12, 65] and has more significant implications for IMVC due to the incomplete cross-view alignment information. Besides, to the best of our knowledge, no generalized framework for solving AUP-ID has been proposed since generating the correct cross-view anchor correspondence under incomplete scenarios would be more difficult due to the variance of the feature dimension and available sample.

To tackle these challenging issues, we propose a novel incomplete anchor graph learning framework termed Scalable Incomplete Multi-View Clustering with Structure Alignment (SIMVC-SA). Specifically, we construct the incomplete anchor graph on each view to capture the complementary information from different views. In order to address the AUP-ID, we adopt a novel structure alignment module to refine the cross-view anchor correspondence mapping adequately. Meanwhile, the anchor graph construction and alignment are jointly optimized in our unified framework to enhance clustering quality. In addition, through the anchor graph construction rather than a full pairwise graph, the time complexity of the SIMVC-SA is effectively reduced from $O(n^3)$ to $O(nm)$.
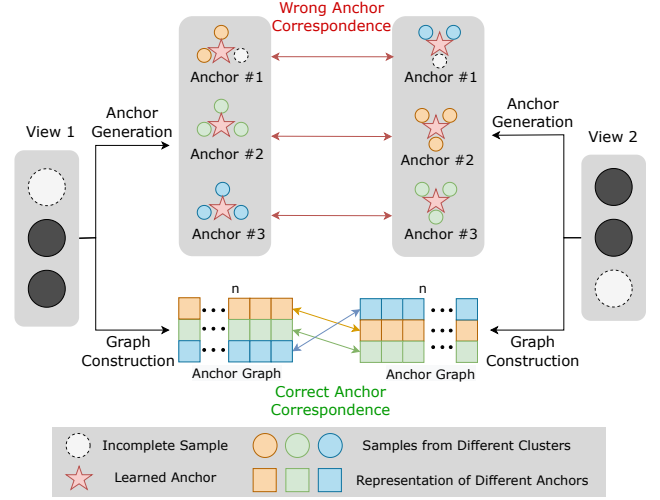
Figure 1: An example illustration of AUP-ID. Samples and representations of different colors represent samples from different clusters and representations of different anchors, respectively. With different missing samples and random anchor initialization, the anchor learned may be unaligned and leads to inaccurate correspondences.

Meanwhile, a convergent five-step alternative algorithm is designed in this paper to tackle the subsequent optimization problem. We summarize the contributions as follows:

- In order to solve the Anchor-Unaligned Problem for Incomplete Data, a novel alignment module is proposed in this paper to capture the view-specific structure. With the guidance of structure information, the cross-view anchor correspondence mapping can be refined adequately.
- We design a novel IMVC approach termed Scalable Incomplete Multi-View Clustering with Structure Alignment (SIM VC-SA). Different from the existing fixed anchor strategy, SIMVC-SA learns the anchor and constructs the respective anchor graph to enhance the clustering performance.
- Extensive experiments on seven incomplete benchmark datasets show the effectiveness and efficiency of the proposed method.

2 RELATED WORK

2.1 Incomplete Multi-View Clustering (IMVC)

In real scenarios, samples are often partially available due to data dropout and sensor corruption, which raises the incomplete multi-view clustering (IMVC) study [10, 13, 31, 59, 73, 74]. The existing Incomplete Muti-View Clustering (IMVC) approaches can be roughly divided into three types: Non-negative matrix factorization (NMF) methods [84], kernel or graph-based methods [8, 49], and deep neural networks [64, 68, 86].

In general, NMF jointly decomposes the raw matrix of each view into a coefficient matrix and a basis matrix and learns a consensus matrix from the coefficient matrices with a group of adaptive view weights. The graph or kernel-based IMVC approach performs matrix complementation and achieves the desired clustering performance by constructing consensus graphs or kernels [40, 54, 60].

For example, Wang et al. [62] propose a novel similarity matrix padding strategy based on matrix perturbation theory. Because of the capacity to extract high-level information, deep neural networks often achieve desirable performance for solving IMVC problems [32, 63, 70]. Lin et al. [33] design a deep IMVC model through the union of representation and cross-view data recovery.

2.2 Graph-based IMVC Method

Graph structures [14, 16, 29, 34, 42, 44, 46, 61, 76–78], which can well describe the relationships of pairwise data, are widely adopted in the field of IMVC. Denoting the indicator vector $\mathbf{w}^{(v)} \in \mathbb{R}^{n_v}$ containing the index for n_v available samples in the v-th view, we define the index matrix $\mathbf{H}_v \in \mathbb{R}^{n \times n_v}$ for v-th view as follows:

$$\mathbf{h}_{ij}^{(v)} = \begin{cases} 1, & \text{if } w_j^{(v)} = i, \\ 0, & \text{otherwise.} \end{cases}$$

where $\mathbf{h}_{ij}^{(v)}$ denotes the element in i-th row and j-th column of $\mathbf{H}_v$. Then, $\mathbf{X}_v \mathbf{H}_v \in \mathbb{R}^{d_v \times n_v}$ denotes the existing data matrix of the v-th view.

As to the incomplete multi-view data, the subgraphs from each view may have a few blanks in the respective rows and columns because of the incomplete setting. Taking this into account, the classical graph-based IMVC paradigms [24, 37] could be mathematically represented in two parts:

$$\min_{\mathbf{S}_v, \mathbf{S}} \|\mathbf{X}_v \mathbf{H}_v - \mathbf{X}_v \mathbf{H}_v \mathbf{S}_v\|_{\mathbf{F}}^2 + \Psi \left(\mathbf{H}_v \mathbf{S}_v \mathbf{H}_v^\top, \mathbf{S}\right),$$
$$\text{s.t. } \mathbf{S}_v \geq 0, \quad \mathbf{S}_v^\top \mathbf{1} = \mathbf{1}, \quad \mathbf{S} \geq 0, \quad \mathbf{S}^\top \mathbf{1} = \mathbf{1} \tag{1}$$

where $\mathbf{S}_v \in \mathbb{R}^{n_v \times n_v}$ is the view-specific subgraph, $\mathbf{S}$ indicates the similarity among all the samples. The $\Psi(\cdot)$ indicates the graph fusion process. However, $O\left(vn^2\right)$ space complexity and $O\left(n^3\right)$ time expenditure prevent this category of algorithms from handling large-scale incomplete multiview tasks [15].

2.3 Anchor-based IMVC Method

As is shown in Eq.(1), the majority of the classical graph-based IMVC methods involve the full graph construction, which makes them suffer from $O\left(n^3\right)$ time complexity. To tackle the issues, Li et al. [20], Liu et al. [36] propose the anchor-based incomplete multi-view clustering (AIMVC). The complexity of AIMVC is effectively reduced by merely building the relationship between representative anchors and samples [11]. The classical AIMVC framework can be formulated as follows:

$$\min_{\mathbf{Z}} \|\mathbf{X}_v \mathbf{H}_v - \mathbf{A}_v \mathbf{Z} \mathbf{H}_v\|_{\mathbf{F}}^2 + \Omega\left(\mathbf{Z}\right),$$
$$\text{s.t. } \mathbf{Z} \geq 0, \quad \mathbf{Z}^\top \mathbf{1} = \mathbf{1}, \tag{2}$$

where $\mathbf{A}_p \in \mathbb{R}^{d_v \times m}$ denotes the anchor matrix from the v-th view, $\mathbf{Z}$ is the consistent anchor graph, m is the number of anchors, and Ω is the regularization term.

On the basis of the above paradigm, many methods adopt different regularization terms to enhance clustering performance [7, 36]. However, most existing approaches overlook the inter-view discrepancy and enforce cross-view representation to be consistent, which would corrupt the representation capability of the model. Moreover, the potential Anchor-Unaligned Problem for Incomplete Data has not been discussed in previous research. Such AUP-ID

would result in inaccurate graph fusion and suboptimal clustering performance. In the next section, we will propose SIMVC-SA to tackle these issues.

