Fast Incomplete Multi-view Clustering with Adaptive Similarity Completion and Reconstruction

Deng Xu, Chao Zhang*, Cong Guo, Chunlin Chen, Huaxiong Li*

Department of Control Science and Intelligence Engineering, Nanjing University
{dengxu, chzhang, congguo}@smail.nju.edu.cn, {clchen, huaxiongli}@nju.edu.cn

Abstract

Recently, anchor-based incomplete multi-view clustering (IMVC) has been widely adopted for fast clustering, but most existing approaches still encounter some issues: (1) They generally rely on the observed samples to construct anchor graphs, ignoring the potentially useful information of missing instances. (2) Most methods attempt to learn a consensus anchor graph, failing to fully excavate the complementary information and high-order correlations across views. (3) They generally apply post-processing on learned anchor graph to seek latent embeddings, making them not globally-optimal. To address these issues, this paper proposes a novel fast IMVC approach with Adaptive Similarity Completion and Reconstruction (ASCR), which unifies anchor learning, anchor-sample similarity construction and completion, and latent multi-view embedding learning in a joint framework. Specifically, ASCR learns an anchor-sample similarity graph for each view, and the missing values are fulfilled to mitigate the adverse effects. To explore the consistent and complementary information across views, ASCR simultaneously seeks the view-specific anchor embeddings and sample embeddings in a latent subspace by similarity reconstruction, which not only preserves the semantic information into latent embeddings but also enhances the low-rank property of similarity graphs, achieving a reliable graph completion process. Furthermore, the high-order cross-view correlations are explored with tensor-based regularization. Extensive experimental results demonstrate the superiority and efficiency of ASCR compared with SOTA approaches.

Code — https://github.com/dengxu-nju/ASCR

Introduction

In real-world scenarios, diverse sources and feature collectors generate multi-view data, whose exponential growth over the past decade has driven significant research in multi-view learning, especially in multimedia and machine learning (Lu et al. 2019; Cai et al. 2024; Han et al. 2022; Fang et al. 2023). Among these tasks, multi-view clustering (MVC) (Sun et al. 2024; Zhang et al. 2024; Wang et al. 2021; Xu et al. 2024; Chen et al. 2022; Li et al. 2024; Gu, Li, and Feng 2024) stands out for its good ability to leverage complementary information across multiple views for clustering. However, in practical scenarios, some samples may be only partially available due to detector failure or data corruption.

To effectively group these incomplete data, there has been considerable attention towards incomplete multi-view clustering (IMVC). Numerious existing IMVC approaches have proven effective, such as those employing matrix factorization (Li, Jiang, and Zhou 2014), graph construction (Zhang et al. 2023a; Wen, Xu, and Liu 2020), subspace learning (Hu and Chen 2018; Liu et al. 2021), and deep learning (Lin et al. 2021; Xu et al. 2023; Chao, Jiang, and Chu 2024). For instance, (Liu et al. 2021) proposed jointly performing data imputation and self-representation learning for better IMVC results. (Wen, Xu, and Liu 2020) exploits the graph learning and spectral clustering techniques to learn the common representation for IMVC. Furthermore, the deep IMVC methods usually try to infer missing data and extract representation with a deep neural network (Wen et al. 2020b). Although remarkable success achieved by these IMVC approaches, their high storage and computational complexity pose challenges to their application on large-scale datasets.

Recently, some anchor-based IMVC methods have been proposed for fast clustering, which are widely applied in large-scale IMVC tasks (Liu et al. 2022; Wang et al. 2022; Yu et al. 2022; Wen et al. 2023; Yu et al. 2024). Anchor-based IMVC focuses on exploring the similarity relations between some anchors and samples to construct anchor graphs. As a representative, (Wang et al. 2022) proposed to learn a consensus anchor matrix and an anchor graph in a latent subspace. (Wen et al. 2023) learned view-specific anchors and anchor graphs, and aligned all graphs with a consensus one. In the latest work by (Yu et al. 2024), they utilized observed samples to generate view-shared anchors with multi-dimensions and multi-sizes for large-scale IMVC tasks. Although these anchor-based IMVC methods have achieved great success in large-scale IMVC applications, most of them still encounter some issues: (1) They typically construct anchor-sample similarity graphs directly from observed samples, neglecting the potentially valuable information of missing samples. (2) Most methods mainly focus on learning a consensus anchor graph, which might not fully exploit complementary information and high-order correlations across views. (3) They generally need post-processing

*Corresponding author.

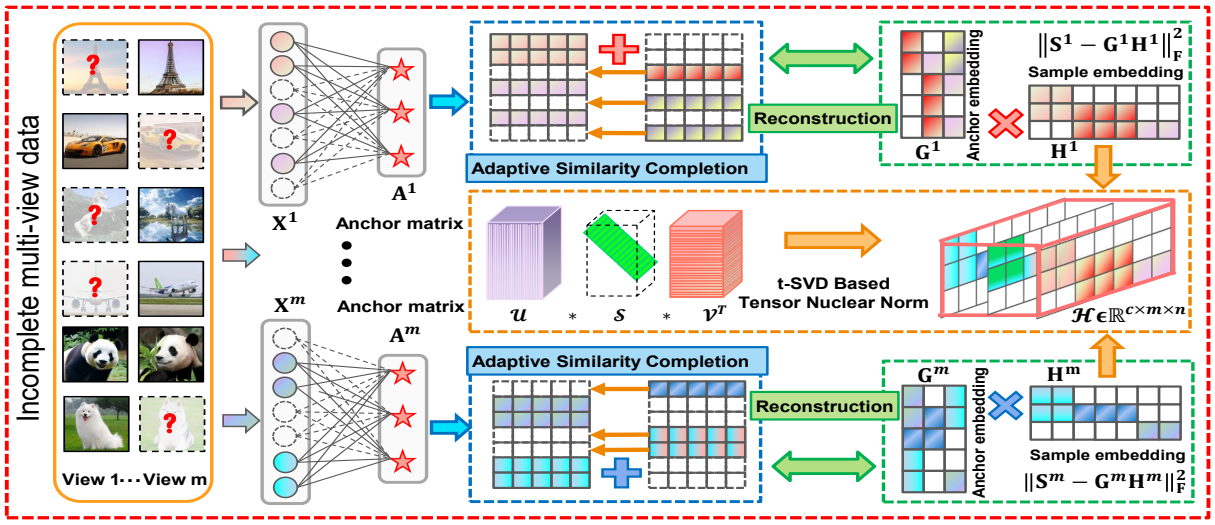

Figure 1: The overall framework of ASCR.

such as singular value decomposition (SVD) (Wen et al. 2023; Chen et al. 2023; Yu et al. 2024) for the learned anchor graph to seek latent embeddings for final k-means, which separates the graph learning and embedding learning into two steps and may lead to sub-optimal results.

To overcome these limitations, in this paper, we proposed a novel fast IMVC approach with Adaptive Similarity Completion and Reconstruction (ASCR), whose framework is shown in Figure 1. Specifically, ASCR learns an anchor-sample similarity graph for each view, and the missing values are effectively fulfilled to alleviate the negative impact of incomplete data. To explore the latent consistent and complementary information across views, ASCR seeks the latent view-specific anchor embedding and sample embedding by reconstructing the similarity graphs, which not only preserves the similarities into latent embeddings but also enhances the low-rank property of similarity graphs, ensuring a robust completion process and mutual reinforcement. Furthermore, ASCR incorporates tensorized regularization on sample embeddings to explore high-order cross-view correlations, enabling a deeper understanding of the underlying data structure. Finally, the multi-view sample embeddings are combined for clustering, leveraging the comprehensive information captured across views. The main contributions of this work are summarized as follows:

- We propose a novel IMVC method termed ASCR, which integrates anchor leaning, similarity graph construction and completion, and latent multi-view embedding learning into a unified framework.
- ASCR adaptively learns and completes anchor-sample similarity graphs, and simultaneously reconstructs the similarities in a latent subspace for discriminative embeddings learning. The high-order cross-view correlations are also explored with tensor-based regularization.
- Extensive experiments on several popular datasets demonstrate the effectiveness and efficiency of ASCR compared to various state-of-the-art IMVC methods.

Related Work

Notations and Preliminaries

We adopt the following notation conventions this paper: bold lowercase letters (e.g., $\mathbf{a}$) represent vectors, uppercase letters (e.g., $\mathbf{A}$) denote matrices, and calligraphic letters (e.g., $\mathcal{A}$) signify tensors. For a matrix $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$, its Frobenius norm and nuclear norm are defined as $\|\mathbf{A}\|_F = \sqrt{\sum_{ij} a_{ij}^2}$ and $\|\mathbf{A}\|_* = \sum_i \delta_i(\mathbf{A})$, respectively, where a_{ij} is the element of $\mathbf{A}$ at position (i, j), and $\delta_i(\mathbf{A})$ is the i-th singular value of $\mathbf{A}$. $\mathbf{I}_k$ is a k-dimensional identity matrix. For a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we denote the i-th frontal, lateral, and horizontal slice as $\mathcal{A}^{(:,:,i)}$, $\mathcal{A}^{(:,i,:)}$, and $\mathcal{A}^{(i,:,:)}$, respectively. Additionally, for convenience, we use $\mathcal{A}^{(i)}$ to represent $\mathcal{A}^{(:,:,i)}$. $\mathcal{A}_f$ denotes the fast Fourier transformation (FFT) of $\mathcal{A}$ along the third dimension, i.e., $\mathcal{A}_f = \text{fft}(\mathcal{A}, [], 3)$, and $\mathcal{A}$ can be recovered from $\mathcal{A}_f$ by the inverse FFT operation, i.e., $\mathcal{A} = \text{ifft}(\mathcal{A}_f, [], 3)$ (Lu et al. 2020).

Definition 1 (**t-SVD** (Kilmer et al. 2013))**.** For a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its t-SVD is defined as

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T,$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ are orthogonal tensors, $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is an f-diagonal tensor, and "$*$" denotes the t-product.

Definition 2 (**t-SVD based tensor nuclear norm** (Semerci et al. 2014))**.** Given a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, its t-SVD based tensor nuclear norm is defined as

$$\|\mathcal{A}\|_{\circledast} = \sum_{k=1}^{n_3} \|\mathbf{A}_f^{(k)}\|_* = \sum_{i=1}^{\min(n_1,n_2)} \sum_{k=1}^{n_3} \delta_i(\mathbf{A}_f^{(k)}),$$

Anchor-Based IMVC

Anchor graph has emerged as a powerful tool in exploring the data structure of large-scale data. Most anchor-based

IMVC methods focus on using some representative anchors as the basis to construct a consensus similarity matrix that captures the relationships between the anchors and samples. Let $\mathbf{X} = \{\mathbf{X}^v\}_{v=1}^m$ with $\mathbf{X}^v \in \mathbb{R}^{d_v \times n_v}$ denote an incomplete multi-view dataset, where m is the view number, d_v denotes the feature dimension and n_v signifies the number of observed samples in view v. The general framework of anchor-based IMVC can be described as

$$\begin{aligned} \min_{\mathbf{A}^v, \mathbf{Z}} \sum_{v=1}^{m} \|\mathbf{X}^v - \mathbf{A}^v \mathbf{Z} \mathbf{Q}^{v^T}\|_F^2 + \alpha \mathcal{R}(\mathbf{Z}) \\ s.t.\ \mathbf{A}^{v^T} \mathbf{A}^v = \mathbf{I}_k,\ \mathbf{Z} \geq 0,\ \mathbf{Z}^T \mathbf{1} = \mathbf{1}, \end{aligned} \tag{1}$$

where $\mathbf{A}^v \in \mathbb{R}^{d_v \times k}$ is the pre-defined or learnable anchor matrix of the v-th view, and k is the number of anchors (usually $k \ll n_v$). $\mathbf{Z} \in \mathbb{R}^{k \times n}$ is the consensus anchor graph. $\mathbf{Q}^v \in \mathbb{R}^{n_v \times n}$ functions as an index matrix to indicate the positions of existing samples, and it is constructed by deleting those rows of an $n \times n$ identity matrix that correspond to the missing samples. By this way, $\mathbf{Z}\mathbf{Q}^{v^T} \in \mathbb{R}^{k \times n_v}$ denotes the relationship between anchors and existing samples of the v-th view. $\mathcal{R}(\mathbf{Z})$ is a regularization term for some properties like sparse or low-rank. When the optimal anchor graph $\mathbf{Z}$ is obtained, SVD is then applied to seek the latent embedding for k-means clustering (Wang et al. 2022; Liu et al. 2022; Chen et al. 2023).

Based on Eq. (1), (Wang et al. 2022) jointly learned a consensus anchor matrix and anchor graph in a latent subspace. Following it, (Liu et al. 2022) considered view-specific anchor matrices $\mathbf{A}^v$ and a shared anchor graph $\mathbf{Z}$. They both use Frobenius norm regularization for a smooth graph. (Wen et al. 2023) learned the view-specific anchor graphs and fused them to a consensus one. However, these methods neglect the potential information of missing samples and also fail to explore the high-order cross-view correlations. Besides, all of them separate the graph learning and embedding learning into two individual steps, may making the embedding not globally-optimal. Although some methods like (Chen et al. 2023) employed the tensor nuclear norm for high-order correlations exploration, the rest two issues are still not addressed. In the upcoming section, we will propose a novel approach named ASCR, which is carefully designed to address these limitations effectively.

The Proposed Method

Model Formulation

From Eq. (1), we can observe that most anchor-based IMVC methods aim to develop a consensus anchor-sample similarity graph $\mathbf{Z}$. However, this approach often neglects the discrepancies between different views. To address this issue, we propose learning a distinct anchor-sample similarity graph $\mathbf{Z}^v$ for each view. This approach enables a more thorough exploration of the anchor-sample similarity structure within each specific view, thereby enhancing the mining of cross-view complementary information.

It should be noted that $\mathbf{Z}^v$ only reveals the similarities between anchors and existing samples, and the similarities w.r.t. missing samples are unavailable. Previous methods directly discard the information. However, as the missing rate increases, the loss of such valuable information becomes increasingly significant. To tackle this issue, we aim to learn a complete anchor-sample similarity graph for each view by adaptively fulfilling the missing values, which can be formulated as

$$\begin{aligned} \min_{\substack{\mathbf{A}^v, \mathbf{Z}^v \\ \mathbf{E}^v, \mathbf{S}^v}} \sum_{v=1}^{m} \|\mathbf{X}^v - \mathbf{A}^v \mathbf{Z}^v\|_F^2 + \|\mathbf{Z}^v \mathbf{Q}^v + \mathbf{E}^v \mathbf{P}^v - \mathbf{S}^v\|_F^2 \\ s.t.\ \mathbf{A}^{v^T} \mathbf{A}^v = \mathbf{I}_k,\ \mathbf{Z}^v \geq 0,\ \mathbf{E}^v \geq 0. \end{aligned} \tag{2}$$

Here, $\mathbf{Z}^v \in \mathbb{R}^{k \times n_v}$ reflects the relationship between anchors and existing samples of the v-th view. $\mathbf{E}^v \in \mathbb{R}^{k \times (n - n_v)}$ is the adaptive anchor-sample similarity completion matrix which reflects the similarity between anchors and missing samples. The nonnegative constraint on $\mathbf{Z}^v$ and $\mathbf{E}^v$ makes the learned similarity matrix and completion matrix physically meaningful. $\mathbf{P}^v \in \mathbb{R}^{(n - n_v) \times n}$ is also an index matrix to indicate the positions of existing samples, which is constructed by deleting those columns of an $n \times n$ identity matrix that corresponds to existing samples of the v-th view. In this way, we can utilize $\mathbf{E}^v$ to fill the missing values and obtain a complete anchor-sample similarity matrix $\mathbf{S}^v \in \mathbb{R}^{k \times n}$ for each view.