3 METHODS

3.1 Problem Formulation

As mentioned before, the main challenge for solving AUP-ID is the variation of the feature dimension and available sample, which results in the anchors from different views being under different metric spaces, and we can't directly measure the distance of cross-view anchors. As a result, a question worth considering: **how to effectively refine the cross-view anchor correspondence under the incomplete scenario?**

An intuitive method [36] to implicitly avoid anchor correspondence is to enforce cross-view anchor and the respective graph to be consistent. While such a strategy overlooks the inter-view discrepancy and corrupts the representation capability of the model. Inspired by [12, 65], we consider such principle: the correspondence probability of the anchors should be high if their corresponding structure is similar. Therefore, the original anchor correspondence problem can be transferred to the structure alignment problem, as depicted in Fig. 1. In this paper, we introduce the alignment matrix $\mathbf{P}_v$ that satisfies $\mathbf{P}_v^\top \mathbf{P}_v = \mathbf{I}_m$ to efficiently tackle the problem. Denoting the fusion representation as $\mathbf{F}$, and the anchor graph alignment problem can be addressed as follow:

$$\min_{\mathbf{P}_v} \|\mathbf{P}_v \mathbf{Z}_v - \mathbf{F}\|_F^2, \quad \text{s.t.} \quad \mathbf{P}_v^\top \mathbf{P}_v = \mathbf{I}_m, \tag{3}$$

where $\mathbf{Z}_v \in \mathbb{R}^{m \times n}$ is the view-specific anchor graph.

Moreover, considering the traditional fixed anchor strategy relies on the quality of anchors initialization and introduces unnecessary time overhead, we adopt the anchor learning strategy to enhance our clustering performance in this paper. In summary, the proposed Scalable Incomplete Multi-View Clustering with Structure Alignment (SIMVC-SA) can be formulated as follows:

$$\min_{\substack{\gamma, \{\mathbf{A}_v\}_{v=1}^V, \\ \{\mathbf{Z}_v\}_{v=1}^V, \mathbf{P}, \mathbf{F}}} \sum_{v=1}^V \gamma_v^2 \|\mathbf{X}_v \mathbf{H}_v - \mathbf{A}_v \mathbf{Z}_v \mathbf{H}_v\|_F^2$$
$$+ \lambda \sum_{v=1}^V \|\mathbf{P}_v \mathbf{Z}_v - \mathbf{F}\|_F^2 + \mu \sum_{v=1}^V \|\mathbf{Z}_v\|_F^2 \tag{4}$$
$$\text{s.t.} \gamma^\top \mathbf{1} = 1, \mathbf{A}_v^\top \mathbf{A}_v = \mathbf{I}_m, \mathbf{P}_v^\top \mathbf{P}_v = \mathbf{I}_m, \mathbf{Z}_v \geq 0,$$
$$\mathbf{Z}_v^\top \mathbf{1}_m = \mathbf{1}_n, \mathbf{F} \mathbf{F}^\top = \mathbf{I}_m,$$

where $\mathbf{Z}_v \mathbf{H}_v$ can be considered as the similarities between m anchors and n_v available samples of the v-th view. For the sake of making the learned anchors $\mathbf{A}_v$ and consistent representation $\mathbf{F}$ more discriminative, we impose orthogonal constraints into them that $\mathbf{A}_v^\top \mathbf{A}_v = \mathbf{I}_m, \mathbf{F} \mathbf{F}^\top = \mathbf{I}_m$. The learned bipartite graph $\mathbf{Z}_v$ should satisfy $\mathbf{Z}_v \geq 0$ and $\mathbf{Z}_v^\top \mathbf{1} = \mathbf{1}$. The $\gamma \in \mathbb{R}^V$ captures the weight contribution of every single view to all. λ is the trade-off to balancing the influence between anchor graph generation and alignment term. μ is the hyperparameter of the regularization term. The framework of our SIMVC-SA is shown in Fig. 2.
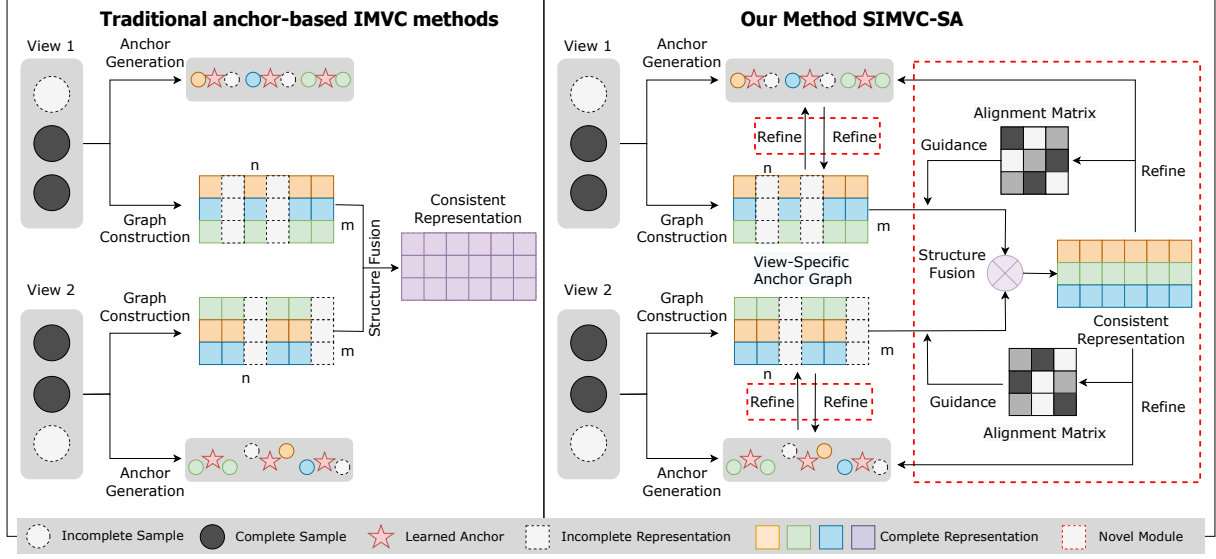
Figure 2: The framework of traditional anchor-based IMVC methods (left) and the proposed SIMVC-SA (right). Different from traditional IMVC methods, the proposed SIMVC-SA proposes a novel structure alignment module and adopts the anchor learning strategy to efficiently enhance the clustering performance.

Although the Eq.(4) appears to be simple, we emphasize the superiority of SIMVC-SA as follows:

(1) **Joint Optimization Model.** Unlike the existing two-stage "aligning then clustering" strategy [65], we propose a joint alignment-clustering framework where the consistent representation $\mathbf{F}$ and the alignment matrix $\mathbf{P}_v$ can be joint optimized to enhance the final clustering performance in our model.

(2) **Flexible Model with No Reference View.** Different from the FMVACC [65], which selects the first view for reference (all views align to the first view) and MvCLN [75] iteratively selects the reference view, we set a consistent representation $\mathbf{F}$ for alignment and optimize it adaptively in our mode which avoids catastrophic performance degradation when the reference view has poor quality[6, 30].

(3) **Soft Alignment Correspondence.** The strict one-to-one mapping proposed in [65] neglects the relationship between the different anchors and brings higher time expenses. Besides, it is too harsh and unreasonable to completely push away the different anchors. Recent work CLIP [5] also noticed this problem. In the proposed method, we relax the original strict constraint to an orthogonal constraint to achieve a soft assignment while effectively reducing the time complexity of the alignment.

3.2 Optimization

The optimization problem in Eq.(4) is a non-convex problem when taking all variables into account. To solve this problem in the section, we develop an iterative optimization algorithm to address it. For the sake of simplifying the optimization procedure, we have that $\mathbf{X}_v \mathbf{H}_v \mathbf{H}_v^\top = \mathbf{X}_v \odot \mathbf{R}_v$, where $\mathbf{R}_v = \mathbf{1}_{d_v} \mathbf{r}^{(v)} \in \mathbb{R}^{d_v \times n}$, $\mathbf{r}^{(v)} = [r_1^{(v)}, \cdots, r_n^{(v)}]$, where $r_i^{(v)} = \sum_{j=1}^{n_v} \mathbf{H}_{ij}^{(v)}$, $\odot$ represents the

Hadamard product. With this transformation, the space complexity drops from $O(vn^2)$ to $O(dn)$.