For a desirable completion, it is usually expected that the similarity matrix $\mathbf{S}^v$ can reveal the cluster structure of samples. Low-rank regularization is widely used in clustering models via nuclear norm for similarity matrices. Whereas, it usually needs post-processing for embedding learning, and also introduces additional terms and hyper-parameters. In our method, inspired by matrix factorization, we factorize the similarity matrix $\mathbf{S}^v$ into two latent factors, i.e., $\mathbf{S}^v = \mathbf{G}^v \mathbf{H}^v$, where $\mathbf{G}^v \in \mathbb{R}^{k \times c}$ and $\mathbf{H}^v \in \mathbb{R}^{c \times n}$ denote the latent anchor embedding and sample embedding, respectively, and c is the number of clusters ($c \leq k$). Then, Eq. (2) becomes

$$\begin{aligned} \min_{\substack{\mathbf{A}^v, \mathbf{Z}^v, \mathbf{E}^v \\ \mathbf{H}^v, \mathbf{G}^v}} \sum_{v=1}^{m} \|\mathbf{X}^v - \mathbf{A}^v \mathbf{Z}^v\|_F^2 + \|\mathbf{Z}^v \mathbf{Q}^v + \mathbf{E}^v \mathbf{P}^v - \mathbf{G}^v \mathbf{H}^v\|_F^2 \\ s.t.\ \mathbf{A}^{v^T} \mathbf{A}^v = \mathbf{I}_k,\ \mathbf{Z}^v \geq 0,\ \mathbf{E}^v \geq 0, \mathbf{G}^{v^T} \mathbf{G}^v = \mathbf{I}_c. \end{aligned} \tag{3}$$

The orthogonal constraint on $\mathbf{G}^v$ is to avoid the arbitrary scale. The anchor-sample similarity matrix factorization not only encourages the low-rank property of $\mathbf{S}^v$, but also incorporates the latent embedding learning into a unified model without introducing additional terms. And the similarity relations are reconstructed by anchor embedding and sample embedding in a latent subspace.

Eq. (3) learns a latent semantic sample embedding for each view individually, and the cross-view correlations are not fully captured. Instead of enforcing a shared representation for them, we adopt the tensor nuclear norm regularization to capture the cross-view consistency. The final objec-

tive function of our ASCR model is

$$\begin{aligned}\min_{\Omega} &\sum_{v=1}^{m} \|\mathbf{X}^v - \mathbf{A}^v\mathbf{Z}^v\|_F^2 \\ &+ \|\mathbf{Z}^v\mathbf{Q}^v + \mathbf{E}^v\mathbf{P}^v - \mathbf{G}^v\mathbf{H}^v\|_F^2 + \alpha\|\mathcal{H}\|_{\circledast} \quad (4)\\ s.t.\ &\mathbf{A}^{v^T}\mathbf{A}^v = \mathbf{I}_k,\ \mathbf{Z}^v \geq 0,\ \mathbf{E}^v \geq 0,\\ &\mathbf{G}^{v^T}\mathbf{G}^v = \mathbf{I}_c,\ \mathcal{H} = \Phi(\mathbf{H}^1, \mathbf{H}^2, ..., \mathbf{H}^m).\end{aligned}$$

where $\Omega = \{\mathbf{A}^v, \mathbf{Z}^v, \mathbf{E}^v, \mathbf{G}^v, \mathcal{H}\}$ is the target variables set. Φ is to convert all sample embeddings $\mathbf{H}^v$ into a three-order tensor $\mathcal{H} \in \mathbb{R}^{c\times m\times n}$. Our model integrates anchor learning, similarity matrix construction and completion, and latent embedding learning in a unified framework. The missing information is recovered and used to enrich the semantic representation. After solving the above model, we concatenate the multi-view latent embeddings by $\mathbf{H} = [\mathbf{H}^{1^T}, \mathbf{H}^{2^T}, ..., \mathbf{H}^{m^T}] \in \mathbb{R}^{n\times mc}$ for k-means clustering.

Optimization

To optimize the objective function with multiple variables, we use an alternate optimization strategy, updating one variable at a time while keeping the others fixed.

Updating $\mathbf{E}^v$: When other variables are fixed, the optimization for similarity completion matrix $\mathbf{E}^v$ is

$$\mathbf{E}^v_{t+1} = \arg\min_{\mathbf{E}^v\geq 0} \|\mathbf{E}^v\mathbf{P}^v - \mathbf{L}^v_t\|_F^2, \quad (5)$$

where $\mathbf{L}^v_t = \mathbf{G}^v_t\mathbf{H}^v_t - \mathbf{Z}^v_t\mathbf{Q}^v$. Since $\mathbf{P}^v$ is an index matrix, Eq. (5) can be solved by the following rule:

$$\mathbf{E}^v_{t+1} = \max(\hat{\mathbf{L}}^v_t, 0), \quad (6)$$

where $\hat{\mathbf{L}}^v_t \in \mathbb{R}^{k\times(n-n_v)}$ is a sub-matrix of $\mathbf{L}^v_t$ that is composed of the rows that correspond to missing samples.

Updating $\mathbf{A}^v$: When other variables are fixed, the optimization for view-specific anchor matrix $\mathbf{A}^v$ is

$$\mathbf{A}^v_{t+1} = \arg\min_{\mathbf{A}^v} \|\mathbf{X}^v - \mathbf{A}^v\mathbf{Z}^v\|_F^2 \quad s.t.\ \mathbf{A}^{v^T}\mathbf{A}^v = \mathbf{I}_k. \quad (7)$$

By transforming the Frobenius norm to the trace and eliminating terms irrelevant to $\mathbf{A}^v$, the above formula can be equivalently reformulated as:

$$\max_{\mathbf{A}^v} \mathrm{Tr}(\mathbf{A}^{v^T}\mathbf{M}^v_t) \quad s.t.\ \mathbf{A}^{v^T}\mathbf{A}^v = \mathbf{I}_k, \quad (8)$$

where $\mathbf{M}^v_t = \mathbf{X}^v\mathbf{Z}^{v^T}_t$. This subproblem can be efficiently solved using SVD. The closed-form solution is given by $\mathbf{A}^v_{t+1} = \mathbf{J}\mathbf{K}^T$, where $\mathbf{J}$ and $\mathbf{K}$ are the left and right singular matrices of $\mathbf{M}^v_t$.