3.2.1 Optimization of Anchor Matrices $\{\mathbf{A}_v\}_{v=1}^V$. When $\{\mathbf{Z}_v\}_{v=1}^V$, $\{\mathbf{P}_v\}_{v=1}^V$, γ and $\mathbf{F}$ are fixed, the optimization for $\{\mathbf{A}_v\}_{v=1}^V$ can be written as follows:

$$\min_{\{\mathbf{A}_v\}_{v=1}^V} \sum_{v=1}^V \gamma_v^2 \|\mathbf{X}_v \mathbf{H}_v - \mathbf{A}_v \mathbf{Z}_v \mathbf{H}_v\|_F^2,$$
$$\text{s.t.} \quad \mathbf{A}_v^\top \mathbf{A}_v = \mathbf{I}_m. \tag{5}$$

Considering the optimization of each $\mathbf{A}_v$ is independent of the corresponding view. Therefore, we extend the Frobenius norm with traces and remove the irrelevant item, Eq.(5) can be formulated as:

$$\max_{\mathbf{A}_v} \text{Tr}\left(\mathbf{A}_v^\top \mathbf{M}_v\right), \text{ s.t. } \mathbf{A}_v^\top \mathbf{A}_v = \mathbf{I}_m, \tag{6}$$

where $\mathbf{M}_v = \mathbf{X}_v \mathbf{H}_v \mathbf{H}_v^\top \mathbf{Z}_v^\top = (\mathbf{X}_v \odot \mathbf{R}_v) \mathbf{Z}_v^\top$. According to [66], the optimal solution for $\mathbf{A}_v$ is $\mathbf{U}_m \mathbf{V}_m^\top$, where $\mathbf{U}_m$ and $\mathbf{V}_m$ represent the matrices which comprise the first m left singular vectors and right singular vectors of $\mathbf{M}_v$, correspondingly. Their time overhead is $O(nmd + m^2 d)$ to obtain all the optimal $\{\mathbf{A}_v\}_{v=1}^V$, where $d = \sum_{v=1}^V d_v$.

3.2.2 Optimization of Anchor Graphs $\{\mathbf{Z}_v\}_{v=1}^V$. When $\{\mathbf{A}_v\}_{v=1}^V$, $\{\mathbf{P}_v\}_{v=1}^V$, γ and $\mathbf{F}$ are fixed, the optimization of $\{\mathbf{Z}_v\}_{v=1}^V$ can be written as follows:

$$\min_{\{\mathbf{Z}_v\}_{v=1}^V} \sum_{v=1}^V \gamma_v^2 \|\mathbf{X}_v \mathbf{H}_v - \mathbf{A}_v \mathbf{Z}_v \mathbf{H}_v\|_F^2$$
$$+ \lambda \sum_{v=1}^V \|\mathbf{P}_v \mathbf{Z}_v - \mathbf{F}\|_F^2 + \mu \sum_{v=1}^V \|\mathbf{Z}_v\|_F^2, \tag{7}$$
$$\text{s.t.} \quad \mathbf{Z}_v \geq 0, \mathbf{Z}_v^\top \mathbf{1}_m = \mathbf{1}_n.$$

Algorithm 1 Scalable Incomplete Multi-View Clustering with Structure Alignment (SIMVC-SA)

Input: v views incomplete dataset $\{X_v\}_{v=1}^{V}$, the missing index $\{H_v\}_{v=1}^{V}$, and the number of cluster k.
1: Initialize $\{Z_v\}_{v=1}^{V}$, $\{P_v\}_{v=1}^{V}$, γ.
2: **repeat**
3:     Obtain $\{A_v\}_{v=1}^{V}$ with Eq. (5).
4:     Obtain F with Eq. (10).
5:     Obtain $\{Z_v\}_{v=1}^{V}$ with Eq. (7).
6:     Obtain $\{P_v\}_{v=1}^{V}$ with Eq. (11).
7:     Obtain γ with Eq. (12).
8: **until** converged
9: Obtain U by performing SVD on F
Output: Perform k-means on U to obtain discrete labels.

Table 1: Incompete Multiview Datasets in our Experments

Dataset	Size	Clusters	Views	Dimensionality
ORL	400	40	3	4096/3304/6750
ProteinFold	694	27	12	27/27/.../27/27/27
BDGP	2500	5	3	1000/500/250
SUNRGBD	10335	45	2	4096/4096
NUSWIDEOBJ	30000	31	5	65/226/145/74/129
Cifar10	50000	10	3	512/2048/1024
MNIST	60000	10	3	342/1024/64

By removing the irrelevant items, the Eq.(7) can be rewrited as:

$$\min_{Z_v} \text{Tr}\left(Z_v^\top Z_v \left(\gamma_v^2 H_v H_v^\top + (\lambda + \mu)\, I\right)\right)$$
$$- 2\,\text{Tr}\left(Z_v^\top \left(\gamma_v^2 A_v^\top X_v H_v H_v^\top + \lambda P_v^\top F\right)\right) \tag{8}$$
$$\text{s.t.} \quad Z_v \geq 0, \quad Z_v^\top 1_m = 1_n,$$

with $X_v H_v H_v^\top = X_v \odot R_v$. Denoting $z_j^{(v)}$ as the j-th column vector of Z_v, we have

$$\min_{z_j^{(v)}} \frac{1}{2}\left\|z_j^{(v)} - f_j^{(v)}\right\|_F^2, \quad \text{s.t.} z_j^{(v)} \geq 0, z_j^{(v)\top} 1_m = 1, \tag{9}$$

where $f_{ij}^{(v)} = \frac{\gamma_v^2\left[A_v^\top(X_v \odot R_v)\right]_{ij} + \lambda\left[P_v^\top F\right]_{ij}}{\gamma_v^2 r_j^{(v)} + \lambda + \mu}$, $\left[P_v^\top F\right]_{ij}$ denotes the element of the i-th row and j-th column of $P_v^\top F$.

We write the Lagrangian function of Eq.(9) as

$$\mathcal{L}\left(z_j^{(v)}, \alpha_j, \eta_j\right) = \frac{1}{2}\left\|z_j^{(v)} - f_j^{(v)}\right\|_F^2 - \alpha_j\left(z_j^{(v)\top} 1_m - 1\right) - \eta_j^\top z_j^{(v)},$$

where α_j and η_j represent the respective Lagrange multipliers. Their Kahn-Kuhn-Tucker (KKT) conditions can write as

$$\begin{cases} z_j^{(v)} - f_j^{(v)} - \alpha_j 1_m - \eta_j = 0, \\ \eta_j \odot z_j^{(v)} = 0. \end{cases}$$

Together with $z_j^{(v)\top} 1_m = 1$, we can derive the equation below:

$$z_j^{(v)} = \max\left(f_j^{(v)} + \alpha_j 1_m, 0\right),$$

where α_j could be addressed by Newton's method effectively. The time complexity of optimizing $\{Z_v\}_{v=1}^{V}$ is $O(nmd)$.

3.2.3 Optimization of Consistent Representation F. When $\{A_v\}_{v=1}^{V}$, $\{P_v\}_{v=1}^{V}$, $\{Z_v\}_{v=1}^{V}$ and γ are fixed, the optimization for Z can be written as follows:

$$\max_{F} \text{Tr}\left(FQ\right), \quad \text{s.t.} FF^\top = I_m, \tag{10}$$

where $Q = \sum_{v=1}^{V} Z_v^\top P_v^\top$, the optimal solution of F is $\Psi_m \Sigma_m^\top$, where Σ_m and Ψ_m indicate the matrices which comprise the first m left singular vectors and the first m right singular vectors of W_v, correspondingly. It costs $O(nm^2V)$ time.

3.2.4 Optimization of Alignment Matrices $\{P_v\}_{v=1}^{V}$. When $\{A_v\}_{v=1}^{V}$, $\{Z_v\}_{v=1}^{V}$, γ and F are fixed, the optimization for $\{P_v\}_{v=1}^{V}$ can be written as follows:

$$\max_{P_v} \text{Tr}\left(P_v^\top W_v\right), \quad \text{s.t.} P_v^\top P_v = I_m, \tag{11}$$

where $W_v = FZ_v^\top$, Similar to Eq. (10), this problem can be efficiently solved by rank-k truncated SVD.

3.2.5 Optimization of View Weight γ. When $\{A_v\}_{v=1}^{V}$, $\{P_v\}_{v=1}^{V}$, $\{Z_v\}_{v=1}^{V}$ and Z are fixed, the optimization for γ can be written as follows:

$$\min_{\gamma} \sum_{v=1}^{V} \gamma_v^2 \tau_v, \quad \text{s.t.} \gamma^\top 1_V = 1, \gamma \geq 0, \tag{12}$$

where $\tau_v = \|X_v H_v - A_v Z_v H_v\|_F^2$. By Cauchy–Schwarz inequality, the view weight γ can be acquired by

$$\gamma_v = \frac{1/\tau_v}{\sum_{v=1}^{V} 1/\tau_v}. \tag{13}$$

It consumes $O(nmd)$ time. Algorithm 1 summarises the entire optimization procedure for addressing Eq.(4).