Updating $\mathbf{Z}^v$: When other variables are fixed, the optimization for $\mathbf{Z}^v$ is

$$\mathbf{Z}^v_{t+1} = \arg\min_{\mathbf{Z}^v\geq 0} \|\mathbf{X}^v - \mathbf{A}^v\mathbf{Z}^v\|_F^2 + \|\mathbf{Z}^v\mathbf{Q}^v + \mathbf{E}^v\mathbf{P}^v - \mathbf{G}^v\mathbf{H}^v\|_F^2. \quad (9)$$

It can be further simplified as

$$\mathbf{Z}^v_{t+1} = \arg\min_{\mathbf{Z}^v\geq 0} \|\mathbf{O}^v_t - \mathbf{Z}^v\|_F^2. \quad (10)$$

where $\mathbf{O}^v_t = ((\mathbf{G}^v_t\mathbf{H}^v_t - \mathbf{E}^v_{t+1}\mathbf{P}^v)\mathbf{Q}^{v^T} + \mathbf{A}^{v^T}_{t+1}\mathbf{X}^v)/2$. The optimal solution of $\mathbf{Z}^v_t$ is given by

$$\mathbf{Z}^v_{t+1} = \max(\mathbf{O}^v_t, 0). \quad (11)$$

Updating $\mathbf{G}^v$: When other variables are fixed, the optimization for anchor embedding $\mathbf{G}^v$ is

$$\begin{aligned}\mathbf{G}^v_{t+1} = \arg\min_{\mathbf{G}^v} &\|\mathbf{Z}^v_{t+1}\mathbf{Q}^v + \mathbf{E}^v_{t+1}\mathbf{P}^v - \mathbf{G}^v\mathbf{H}^v_t\|_F^2 \\ &s.t.\ \mathbf{G}^{v^T}\mathbf{G}^v = \mathbf{I}_c.\end{aligned} \quad (12)$$

Similar to $\mathbf{A}^v$, Eq. (12) is converted to

$$\max_{\mathbf{G}^v} \mathrm{Tr}(\mathbf{G}^{v^T}\mathbf{N}^v_t) \quad s.t.\ \mathbf{G}^{v^T}\mathbf{G}^v = \mathbf{I}_c, \quad (13)$$

where $\mathbf{N}^v_t = (\mathbf{Z}^v_{t+1}\mathbf{Q}^v + \mathbf{E}^v_{t+1}\mathbf{P}^v)\mathbf{H}^{v^T}_t$. As in solving Eq. (8), the optimal solution for $\mathbf{G}^v$ is $\mathbf{G}^v_{t+1} = \mathbf{R}\mathbf{T}^T$, where $\mathbf{R}$ and $\mathbf{T}$ are the left and right singular matrices of $\mathbf{N}^v_t$.

Updating $\mathcal{H}$: When other variables are fixed, the optimization for sample embedding tensor $\mathcal{H}$ is

$$\mathcal{H}_{t+1} = \arg\min_{\mathcal{H}} \alpha\|\mathcal{H}\|_{\circledast} + \|\mathcal{H} - \mathcal{F}_{t+1}\|_F^2, \quad (14)$$

where $\mathcal{F}_{t+1} = \Phi(\mathbf{F}^1_{t+1}, ..., \mathbf{F}^m_{t+1})$ is constructed by $\mathbf{F}^v_t = \mathbf{G}^{v^T}_{t+1}(\mathbf{Z}^v_{t+1}\mathbf{Q}^v + \mathbf{E}^v_{t+1}\mathbf{P}^v)$. Eq. (14) is a typical low-rank tensor norm minimization problem that has a closed-form solution and can be solved by Theorem 1 (Xie et al. 2018).

Theorem 1. Given two tensor $\mathcal{B} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathcal{D} \in \mathbb{R}^{n_1\times n_2\times n_3}$ with a constant ρ, the globally optimal solution of the problem

$$\min_{\mathcal{B}} \rho\|\mathcal{B}\|_{\circledast} + \frac{1}{2}\|\mathcal{B} - \mathcal{D}\|_F^2$$

can be obtained by the tensor tubal-shrinkage operator

$$\mathcal{B} = \mathcal{C}_{n_3\rho}(\mathcal{D}) = \mathcal{U} * \mathcal{C}_{n_3\rho}(\mathcal{S}) * \mathcal{V}^T,$$

where $\mathcal{D} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ and $\mathcal{C}_{n_3\rho}(\mathcal{D}) = \mathcal{S} * \mathcal{K}$. $\mathcal{K} \in \mathbb{R}^{n_1\times n_2\times n_3}$ is a f -diagonal tensor and its diagonal element in the Fourier domain is $\mathcal{K}_f(i,i,j) = (1 - \frac{n_3\rho}{\mathcal{S}^{(j)}_f(i,i)})_+$.

Algorithm 1 summarizes the optimization process.

Computational Complexity Analysis

Time Complexity When updating $\mathbf{E}^v$, the primary time cost is from matrix multiplication, which takes $O(ckn + d_vkn)$. Updating $\mathbf{A}^v$ and $\mathbf{G}^v$ involves the SVD operation and matrix multiplication, requiring $O(k^2d_v + d_vkn_v)$ and $O(kc^2 + ckn)$, respectively. The $\mathbf{Z}^v$ step costs $O(d_vkn_v + ckn)$ for matrix multiplication. Updating $\mathcal{H}$ involves matrix multiplication, FFT, inverse FFT, and SVD operations, with matrix multiplication taking $\mathcal{O}(kcn)$. For a $c \times m \times n$ tensor, FFT and inverse FFT operations take $\mathcal{O}(cmn\log(n))$, and the SVD operation needs $\mathcal{O}(cnm^2)$. Since (k, c, m) $\ll$ (n, d_v), and $n_v < n$, the overall time complexity of Algorithm 1 is $O(\tau(cmn\log(n) + kdn))$, where $d = \sum_{v=1}^m d_v$, and τ is the number of iterations.

Space Complexity The major memory costs of our algorithm are for variables $\mathbf{A}^v \in \mathbb{R}^{d_v\times k}$, $\mathbf{Z}^v\mathbf{Q}^v \in \mathbb{R}^{k\times n}$, $\mathbf{P}^v\mathbf{E}^v \in \mathbb{R}^{k\times n}$, $\mathbf{G}^v \in \mathbb{R}^{k\times c}$, and $\mathbf{H}^v \in \mathbb{R}^{c\times n}$. Thus, the space complexity of Algorithm 1 is $O(n)$.

Algorithm 1: ASCR algorithm

Input: Incomplete multi-view data $\{\mathbf{X}^v\}_{v=1}^m$, parameter α, the number of anchor k and cluster c.
Output: Perform k-means on $\mathbf{H}$.

Initialize $\mathbf{E}^v = \mathbf{0}$, $\mathbf{A}^v = \mathbf{0}$, $\mathbf{Z}^v = \mathbf{0}$, $\mathbf{G}^v = \mathbf{0}$, $\mathcal{H} = 0$, $\epsilon = 1e-5$.
Construct the index matrices $\{\mathbf{Q}^v\}_{v=1}^m$ and $\{\mathbf{P}^v\}_{v=1}^m$.
while not converged **do**
  Update $\mathbf{E}^v$ by Eq. (6);
  Update $\mathbf{A}^v$ by solving (8);
  Update $\mathbf{Z}^v$ by Eq. (11);
  Update $\mathbf{G}^v$ by solving (13);
  Update $\mathcal{H}$ by solving. (14);
  Check the convergence conditions: $\sum_{v=1}^m \|\mathbf{H}_t^v - \mathbf{H}_{t-1}^v\|_F^2 / \sum_{v=1}^m \|\mathbf{H}_{t-1}^v\|_F^2 \leq \epsilon$;
  $t \longleftarrow t+1$;
end while
Obtain $\mathbf{H}$ by concatenating $[\mathbf{H}^{1^T}, \mathbf{H}^{2^T}, ..., \mathbf{H}^{m^T}]$.

Experiment

Experimental Setup

Datasets Eight popular datasets from diverse applications were used to validate our ASCR, consisting of text datasets **NGs**[1], digit dataset **HW**[2], scene datasets **Scene15** (Xie et al. 2018) and **SUNRGBD**[3], face dataset **NHface** (Cao et al. 2015), video dataset **CCV** (Wang et al. 2021), as well as object datasets **Caltech101**[4] and **NUSWIDE** (Chua et al. 2009). More details are shown in Table 1.

Incomplete Data Construction To construct incomplete multi-view data, following (Wen et al. 2020a), we set varying missing rates $p \in [0.1 : 0.2 : 0.9]$ to comprehensively investigate the robustness of our method against missing data. When the missing rate $p = 0.1$, we randomly select 90% samples as complete data and randomly drop some views of the remaining 10% samples. At least one view is preserved for incomplete instances.

Evaluation Metrics Five commonly used metrics are employed to evaluate clustering performance: accuracy (ACC), normalized mutual information (NMI), purity (PUR), adjusted Rand index (ARI), and F-score. Higher values for each metric indicate better clustering performance.

Baselines We compare ASCR's clustering performance with eight SOTA IMVC methods: **IMVC-CBG** (Wang et al. 2022), **FIMVC-VIA** (Liu et al. 2022), **SIMVC-SA** (Wen et al. 2023), **DVSAI** (Yu et al. 2024), **LSIMVC** (Liu et al. 2023), **SEC-IMVC** (Zhang et al. 2023b), **TDASC** (Chen et al. 2023), **sFSR-IMVC** (Long et al. 2023).

Parameter Settings The hyper-parameters in the baselines are tuned according to the corresponding papers. For ASCR, we fix the anchor number $k = c$ and tune α within the range $2^{\{2:7\}}$. All methods are run 10 times to obtain clustering results and standard deviations for fair comparison. The experiments are conducted using MATLAB R2021a on a PC with an i5-12400 CPU and 16GB RAM.

[1]https://lig-membres.imag.fr/grimal/data.html
[2]https://archive.ics.uci.edu/ml/datasets/Multiple+Features
[3]https://rgbd.cs.princeton.edu/
[4]https://data.caltech.edu/records/mzrjq-6wc02

Dataset	Sample	Cluster	View	View dimension
NGs	500	5	3	2000, 2000, 2000
HW	2000	10	6	240, 76, 216, 47, 64, 6
Scene15	4485	15	3	1800, 1180, 1240
NHface	4660	5	3	6750, 2000, 3304
CCV	6773	20	3	20, 20, 20
Caltech101	9144	102	5	48, 40, 254, 512, 928
SUNRGBD	10335	45	2	4096, 4096
NUSWIDE	30000	31	5	64, 225, 144, 73, 128

Table 1: Details of the used datasets.

Experimental Results

Figure 2 shows the ACC value of all approaches on all datasets with varying missing rates (5 cases). Additionally, Table 2 presents the average clustering results of the 5 cases on all datasets. The best and second-best results are highlighted in **red** and **blue**, respectively. "OM" denotes unavailable results due to out-of-memory errors. Based on the results, we can draw the following conclusions:

(1) Our method generally significantly outperforms the baselines. For instance, on the CCV datasets, it achieves an average improvement of 34.21%, 36.90%, 32.46%, 27.96%, and 35.64% over the second-best method (i.e., sFSR-IMVC) in terms of five metrics, respectively. Notably, as shown in Figure 2, our method exhibits remarkable stability compared to others as the missing rate increases, demonstrating the robustness of our ASCR against missing data.

(2) Compared to both anchor-based (e.g., IMVC-CBG, FIMVC-VIA, SIMVC-SA, DVSAI, and TDASC) and graph-based (e.g., LSIMVC) methods, which rely on observed samples for anchor graph construction or consensus representation learning, our ASCR consistently outperforms them. As the missing rate increases, the superiority of ASCR becomes more pronounced, demonstrating the advantages of our adaptive similarity completion strategy.

(3) While SEC-IMVC and sFSR-IMVC tend to yield better results than the other five baselines in most datasets due to their spectral embedding completion or feature space recovery approach, our ASCR consistently achieves superior performance across nearly all datasets. This is attributed to our approach of reconstructing anchor-sample similarity graphs using sample embedding and anchor embedding. This process not only preserves similarities in latent embeddings but also enhances the low-rank property of similarity graphs, leading to a more robust completion process and obtaining more discriminative embeddings.

Parameter Analysis

To assess the influence of α and the anchor number k in ASCR, Figure 3 shows ACC values on NGs and NHface for different α and k. The results show that ASCR is not sensitive to k, consistently delivering excellent and stable clustering performance. Therefore, we set $k = c$ for all datasets to

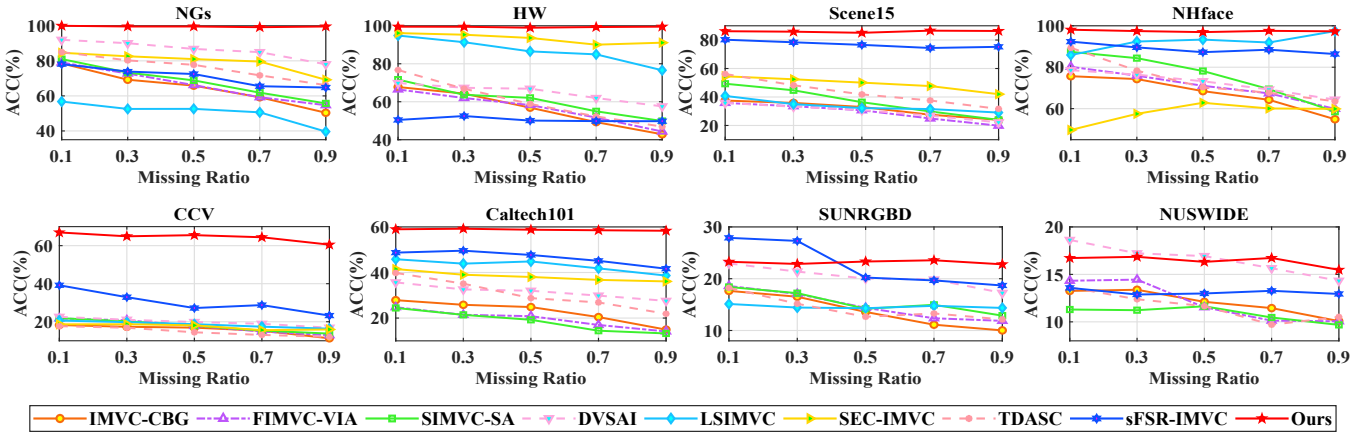

Figure 2: Clustering performances w.r.t. ACC on 8 datasets with varying missing rates.