3.3 Discussions

3.3.1 Convergence. As the iterations proceed, five variables of the above optimization procedure will be separately addressed. As each sub-optimization problem reaches the global optimum, the objective value monotonically decreases until the convergence condition is attained [1]. Furthermore, because it is easy to prove that the lower boundary of the objective function is zero, our proposed SIMVC-SA can converge to the local optimum.

3.3.2 Time Complexity. The time overhead of SIMVC-SA is composed of five optimization processes, as previously mentioned. The time overhead of updating $\{A_v\}_{v=1}^{V}$ is $O(nmd + m^2d)$. When updating $\{Z_v\}_{v=1}^{V}$ and γ need $O(nmd)$. When analytically obtaining $\{P_v\}_{v=1}^{V}$, it costs $O((nm^2 + m^3)V)$ for all columns. The time overhead of calculating F is $O(nm^2V)$. As a result, the total time overhead of the optimization procedure is $O\left(n\left(md + m^2V\right) + m^3V + m^2d\right)$. Consequently, the computational complexity of SIMVC-SA is $O(n)$, which is linearly related to the number of samples.

4 EXPERIMENT

4.1 Datasets

Seven wide-used datasets are adopted to evaluate the effectiveness of the proposed algorithm, including ORL, ProteinFold, BDGP,

Table 2: Empirical evaluation and comparison of SIMVC-SA with twelve baseline methods on seven benchmark datasets.

Methods	BSV	MIC	MKKM-IK	AWP	DAIMC	APMC	UEAF	MKKM-IK-MKC	EEIMVC	FLSD	V^3H	FIMVC-VIA	Proposed
ACC (%)													
ORL	24.32±0.89	37.56±1.66	59.80±2.44	68.69±0.00	68.03±2.32	65.58±1.91	60.25±2.50	64.95±2.62	73.24±2.54	48.09±1.85	67.03±1.45	76.36±2.79	76.06±2.36
ProteinFold	22.25±0.53	15.99±0.78	26.03±1.06	28.00±0.00	28.65±1.65	N/A	28.72±1.53	17.99±0.83	27.75±1.67	25.98±1.33	17.33±0.48	28.15±1.26	30.17±1.23
BDGP	34.96±1.06	25.37±0.61	32.17±0.24	23.62±0.00	28.12±0.01	28.12±0.01	44.88±0.02	40.77±0.20	44.00±0.05	42.96±0.53	43.63±0.75	39.84±0.16	48.11±0.21
SUNRGBD	6.14±0.08	14.61±0.54	11.35±0.31	17.01±0.00	17.03±0.65	17.34±0.57	15.35±0.41	16.81±0.49	16.74±0.49	14.42±0.34	N/A	16.88±0.48	17.18±0.48
NUSWIDEOBJ	12.05±0.03	N/A	N/A	N/A	13.79±0.37	N/A	N/A	N/A	12.73±0.16	N/A	N/A	12.96±0.12	15.40±0.32
Cifar10	N/A	N/A	N/A	N/A	95.81±0.45	N/A	N/A	N/A	N/A	N/A	N/A	96.16±0.00	96.32±0.22
MNIST	N/A	N/A	N/A	N/A	97.57±0.31	N/A	N/A	N/A	N/A	N/A	N/A	98.16±0.01	98.40±0.04
NMI (%)													
ORL	48.49±0.90	56.44±1.00	75.95±1.33	83.79±0.00	82.89±1.06	80.20±0.82	76.16±1.25	79.76±1.41	85.37±1.32	67.91±1.28	81.05±0.61	88.08±1.31	87.53±1.25
ProteinFold	27.60±0.59	16.64±1.02	33.70±0.84	36.17±0.00	37.67±1.08	N/A	38.18±0.88	24.88±0.84	36.03±0.98	35.66±0.79	22.75±0.53	36.22±0.96	37.72±0.98
BDGP	12.88±0.94	4.47±0.70	7.41±0.16	4.68±0.00	8.68±0.01	8.68±0.01	23.55±0.04	16.35±0.13	19.91±0.09	18.95±0.06	24.15±0.40	15.11±0.10	23.67±0.14
SUNRGBD	3.27±0.08	21.27±0.35	15.27±0.25	23.60±0.00	21.53±0.43	22.46±0.25	21.72±0.22	20.48±0.28	20.84±0.28	20.82±0.17	N/A	21.48±0.33	22.52±0.23
NUSWIDEOBJ	2.68±0.03	N/A	N/A	N/A	11.70±0.36	N/A	N/A	N/A	10.31±0.16	N/A	N/A	10.27±0.07	11.78±0.11
Cifar10	N/A	N/A	N/A	N/A	90.47±0.35	N/A	N/A	N/A	N/A	N/A	N/A	91.18±0.00	91.21±0.18
MNIST	N/A	N/A	N/A	N/A	93.89±0.53	N/A	N/A	N/A	N/A	N/A	N/A	95.76±0.02	95.58±0.00
Purity (%)													
ORL	26.80±0.92	40.81±1.40	62.79±2.11	70.42±0.00	71.82±1.79	69.24±1.48	63.90±1.90	67.68±2.34	76.09±2.19	50.88±1.72	70.22±1.09	78.59±2.29	78.79±2.08
ProteinFold	25.89±0.60	19.78±0.84	30.91±1.04	31.97±0.00	34.99±1.54	N/A	35.47±1.16	22.73±0.87	33.13±1.30	32.82±0.97	22.24±0.51	33.67±1.12	35.97±1.14
BDGP	36.75±0.89	25.67±0.59	33.42±0.18	24.02±0.00	28.46±0.01	28.46±0.01	45.92±0.01	41.10±0.13	46.40±0.05	44.41±0.03	45.38±0.50	40.12±0.15	49.05±0.19
SUNRGBD	13.06±0.16	32.36±0.59	27.06±0.45	37.56±0.00	34.39±0.59	33.19±0.47	33.37±0.48	32.92±0.50	33.58±0.48	32.41±0.39	N/A	34.34±0.61	34.51±0.51
NUSWIDEOBJ	13.72±0.04	N/A	N/A	N/A	23.41±0.63	N/A	N/A	N/A	21.83±0.21	N/A	N/A	21.97±0.12	23.71±0.27
Cifar10	N/A	N/A	N/A	N/A	95.81±0.45	N/A	N/A	N/A	N/A	N/A	N/A	96.23±0.00	96.29±0.25
MNIST	N/A	N/A	N/A	N/A	97.57±0.31	N/A	N/A	N/A	N/A	N/A	N/A	98.24±0.02	98.42±0.05
Fscore (%)													
ORL	9.01±0.69	17.30±1.18	46.32±2.50	58.73±0.00	56.84±2.87	50.70±2.55	42.53±2.74	53.30±2.91	63.67±2.85	31.17±2.00	54.27±1.40	68.25±3.25	67.62±2.85
ProteinFold	12.30±0.08	10.35±0.38	14.35±0.81	12.38±0.00	16.97±1.09	N/A	16.31±1.22	8.92±0.48	15.62±1.19	14.57±1.02	10.33±0.13	15.53±1.00	16.81±1.04
BDGP	28.76±0.61	29.88±0.06	25.25±0.08	32.53±0.00	31.21±0.00	31.21±0.00	33.69±0.03	30.15±0.10	32.93±0.04	34.29±0.01	35.31±0.24	31.65±0.10	35.40±0.12
SUNRGBD	6.89±0.01	9.45±0.24	7.10±0.14	11.58±0.00	10.71±0.35	10.96±0.25	10.20±0.16	10.06±0.21	10.20±0.20	12.09±0.00	N/A	10.33±0.24	11.74±0.10
NUSWIDEOBJ	10.95±0.00	N/A	N/A	N/A	8.58±0.19	N/A	N/A	N/A	7.81±0.08	N/A	N/A	7.87±0.07	11.77±0.12
Cifar10	N/A	N/A	N/A	N/A	92.16±0.68	N/A	N/A	N/A	N/A	N/A	N/A	92.79±0.00	92.90±0.19
MNIST	N/A	N/A	N/A	N/A	95.28±0.57	N/A	N/A	N/A	N/A	N/A	N/A	96.66±0.00	96.92±0.00

SUNRGBD, NUSWIDEOBJ, Cifar10, and MNIST. The elaborate information of these datasets is listed in Tab. 1. For the above datasets, we remove samples randomly on each view to obtain its incomplete version. Specifically, according to [26], with the principle that each sample appears in at least one view, we generate incomplete datasets at 0.1 intervals from 0.1 to 0.9.