	Metric(%)	IMVC-CBG	FIMVC-VIA	SIMVC-SA	DVSAI	LSIMVC	SEC-IMVC	TDASC	sFSR-IMVC	Ours
NGs	ACC	64.65±0.42	66.38±0.46	68.14±0.22	**86.37±0.16**	50.44±0.21	79.40±0.00	76.17±0.29	70.95±1.33	**99.60±0.00**
	NMI	53.73±0.31	57.38±0.27	58.08±0.28	**71.01±0.16**	33.86±0.22	56.62±0.00	62.78±0.20	66.57±1.48	**98.68±0.00**
	PUR	64.65±0.47	66.35±0.28	67.30±0.07	**86.37±0.11**	50.46±0.21	79.20±0.00	76.17±0.13	71.64±1.31	**99.60±0.00**
	ARI	40.63±0.26	47.29±0.26	50.82±0.19	**70.45±0.04**	19.49±0.18	56.83±0.00	54.28±0.24	60.99±1.71	**99.00±0.00**
	F-score	46.20±0.27	47.54±0.15	52.64±0.55	**77.52±0.06**	36.56±0.15	65.68±0.00	62.00±0.23	69.41±1.21	**99.20±0.00**
HW	ACC	56.20±0.60	56.53±0.67	60.43±0.39	64.71±0.32	86.85±1.56	**93.24±0.00**	60.31±0.45	50.50±3.20	**99.38±0.03**
	NMI	43.74±0.35	44.87±0.59	50.62±0.42	53.78±0.37	80.85±1.15	**86.57±0.00**	57.94±0.42	59.25±4.97	**98.47±0.07**
	PUR	56.98±0.62	57.91±0.67	63.27±0.38	64.55±0.40	86.65±1.32	**93.24±0.00**	67.05±0.31	50.72±5.27	**99.38±0.03**
	ARI	36.76±0.51	35.97±0.42	41.52±0.29	39.03±0.34	83.92±0.81	**85.84±0.00**	58.06±0.54	43.27±4.33	**98.61±0.06**
	F-score	43.92±0.37	44.72±0.37	50.19±0.31	47.53±0.52	80.54±0.91	**87.25±0.00**	54.09±0.32	51.56±4.01	**98.75±0.06**
Scene15	ACC	31.57±1.43	28.96±0.67	36.83±1.96	29.79±1.79	33.71±1.27	49.35±0.00	43.09±0.31	**76.98±5.12**	**86.01±3.78**
	NMI	26.74±0.38	29.21±0.49	35.47±0.62	31.73±0.39	36.63±0.58	43.46±0.00	37.34±0.19	**84.27±2.46**	**90.90±1.35**
	PUR	30.27±1.16	28.90±0.58	40.13±0.81	35.13±0.69	37.40±0.95	51.10±0.00	40.98±0.37	**82.32±3.76**	**89.81±2.35**
	ARI	20.40±0.92	19.58±0.24	20.86±0.54	22.36±1.21	16.06±0.90	30.02±0.00	27.79±0.19	**72.08±5.40**	**83.12±3.93**
	F-score	27.59±0.35	27.54±0.20	32.59±0.49	30.32±0.54	23.35±0.74	35.03±0.00	29.85±0.20	**74.17±4.95**	**84.31±3.57**
NHface	ACC	67.51±0.05	70.84±1.04	75.58±0.14	72.36±0.09	**92.19±0.46**	58.00±0.00	73.86±0.00	88.80±0.07	**97.48±0.01**
	NMI	59.12±0.17	57.52±0.60	64.45±0.09	61.11±0.20	**84.70±0.73**	40.10±0.00	64.07±0.01	82.29±0.39	**94.19±0.01**
	PUR	68.91±0.04	68.59±0.25	75.20±0.16	72.76±0.09	**92.28±0.46**	63.16±0.00	72.02±0.00	84.88±0.07	**97.48±0.01**
	ARI	65.79±0.11	59.49±1.37	63.67±0.11	58.02±0.23	**85.80±0.85**	34.34±0.00	60.73±0.01	83.56±0.64	**95.43±0.01**
	F-score	60.03±0.09	57.28±0.81	60.44±0.08	59.69±0.27	**88.93±0.49**	49.15±0.00	69.14±0.00	86.85±0.34	**96.33±0.01**
CCV	ACC	15.81±0.36	17.61±0.25	17.43±0.32	19.85±0.29	18.57±0.34	17.39±0.00	14.81±0.03	**30.25±1.45**	**64.46±2.72**
	NMI	11.43±0.18	12.21±0.20	12.31±0.23	15.32±0.27	16.07±0.19	11.94±0.00	9.04±0.09	**34.23±0.71**	**71.13±0.77**
	PUR	21.15±0.49	20.21±0.27	20.14±0.32	23.25±0.45	22.37±0.27	18.86±0.00	17.11±0.05	**35.78±1.12**	**68.24±1.88**
	ARI	14.11±0.36	14.30±0.11	14.36±0.11	16.08±0.28	15.69±0.11	14.77±0.00	12.56±0.04	**26.71±0.85**	**54.67±2.15**
	F-score	9.96±0.21	10.54±0.20	10.61±0.11	12.37±0.32	11.24±0.12	10.12±0.00	9.00±0.03	**21.50±0.75**	**57.14±2.00**
Caltech101	ACC	22.77±0.97	18.02±0.54	17.56±0.42	31.52±0.33	43.00±2.86	38.28±0.00	24.48±0.40	**46.55±1.40**	**58.76±2.15**
	NMI	40.28±0.43	32.95±0.42	32.31±0.44	51.48±0.34	64.75±2.47	55.52±0.00	45.39±0.57	**79.71±0.34**	**86.66±0.68**
	PUR	42.15±0.52	30.55±0.43	29.92±0.45	48.43±0.32	59.85±2.43	55.65±0.00	40.65±0.24	**72.33±0.73**	**81.69±1.45**
	ARI	10.90±0.58	4.90±0.48	5.98±0.50	19.62±0.39	23.19±0.40	23.26±0.00	16.90±0.50	**36.45±1.43**	**45.22±2.86**
	F-score	14.95±1.51	8.81±0.31	9.66±0.40	20.96±0.26	20.72±1.02	24.48±0.00	18.44±0.46	**37.50±1.42**	**46.16±2.64**
SUNRGBD	ACC	13.78±0.35	14.85±0.36	15.53±0.40	20.29±0.44	14.59±0.35	OM	14.29±0.40	**22.75±1.04**	**23.16±0.86**
	NMI	19.27±0.42	18.38±0.27	19.76±0.31	21.48±0.31	20.88±0.17	OM	13.62±0.14	**27.61±1.05**	**39.05±0.36**
	PUR	29.19±0.35	30.01±0.42	31.85±0.52	32.63±0.30	32.24±0.33	OM	23.45±0.16	**36.58±0.89**	**47.73±0.67**
	ARI	5.16±0.22	5.58±0.11	5.62±0.14	5.82±0.20	6.04±0.14	OM	2.59±0.07	**12.14±1.08**	**14.24±0.65**
	F-score	8.87±0.32	10.30±0.11	9.85±0.14	10.41±0.17	9.48±0.15	OM	10.39±0.17	**16.87±1.01**	**17.36±0.61**
NUSWIDE	ACC	12.06±0.33	12.10±0.20	10.86±0.13	**16.55±0.21**	OM	OM	11.58±0.12	13.13±0.65	**16.42±0.29**
	NMI	9.89±0.10	8.16±0.10	10.24±0.18	**11.67±0.24**	OM	OM	8.59±0.11	9.54±0.40	**18.79±0.11**
	PUR	21.24±0.37	19.87±0.15	15.81±0.18	**21.59±0.13**	OM	OM	20.21±0.13	19.41±0.58	**28.22±0.25**
	ARI	2.71±0.24	2.84±0.12	2.42±0.09	**5.82±0.20**	OM	OM	2.40±0.03	2.66±0.20	**6.70±0.12**
	F-score	7.94±0.28	8.91±0.17	**9.62±0.16**	9.30±0.28	OM	OM	7.87±0.04	7.77±0.25	**10.72±0.13**

Table 2: Clustering results (mean±std) of different methods on 8 datasets. OM indicates *out of memory*.

balance time consumption and clustering performance. The parameter α controls the weight of tensor low-rank regularization, and satisfactory results are obtained with α values from $2^{\{2:5\}}$ for NGs and $2^{\{3:6\}}$ for NHface.

Convergence and Time Comparison

For Algorithm 1, we set the convergence condition as $Loss(\mathbf{H}) = \sum_{v=1}^{m} \|\mathbf{H}_t^v - \mathbf{H}_{t-1}^v\|_F^2 / \sum_{v=1}^{m} \|\mathbf{H}_{t-1}^v\|_F^2 \leq 1e-5$. Figure 4 displays the change of $Loss(\mathbf{H})$ and ACC

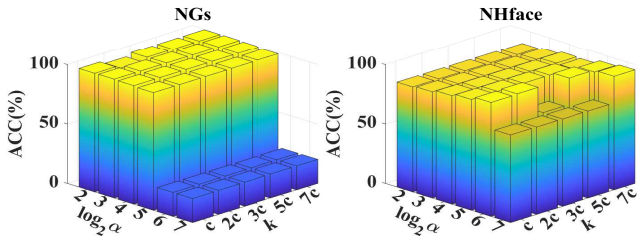

Figure 3: Clustering performance of ASCR under different parameter settings on NGs and NHface datasets ($p = 0.1$).

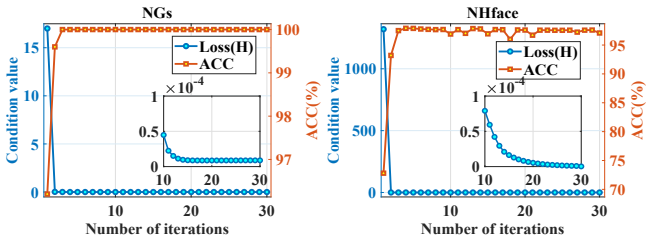

Figure 4: Convergence curves and the change of ACC value on NGs and NHface datasets ($p = 0.1$).

Method	NGs	HW	Scene15	NHface	CCV	Caltech	SUN	NUS
IMVC-CBG	0.68	1.09	5.14	8.85	5.54	27.96	25.62	74.56
FIMVC-VIA	0.31	0.51	2.29	3.76	1.92	5.87	13.46	18.29
SIMVC-SA	0.39	0.94	3.28	8.65	8.61	36.41	15.68	82.24
DVSAI	3.26	0.85	3.72	6.83	12.25	45.29	69.25	164.82
LSIMVC	5.12	3.43	2.85	5.89	6.97	19.52	17.71	OM
SEC-IMVC	0.81	29.27	170.89	131.96	556.02	2135.67	OM	OM
TDASC	1.12	4.54	8.05	11.53	9.26	62.03	42.98	128.78
sFSR-IMVC	6.08	46.11	52.57	63.19	5.77	47.05	276.8	354.22
Ours	0.1	0.32	0.6	1.99	1.45	15.53	9.27	20.41

Table 3: Running time (second) of all methods on 8 datasets.

values w.r.t. iterations on NGs and NHface datasets. It is noticeable that $Loss(\mathbf{H})$ decreases rapidly and converges to a stable value within 30 iterations, which satisfies our convergence condition. Meanwhile, the ACC value also increases and remains stable quickly. The outcomes demonstrate the favorable convergence property of our algorithm.

Table 3 shows the running time of all methods across benchmark datasets. Our ASCR demonstrates competitive computational efficiency compared to most baselines. SEC-IMVC has $O(n_v^3)$ complexity to learn spectral embeddings, and it is not efficient in practice. Although some anchor-based methods have $O(n)$ complexity, they compute the anchor-sample similarities for each sample one-by-one. However, all variables of ASCR are updated directly with closed-form solutions. Thus, our method generally spends less time than those methods. This efficiency highlights the potential of ASCR for large-scale incomplete tasks in practical applications.

Ablation Study

To assess the effectiveness of the anchor-sample similarity graph completion and reconstruction strategy, as well as the tensor-based regularization, we derive three variants based on the original ASCR model: ASCR-C, ASCR-R, and ASCR-T. ASCR-C drops the similarity completion matrix

Datasets	Method	p=0.1	p=0.3	p=0.5	p=0.7	p=0.9
NGs	ASCR-C	90.33	86.54	82.71	78.62	75.92
	ASCR-R	98.62	97.14	97.14	94.62	96.62
	ASCR-T	87.56	78.91	71.29	64.88	57.25
	ASCR	**100**	**99.60**	**99.60**	**99.20**	**99.60**
HW	ASCR-C	87.51	85.42	81.97	78.04	72.59
	ASCR-R	98.30	97.45	97.31	96.90	97.03
	ASCR-T	82.16	75.27	67.21	60.03	51.97
	ASCR	**99.53**	**99.49**	**98.99**	**99.34**	**99.54**
Scene15	ASCR-C	81.75	77.96	74.13	70.29	67.84
	ASCR-R	85.90	84.56	81.66	83.18	81.08
	ASCR-T	67.59	59.99	51.25	48.56	42.95
	ASCR	**86.18**	**85.86**	**85.08**	**86.58**	**86.36**
NHface	ASCR-C	88.25	83.27	80.52	77.21	73.19
	ASCR-R	94.66	94.96	93.88	94.61	93.56
	ASCR-T	86.55	79.24	73.57	64.85	53.92
	ASCR	**98.10**	**97.45**	**96.95**	**97.56**	**97.33**

Table 4: Ablation study results w.r.t. ACC values.

$\mathbf{E}^v$ and just uses the observed samples to construct anchor graphs. ASCR-R replace $\mathbf{G}^v\mathbf{H}^v$ with $\mathbf{S}^v$ and perform SVD on $\frac{1}{m}\sum_{v=1}^{m}\mathbf{S}^v$ to get latent embedding $\mathbf{U}$ and then apply k-means like TDASC. ASCR-T omits tensor nuclear-norm regularization and utilizes traditional nuclear norm on $\mathbf{H}^v$.

Table 4 shows the clustering performance of the three variants and our ASCR with different missing ratios. We can observe that the performance of ASCR-C is not desirable. Especially when the missing rate increases, the ACC value decreases sharply. This proves the effectiveness of our anchor-sample similarity completion. ASCR-R can achieve very competitive results. However, our ASCR still outperforms it with large missing rates. It is attributed to our reconstruction strategy, which seeks the latent embedding for a low-rank graph completion process and then achieves more optimal clustering results. In addition, the poor performance of ASCR-T also demonstrates the effectiveness of tensor regularization on $\mathbf{H}^v$, which helps to fully explore the high-order cross-view correlations, and improves the semantic consistency discovery.

Conclusion

In this paper, we propose a novel fast IMVC approach with Adaptive Similarity Completion and Reconstruction (ASCR), which unifies anchor learning, anchor-sample similarity completion, and latent multi-view embedding learning into a joint framework. ASCR learns an anchor-sample similarity graph for each view and fulfills the missing values to mitigate the adverse effects. To explore consistent and complementary information across views, ASCR simultaneously seeks latent view-specific anchor embedding and sample embedding by reconstructing similarity relations. It not only preserves the similarities into latent embeddings but also enhances the low-rank property of similarity graphs, achieving a robust graph completion process and mutual reinforcement. Furthermore, the high-order cross-view semantic consistency is captured via low-rank tensor regularization. Extensive experimental results on several datasets demonstrate the superiority and efficiency of our proposed ASCR compared with state-of-the-art approaches.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants 62176116, 62276136, 62073160.

References

Cai, J.; Zhang, Y.; Fan, J.; Du, Y.; and Guo, W. 2024. Dual Contrastive Graph-Level Clustering with Multiple Cluster Perspectives Alignment. In *IJCAI*, 3770–3779.

Cao, X.; Zhang, C.; Zhou, C.; Fu, H.; and Foroosh, H. 2015. Constrained multi-view video face clustering. *IEEE TIP*, 24(11): 4381–4393.