4.2 Compared Methods and Setting

Along with our proposed SIMVC-SA, we run twelve state-of-the-art incomplete multi-view clustering methods for comparison, including Best Single View (BSV) [47], MIC Views Clustering via Weighted NMF With $\ell_{2,1}$ Regularization (MIC) [55], Multiple Kernel k-Means With Incomplete Kernels (MKKM-IK) [39], Multiview Clustering via Adaptively Weighted Procrustes (AWP) [50], Doubly Aligned Incomplete Multi-view Clustering (DAIMC) [9], Anchor-based partial multi-view clustering (APMC) [7], Unified Embedding Alignment With Missing Views Inferring for Incomplete Multiview Clustering (UEAF) [69], Multiple Kernel k k-Means With Incomplete Kernels and Multiple Kernel Clustering (MKKM-IK-MKC) [40], Efficient and Effective Regularized Incomplete Multiview Clustering (EEIMVC) [38], Generalized IMVC With Flexible locality Structure Diffusion (FLSD) [71], View Variation and View Heredity for Incomplete Multiview Clustering (V^3H) [3] and Fast Incomplete Multi-View Clustering With View Independent Anchors (FIMVC-VIA) [36].

For all the algorithms mentioned above, we set their parameters as their recommended range. In the proposed method, we adjusted λ to $[10^{-4}, 10^{-2}, 1, 10^2, 10^4]$, μ to $[0, 10^{-4}, 10^{-2}, 1, 10^2, 10^4]$, and the anchor numbers of $[k, 2k, 5k]$ using a mesh search scheme. In

addition, we repeated each experiment 10 cycles to calculate the average performance and standard bias. To assess the clustering performance, we employ four well-used criteria consisting of accuracy (ACC), normalized mutual information (NMI), Purity, and Fscore. All experiments were conducted on a desktop computer with Intel core i9-10900X CPU and 64G RAM, MATLAB 2020b (64-bit).

4.3 Experimental Results

Tab. 2 reports the clustering results on seven benchmark datasets. The best results are marked in red, while the second-best results are marked in blue. "N/A" indicates the unavailable results due to time-out or out-of-memory errors. Besides, we compare the ACC of all methods with different missing rates in Fig. 3. According to the results, we have the following conclusions:

(1) Compared with existing IMVC methods, our proposed algorithm demonstrates the best performance in most datasets. The recently proposed FIMVC-VIA method shows better performance than other methods, which demonstrates its superiority in incomplete datasets. In terms of ACC, our SIMVC-SA achieves better performance than FIMVC-VIA on the Protein-Fold, BDGP, SUNRGBD, NUSWIDEOBJ, and Cifar10 datasets, i.e., 2.02%, 8.27%, 0.3%, 2.44%, and 0.16%, which demonstrates the effectiveness of view-specific representation and cross-view alignment strategy.

(2) Compared to traditional subspace-based IMVC methods, our anchor-based method achieves the best performance in most cases and is applicable to various large-scale datasets.

(3) As shown in Fig. 3, we can observe that most IMVC methods show greater fluctuations in performance with the missing
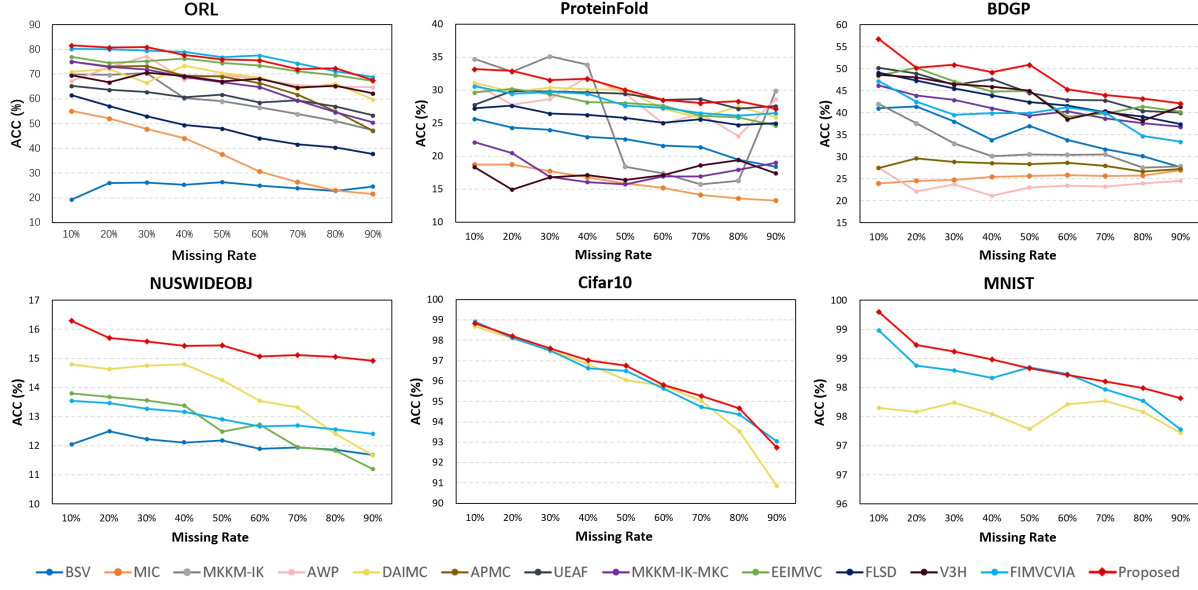
Figure 3: Clustering performance of SIMVC-SA on benchmark datasets with different missing ratio.
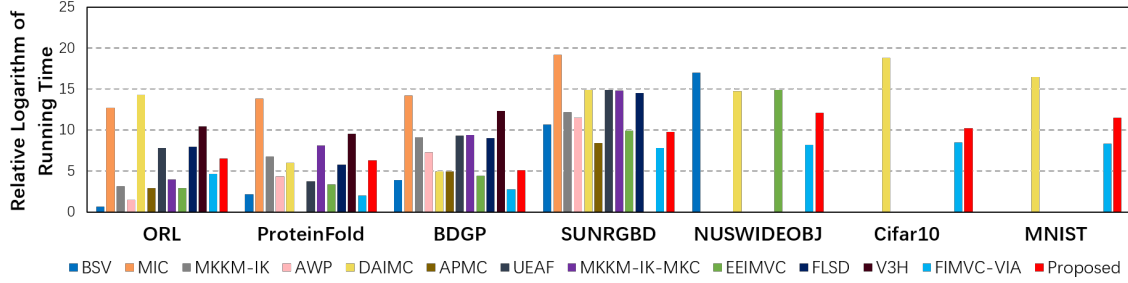

Figure 4: Time Comparison of Different IMVC Methods on Seven Incomplete Datasets
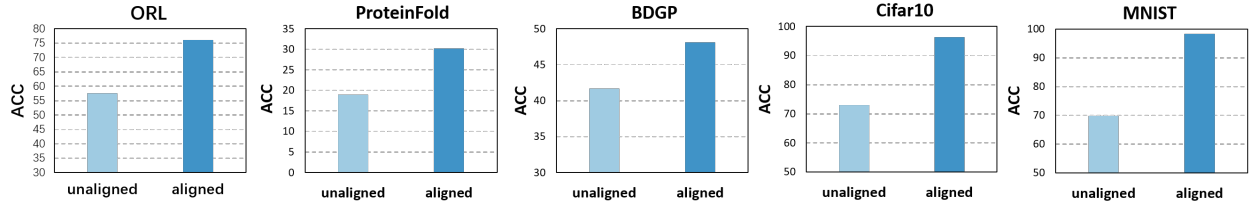

Figure 5: The ablation study of our structural alignment strategy on five benchmark datasets. "Unaligned" indicates without our structural alignment strategy.

rate rising, while our method is more stable. We conjecture that this is because the alignment of the representations well complements the missing information of different views.

4.4 Running Time Comparison

To validate the computational efficiency of the proposed SIMVC-SA, we plot the average running time of each algorithm on seven benchmark datasets in Fig. 4. The results of some compared algorithms

on large-scale datasets are not reported due to memory overflow errors. As shown in the Fig. 4, we can observe that

(1) Compared to full graph-based clustering methods, the proposed SIMVC-SA significantly reduces run time through the construction of anchor graphs.