Chao, G.; Jiang, Y.; and Chu, D. 2024. Incomplete Contrastive Multi-View Clustering with High-Confidence Guiding. In *AAAI*, volume 38, 11221–11229.

Chen, M.; Wang, C.; Huang, D.; Lai, J.; and Yu, P. S. 2022. Efficient orthogonal multi-view subspace clustering. In *ACM SIGKDD*, 127–135.

Chen, Y.; Zhao, X.; Zhang, Z.; Liu, Y.; Su, J.; and Zhou, Y. 2023. Tensor Learning Meets Dynamic Anchor Learning: From Complete to Incomplete Multiview Clustering. *IEEE TNNLS*.

Chua, T.-S.; Tang, J.; Hong, R.; Li, H.; Luo, Z.; and Zheng, Y. 2009. Nus-wide: a real-world web image database from national university of singapore. In *ACM CIVR*, 1–9.

Fang, U.; Li, M.; Li, J.; Gao, L.; Jia, T.; and Zhang, Y. 2023. A comprehensive survey on multi-view clustering. *IEEE TKDE*, 35(12): 12350–12368.

Gu, Z.; Li, Z.; and Feng, S. 2024. EDISON: Enhanced Dictionary-Induced Tensorized Incomplete Multi-View Clustering with Gaussian Error Rank Minimization. In *ICML*.

Han, Z.; Zhang, C.; Fu, H.; and Zhou, J. T. 2022. Trusted multi-view classification with dynamic evidential fusion. *IEEE TPAMI*, 45(2): 2551–2566.

Hu, M.; and Chen, S. 2018. Doubly aligned incomplete multi-view clustering. In *IJCAI*, 2262–2268.

Kilmer, M. E.; Braman, K.; Hao, N.; and Hoover, R. C. 2013. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM J. Matrix Anal. Appl.*, 34(1): 148–172.

Li, S.-Y.; Jiang, Y.; and Zhou, Z.-H. 2014. Partial multi-view clustering. In *AAAI*, volume 28.

Li, X.; Pan, Y.; Sun, Y.; Sun, Q.; Tsang, I. W.; and Ren, Z. 2024. Fast Unpaired Multi-view Clustering. In *IJCAI*, 4488–4496.

Lin, Y.; Gou, Y.; Liu, Z.; Li, B.; Lv, J.; and Peng, X. 2021. Completer: Incomplete multi-view clustering via contrastive prediction. In *CVPR*, 11174–11183.

Liu, C.; Wu, Z.; Wen, J.; Xu, Y.; and Huang, C. 2023. Localized sparse incomplete multi-view clustering. *IEEE TMM*, 25: 5539–5551.

Liu, J.; Liu, X.; Zhang, Y.; Zhang, P.; Tu, W.; Wang, S.; Zhou, S.; Liang, W.; Wang, S.; and Yang, Y. 2021. Self-representation subspace clustering for incomplete multi-view data. In *ACM MM*, 2726–2734.

Liu, S.; Liu, X.; Wang, S.; Niu, X.; and Zhu, E. 2022. Fast incomplete multi-view clustering with view-independent anchors. *IEEE TNNLS*.

Long, Z.; Zhu, C.; Comon, P.; Ren, Y.; and Liu, Y. 2023. Feature Space Recovery for Efficient Incomplete Multi-View Clustering. *IEEE TKDE*.

Lu, C.; Feng, J.; Chen, Y.; Liu, W.; Lin, Z.; and Yan, S. 2020. Tensor Robust Principal Component Analysis with a New Tensor Nuclear Norm. *IEEE TPAMI*, 42(4): 925–938.

Lu, X.; Zhu, L.; Li, J.; Zhang, H.; and Shen, H. T. 2019. Efficient supervised discrete multi-view hashing for large-scale multimedia search. *IEEE TMM*, 22(8): 2048–2060.

Semerci, O.; Hao, N.; Kilmer, M. E.; and Miller, E. L. 2014. Tensor-based formulation and nuclear norm regularization for multienergy computed tomography. *IEEE TIP*, 23(4): 1678–1693.

Sun, Y.; Qin, Y.; Li, Y.; Peng, D.; Peng, X.; and Hu, P. 2024. Robust Multi-View Clustering With Noisy Correspondence. *IEEE TKDE*, 36(12): 9150–9162.

Wang, S.; Liu, X.; Liu, L.; Tu, W.; Zhu, X.; Liu, J.; Zhou, S.; and Zhu, E. 2022. Highly-efficient incomplete large-scale multi-view clustering with consensus bipartite graph. In *CVPR*, 9776–9785.

Wang, S.; Liu, X.; Zhu, X.; Zhang, P.; Zhang, Y.; Gao, F.; and Zhu, E. 2021. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE TIP*, 31: 556–568.

Wen, J.; Xu, Y.; and Liu, H. 2020. Incomplete multi-view spectral clustering with adaptive graph learning. *IEEE TCYB*, 50(4): 1418–1429.

Wen, J.; Yan, K.; Zhang, Z.; Xu, Y.; Wang, J.; Fei, L.; and Zhang, B. 2020a. Adaptive graph completion based incomplete multi-view clustering. *IEEE TMM*, 23: 2493–2504.

Wen, J.; Zhang, Z.; Zhang, Z.; Wu, Z.; Fei, L.; Xu, Y.; and Zhang, B. 2020b. Dimc-net: Deep incomplete multi-view clustering network. In *ACM MM*, 3753–3761.

Wen, Y.; Wang, S.; Liang, K.; Liang, W.; Wan, X.; Liu, X.; Liu, S.; Liu, J.; and Zhu, E. 2023. Scalable Incomplete Multi-View Clustering with Structure Alignment. In *ACM MM*, 3031–3040.

Xie, Y.; Tao, D.; Zhang, W.; Liu, Y.; Zhang, L.; and Qu, Y. 2018. On unifying multi-view self-representations for clustering by tensor multi-rank minimization. *IJCV*, 126: 1157–1179.

Xu, D.; Zhang, C.; Li, Z.; Chen, C.; and Li, H. 2024. Fast Disentangled Slim Tensor Learning for Multi-view Clustering. *arXiv preprint arXiv:2411.07685*.

Xu, J.; Li, C.; Peng, L.; Ren, Y.; Shi, X.; Shen, H. T.; and Zhu, X. 2023. Adaptive Feature Projection With Distribution Alignment for Deep Incomplete Multi-View Clustering. *IEEE TIP*, 32: 1354–1366.

Yu, S.; Wang, S.; Zhang, P.; Wang, M.; Wang, Z.; Liu, Z.; Fang, L.; Zhu, E.; and Liu, X. 2024. DVSAI: Diverse View-Shared Anchors Based Incomplete Multi-View Clustering. In *AAAI*, volume 38, 16568–16577.

Yu, X.; Liu, H.; Lin, Y.; Wu, Y.; and Zhang, C. 2022. Auto-weighted sample-level fusion with anchors for incomplete multi-view clustering. *Pattern Recognition*, 130: 108772.

Zhang, C.; Li, H.; Lv, W.; Huang, Z.; Gao, Y.; and Chen, C. 2023a. Enhanced Tensor Low-Rank and Sparse Representation Recovery for Incomplete Multi-View Clustering. In *AAAI*, 11174–11182.

Zhang, C.; Wei, J.; Wang, B.; Li, Z.; Chen, C.; and Li, H. 2023b. Robust Spectral Embedding Completion Based Incomplete Multi-view Clustering. In *ACM MM*, 300–308.

Zhang, C.; Xu, D.; Jia, X.; Chen, C.; and Li, H. 2024. Continual Multi-View Clustering with Consistent Anchor Guidance. In *IJCAI*, 5434–5442.