(2) Compared to the anchor-based IMVC approach, i.e., FIMVC-VIA, the proposed SIMVC-SA requires more time consumption, mainly due to our view-specific representation and structure alignment strategy, the extra computational complexity
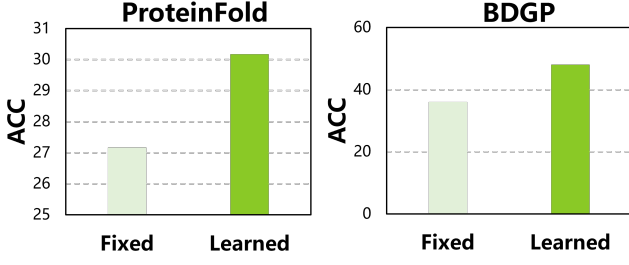
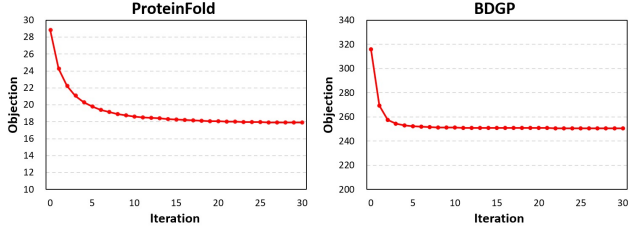**Figure 6: The ablation study of our anchor learning strategy.**

Figure 7: Objectives of the proposed method.

increases with the number of views, which is most obvious in NUSWIDEOBJ (5 views). General, the extra time spent is worthwhile since our proposed SIMVC-SA demonstrates its superiority over FIMVC-VIA in most datasets.

4.5 Ablation Study

Structural Alignment Strategy. The structural alignment strategy is the main contribution of this paper. To further demonstrate the effectiveness of this strategy, we present the experimental results of the ablation study in Fig. 5, where "Unaligned" indicates not using our structural alignment strategy. In our experimental setting, we fixed the alignment matrix $\mathbf{P}$ in the optimization process to obtain the final clustering result. The effectiveness of the proposed strategy can be clearly demonstrated in Fig. 5. In terms of ACC, the proposed structural alignment strategy improves the algorithm performance on the ORL, ProteinFold, BDGP, Cifar10, and MNIST datasets by **18.46%,11.16%, 6.43%, 23.38%**, and **28.55**% respectively, which demonstrates the effectiveness of our strategy.

Anchor Learning Strategy. We conducted ablation experiments with the proposed anchor learning strategy, as shown in Fig. 7. "Fixed" indicates initializing anchors by k-means without updating during the optimization process. Compared to the above methods, our approach significantly improves the clustering performance and avoids the high time expenditure of k-means.

4.6 Convergence and Sensitivity

We conducted several experiments to exhibit the convergence of the proposed SIMVC-SA. As shown in Fig. 6, the objective value of our algorithm is monotonically decreasing in each iteration. These results clearly verify the convergence of our proposed algorithm.

To investigate the sensitivity of SIMVC-SA to the number of anchors m, we investigated how our performance shifts for different numbers of anchors. As shown in Fig. 8, the number of anchors has little effect on the performance of our algorithm. Moreover, two
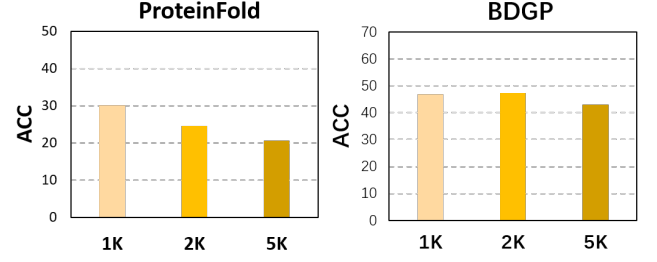

Figure 8: Sensitivity analysis of anchor number m of out method on two benchmark datasets.
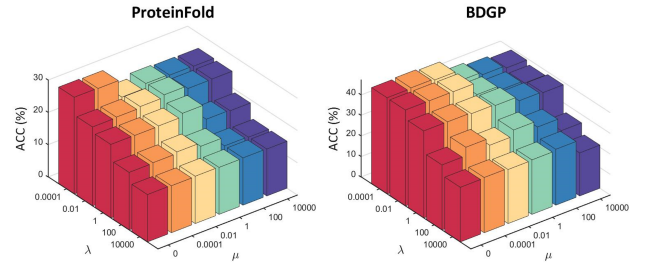

Figure 9: Sensitivity analysis of λ and μ of out method on two benchmark datasets.

hyperparameters, λ, and μ, are used in our method, λ is the structural alignment parameter, and μ is the coefficient of the sparsity regularization term. As is shown in Fig. 9, we conducted comparative experiments to indicate the effect of these two parameters on performance.

5 CONCLUSION

In this paper, we propose a novel incomplete anchor graph learning framework termed Scalable Incomplete Multi-View Clustering with Structure Alignment (SIMVC-SA). Specially, we construct the incomplete anchor graph on each view in terms of the unaligned anchor. Besides, a novel structure alignment module is proposed to refine the cross-view anchor correspondence. Meanwhile, the anchor graph construction and alignment are jointly optimized in our unified framework to enhance clustering quality. Through anchor graph construction instead of full graphs, the time and space complexity of our proposed SIMVC-SA is proven to be linearly related to the number of samples. Extensive experiments on seven incomplete benchmark datasets demonstrate the effectiveness and efficiency of our proposed method. In the future, we will explore more flexible alignment strategies. For example, how to align the anchor with different numbers.

6 ACKNOWLEDGMENTS

REFERENCES

[1] James C. Bezdek and Richard J. Hathaway. 2003. Convergence of Alternating Optimization. Neural, Parallel Sci. Comput. (2003).

[2] Xiao Cai, Feiping Nie, and Heng Huang. 2013. Multi-view k-means clustering on big data. In Twenty-Third International Joint conference on artificial intelligence.

[3] Xiang Fang, Yuchong Hu, Pan Zhou, and Dapeng Oliver Wu. 2020. V ^{3} H: View Variation and View Heredity for Incomplete Multiview Clustering. IEEE Transactions on Artificial Intelligence (2020).

[4] Hongchang Gao, Feiping Nie, Xuelong Li, and Heng Huang. 2015. Multi-view subspace clustering. In Proceedings of the IEEE international conference on computer vision. 4238–4246.

[5] Yuting Gao, Jinfeng Liu, Zihan Xu, Tong Wu, Wei Liu, Jie Yang, Ke Li, and Xing Sun. 2023. SoftCLIP: Softer Cross-modal Alignment Makes CLIP Stronger. arXiv preprint arXiv:2303.17561 (2023).

[6] Fengjiao Gong, Yuzhou Nie, and Hongteng Xu. 2022. Gromov-Wasserstein multimodal alignment and clustering. In Proceedings of the 31st ACM International Conference on Information & Knowledge Management. 603–613.

[7] Jun Guo and Jiahui Ye. 2019. Anchors bring ease: An embarrassingly simple approach to partial multi-view clustering. In Proc. of AAAI.

[8] Xuemei Han, Zhenwen Ren, Chuanyun Zou, and Xiaojian You. 2022. Incomplete multi-view subspace clustering based on missing-sample recovering and structural information learning. Expert Systems with Applications 208 (2022), 118165.

[9] Menglei Hu and Songcan Chen. 2018. Doubly Aligned Incomplete Multi-view Clustering. In Proc. of IJCAI.

[10] Menglei Hu and Songcan Chen. 2019. One-pass incomplete multi-view clustering. In Proc. of AAAI.

[11] Shudong Huang, Zhao Kang, Ivor W Tsang, and Zenglin Xu. 2019. Auto-weighted multi-view clustering via kernelized graph learning. Pattern Recognition (2019).

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

[13] Jiaqi Jin, Siwei Wang, Zhibin Dong, Xinwang Liu, and En Zhu. 2023. Deep Incomplete Multi-view Clustering with Cross-view Partial Sample and Prototype Alignment. arXiv preprint arXiv:2303.15689 (2023).

[14] Zhao Kang, Xinjia Zhao, Chong Peng, Hongyuan Zhu, Joey Tianyi Zhou, Xi Peng, Wenyu Chen, and Zenglin Xu. 2020. Partition level multiview subspace clustering. Neural Networks (2020).

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

[16] Aparajita Khan and Pradipta Maji. 2019. Approximate graph Laplacians for multimodal data clustering. IEEE transactions on pattern analysis and machine intelligence 43, 3 (2019), 798–813.

[17] Abhishek Kumar, Piyush Rai, and Hal Daume. 2011. Co-regularized multi-view spectral clustering. In Advances in neural information processing systems.

[18] Liang Li, Siwei Wang, Xinwang Liu, En Zhu, Li Shen, Kenli Li, and Keqin Li. 2022. Local Sample-Weighted Multiple Kernel Clustering With Consensus Discriminative Graph. IEEE Transactions on Neural Networks and Learning Systems (2022), 1–14. https://doi.org/10.1109/TNNLS.2022.3184970

[19] Liang Li, Junpu Zhang, Siwei Wang, Xinwang Liu, Kenli Li, and Keqin Li. 2023. Multi-View Bipartite Graph Clustering With Coupled Noisy Feature Filter. IEEE Transactions on Knowledge and Data Engineering (2023), 1–13. https://doi.org/10.1109/TKDE.2023.3268215

[20] Miaomiao Li, Siwei Wang, Xinwang Liu, and Suyuan Liu. 2022. Parameter-Free and Scalable Incomplete Multiview Clustering With Prototype Graph. IEEE Transactions on Neural Networks and Learning Systems (2022).

[21] Qian Li, Shu Guo, Cheng Ji, Xutan Peng, Shiyao Cui, and Jianxin Li. 2023. Dual-Gated Fusion with Prefix-Tuning for Multi-Modal Relation Extraction. arXiv preprint arXiv:2306.11020 (2023).

[22] Ruihuang Li, Changqing Zhang, Qinghua Hu, Pengfei Zhu, and Zheng Wang. 2019. Flexible Multi-View Representation Learning for Subspace Clustering.. In Proc. of IJCAI.

[23] Shao-Yuan Li, Yuan Jiang, and Zhi-Hua Zhou. 2014. Partial multi-view clustering. In Proc. of AAAI.

[24] Xiang-Long Li, Man-Sheng Chen, Chang-Dong Wang, and Jian-Huang Lai. 2022. Refining graph structure for incomplete multi-view clustering. IEEE Transactions on Neural Networks and Learning Systems (2022).

[25] Yingming Li, Ming Yang, and Zhongfei Zhang. 2018. A survey of multi-view representation learning. IEEE transactions on knowledge and data engineering 31, 10 (2018), 1863–1883.

[26] Zhenglai Li, Chang Tang, Xiao Zheng, Xinwang Liu, Wei Zhang, and En Zhu. 2022. High-order correlation preserved incomplete multi-view subspace clustering. IEEE Transactions on Image Processing 31 (2022), 2067–2080.

[27] Ke Liang, Lingyuan Meng, Meng Liu, Yue Liu, Wenxuan Tu, Siwei Wang, Sihang Zhou, X Liu, and F Sun. 2022. A Survey of Knowledge Graph Reasoning on Graph Types: Static, Dynamic, and Multimodal. (2022).

[28] Ke Liang, Sifan Wu, Jiayi Gu, et al. 2021. Mka: A scalable medical knowledge-assisted mechanism for generative models on medical conversation tasks. Computational and mathematical methods in medicine 2021 (2021).

[29] Ke Liang, Sihang Zhou, Yue Liu, Lingyuan Meng, Meng Liu, and Xinwang Liu. 2023. Structure Guided Multi-modal Pre-trained Transformer for Knowledge Graph Reasoning. arXiv preprint arXiv:2307.03591 (2023).

[30] Jia-Qi Lin, Man-Sheng Chen, Chang-Dong Wang, and Haizhang Zhang. 2022. A Tensor Approach for Uncoupled Multiview Clustering. IEEE Transactions on Cybernetics (2022).

[31] Jia-Qi Lin, Xiang-Long Li, Man-Sheng Chen, Chang-Dong Wang, and Haizhang Zhang. 2022. Incomplete Data Meets Uncoupled Case: A Challenging Task of Multiview Clustering. IEEE Transactions on Neural Networks and Learning Systems (2022).

[32] Yijie Lin, Yuanbiao Gou, Xiaotian Liu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. 2022. Dual contrastive prediction for incomplete multi-view representation learning. IEEE Transactions on Pattern Analysis and Machine Intelligence (2022).

[33] Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. 2021. COMPLETER: Incomplete multi-view clustering via contrastive prediction. In Proc. of CVPR.

[34] Meng Liu, Ke Liang, Bin Xiao, Sihang Zhou, Wenxuan Tu, Yue Liu, Xihong Yang, and Xinwang Liu. 2023. Self-Supervised Temporal Graph learning with Temporal and Structural Intensity Alignment. arXiv preprint arXiv:2302.07491 (2023).

[35] Meng Liu, Yue Liu, Ke Liang, Siwei Wang, Sihang Zhou, and Xinwang Liu. 2023. Deep Temporal Graph Clustering. arXiv preprint arXiv:2305.10738 (2023).

[36] Suyuan Liu, Xinwang Liu, Siwei Wang, Xin Niu, and En Zhu. 2022. Fast Incomplete Multi-View Clustering With View-Independent Anchors. IEEE Transactions on Neural Networks and Learning Systems (2022).

[37] Wei Liu, Junfeng He, and Shih-Fu Chang. 2010. Large graph construction for scalable semi-supervised learning. In ICML.

[38] X Liu, M Li, C Tang, J Xia, J Xiong, L Liu, M Kloft, and E Zhu. 2020. Efficient and Effective Regularized Incomplete Multi-view Clustering. IEEE transactions on pattern analysis and machine intelligence (2020).

[39] Xinwang Liu, Miaomiao Li, Lei Wang, Yong Dou, Jianping Yin, and En Zhu. 2017. Multiple Kernel k-Means with Incomplete Kernels. In Thirty-First AAAI Conference on Artificial Intelligence.

[40] Xinwang Liu, Xinzhong Zhu, Miaomiao Li, Lei Wang, En Zhu, Tongliang Liu, Marius Kloft, Dinggang Shen, Jianping Yin, and Wen Gao. 2020. Multiple Kernel k k -Means with Incomplete Kernels. IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).

[41] Yue Liu, Ke Liang, Jun Xia, Sihang Zhou, Xihong Yang, , Xinwang Liu, and Z. Stan Li. 2023. Dink-Net: Neural Clustering on Large Graphs. In Proc. of ICML.

[42] Yue Liu, Wenxuan Tu, Sihang Zhou, Xinwang Liu, Linxuan Song, Xihong Yang, and En Zhu. 2022. Deep Graph Clustering via Dual Correlation Reduction. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 36. 7603–7611.

[43] Yue Liu, Jun Xia, Sihang Zhou, Siwei Wang, Xifeng Guo, Xihong Yang, Ke Liang, Wenxuan Tu, Z. Stan Li, and Xinwang Liu. 2022. A Survey of Deep Graph Clustering: Taxonomy, Challenge, and Application. arXiv preprint arXiv:2211.12875 (2022).

[44] Yue Liu, Xihong Yang, Sihang Zhou, and Xinwang Liu. 2023. Simple contrastive graph clustering. IEEE Transactions on Neural Networks and Learning Systems (2023).

[45] Yue Liu, Xihong Yang, Sihang Zhou, Xinwang Liu, Zhen Wang, Ke Liang, Wenxuan Tu, Liang Li, Jingcan Duan, and Cancan Chen. 2023. Hard Sample Aware Network for Contrastive Deep Graph Clustering. In Proc. of AAAI.

[46] Yujie Mo, Liang Peng, Jie Xu, Xiaoshuang Shi, and Xiaofeng Zhu. 2022. Simple Unsupervised Graph Representation Learning. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). 7797–7805.

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

[48] Feiping Nie, Guohao Cai, Jing Li, and Xuelong Li. 2017. Auto-weighted multi-view learning for image clustering and semi-supervised classification. IEEE Transactions on Image Processing (2017).

[49] Feiping Nie, Jing Li, Xuelong Li, et al. 2016. Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification.. In IJCAI. 1881–1887.

[50] Feiping Nie, Lai Tian, and Xuelong Li. 2018. Multiview clustering via adaptively weighted procrustes. In Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining. 2022–2030.

[51] Xi Peng, Zhenyu Huang, Jiancheng Lv, Hongyuan Zhu, and Joey Tianyi Zhou. 2019. COMIC: Multi-view clustering without parameter selection. In Proc. of ICML.

[52] Qianyao Qiang, Bin Zhang, Fei Wang, and Feiping Nie. 2021. Fast multi-view discrete clustering with anchor graphs. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35. 9360–9367.

[53] Yazhou Ren, Kangrong Hu, Xinyi Dai, Lili Pan, Steven CH Hoi, and Zenglin Xu. 2019. Semi-supervised deep embedded clustering. Neurocomputing 325 (2019), 121–130.

[54] Zhenwen Ren, Simon X Yang, Quansen Sun, and Tao Wang. 2020. Consensus affinity graph learning for multiple kernel clustering. IEEE Transactions on Cybernetics 51, 6 (2020), 3273–3284.

[55] Weixiang Shao, Lifang He, and S Yu Philip. 2015. Multiple incomplete views clustering via weighted nonnegative matrix factorization with $\ell_{2,1}$ regularization. In Joint European conference on machine learning and knowledge discovery in databases.

[56] Xinhang Wan, Jiyuan Liu, Weixuan Liang, Xinwang Liu, Yi Wen, and En Zhu. 2022. Continual Multi-View Clustering. In Proceedings of the 30th ACM International Conference on Multimedia (Lisboa, Portugal) (MM '22). Association for Computing Machinery, New York, NY, USA, 3676–3684. https://doi.org/10.1145/3503161.3547864

[57] Xinhang Wan, Jiyuan Liu, Xinwang Liu, Siwei Wang, Yi Wen, Tianjiao Wan, Li Shen, and En Zhu. 2023. One-step Multi-view Clustering with Diverse Representation. arXiv:2306.05437 [cs.LG]

[58] Xinhang Wan, Xinwang Liu, Jiyuan Liu, Siwei Wang, Yi Wen, Weixuan Liang, En Zhu, Zhe Liu, and Lu Zhou. 2023. Auto-weighted Multi-view Clustering for Large-scale Data. arXiv:2303.01983 [cs.LG]

[59] Xinhang Wan, Bin Xiao, Xinwang Liu, Jiyuan Liu, Weixuan Liang, and En Zhu. 2023. Fast Continual Multi-View Clustering with Incomplete Views. arXiv:2306.02389 [cs.LG]

[60] Chang-Dong Wang, Man-Sheng Chen, Ling Huang, Jian-Huang Lai, and S Yu Philip. 2020. Smoothness regularized multiview subspace clustering with kernel learning. IEEE Transactions on Neural Networks and Learning Systems (2020).

[61] Hao Wang, Yan Yang, and Bing Liu. 2019. GMC: Graph-based multi-view clustering. IEEE Transactions on Knowledge and Data Engineering (2019).

[62] Hao Wang, Linlin Zong, Bing Liu, Yan Yang, and Wei Zhou. 2019. Spectral perturbation meets incomplete multi-view data. arXiv preprint arXiv:1906.00098 (2019).

[63] Qianqian Wang, Zhengming Ding, Zhiqiang Tao, Quanxue Gao, and Yun Fu. 2018. Partial multi-view clustering via consistent GAN. In Proc. of ICDM.

[64] Qianqian Wang, Huanhuan Lian, Gan Sun, Quanxue Gao, and Licheng Jiao. 2020. ICMSC: Incomplete cross-modal subspace clustering. IEEE Transactions on Image Processing 30 (2020), 305–317.

[65] Siwei Wang, Xinwang Liu, Suyuan Liu, Jiaqi Jin, Wenxuan Tu, Xinzhong Zhu, and En Zhu. 2022. Align then Fusion: Generalized Large-scale Multi-view Clustering with Anchor Matching Correspondences. arXiv preprint arXiv:2205.15075 (2022).

[66] Siwei Wang, Xinwang Liu, En Zhu, Chang Tang, Jiyuan Liu, Jingtao Hu, Jingyuan Xia, and Jianping Yin. 2019. Multi-view Clustering via Late Fusion Alignment Maximization.. In IJCAI. 3778–3784.

[67] Yang Wang. 2021. Survey on deep multi-modal data analytics: Collaboration, rivalry, and fusion. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 17, 1s (2021), 1–25.

[68] Jie Wen, Zhihao Wu, Zheng Zhang, Lunke Fei, Bob Zhang, and Yong Xu. 2021. Structural deep incomplete multi-view clustering network. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 3538–3542.

[69] Jie Wen, Zheng Zhang, Yong Xu, Bob Zhang, Lunke Fei, and Hong Liu. 2019. Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In Proc. of AAAI.

[70] Jie Wen, Zheng Zhang, Yong Xu, Bob Zhang, Lunke Fei, and Guo-Sen Xie. 2020. CDIMC-net: Cognitive Deep Incomplete Multi-view Clustering Network.. In Proc. of IJCAI.

[71] Jie Wen, Zheng Zhang, Zhao Zhang, Lunke Fei, and Meng Wang. 2020. Generalized incomplete multiview clustering with flexible locality structure diffusion. IEEE transactions on cybernetics (2020).

[72] Jie Xu, Chao Li, Liang Peng, Yazhou Ren, Xiaoshuang Shi, Heng Tao Shen, and Xiaofeng Zhu. 2023. Adaptive Feature Projection With Distribution Alignment for Deep Incomplete Multi-View Clustering. IEEE Transactions on Image Processing 32 (2023), 1354–1366.

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

[74] Mouxing Yang, Yunfan Li, Peng Hu, Jinfeng Bai, Jiancheng Lv, and Xi Peng. 2022. Robust multi-view clustering with incomplete information. IEEE Transactions on Pattern Analysis and Machine Intelligence 45, 1 (2022), 1055–1069.

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

[76] Xihong Yang, Yue Liu, Sihang Zhou, Xinwang Liu, and En Zhu. 2022. Mixed Graph Contrastive Network for Semi-Supervised Node Classification. arXiv preprint arXiv:2206.02796 (2022).

[77] Xihong Yang, Yue Liu, Sihang Zhou, Siwei Wang, Xinwang Liu, and En Zhu. 2022. Contrastive Deep Graph Clustering with Learnable Augmentation. arXiv preprint arXiv:2212.03559 (2022).

[78] Xihong Yang, Yue Liu, Sihang Zhou, Siwei Wang, Wenxuan Tu, Qun Zheng, Xinwang Liu, Liming Fang, and En Zhu. 2023. Cluster-guided Contrastive Graph Clustering Network. arXiv preprint arXiv:2301.01098 (2023).

[79] Kun Zhan, Feiping Nie, Jing Wang, and Yi Yang. 2018. Multiview consensus graph clustering. IEEE Transactions on Image Processing (2018).

[80] Kun Zhan, Changqing Zhang, Junpeng Guan, and Junsheng Wang. 2017. Graph learning for multiview clustering. IEEE transactions on cybernetics 48, 10 (2017), 2887–2895.

[81] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao. 2017. Latent multi-view subspace clustering. In Proc. of CVPR.

[82] Junpu Zhang, Liang Li, Siwei Wang, Jiyuan Liu, Yue Liu, Xinwang Liu, and En Zhu. 2022. Multiple Kernel Clustering with Dual Noise Minimization. In Proceedings of the 30th ACM International Conference on Multimedia (Lisboa, Portugal) (MM '22). Association for Computing Machinery, New York, NY, USA, 3440–3450. https://doi.org/10.1145/3503161.3548334

[83] Handong Zhao, Zhengming Ding, and Yun Fu. 2017. Multi-view clustering via deep matrix factorization. In Proc. of AAAI.

[84] Handong Zhao, Hongfu Liu, and Yun Fu. 2016. Incomplete multi-modal visual data grouping.. In Proc. of IJCAI.

[85] Yu Zhao, Xiangrui Cai, Yike Wu, Haiwei Zhang, Ying Zhang, Guoqing Zhao, and Ning Jiang. 2022. Mose: Modality split and ensemble for multimodal knowledge graph completion. arXiv preprint arXiv:2210.08821 (2022).

[86] Pengfei Zhu, Binyuan Hui, Changqing Zhang, Dawei Du, Longyin Wen, and Qinghua Hu. 2019. Multi-view deep subspace clustering networks. arXiv preprint arXiv:1908.01978 (2019).