One-Step Adaptive Graph Learning for Incomplete Multiview Subspace Clustering

Jie Chen, *Member, IEEE*, Hua Mao, Wai Lok Woo, *Senior Member, IEEE*, Chuanbin Liu, Zhu Wang, and Xi Peng, *Senior Member, IEEE*

***Abstract*—Incomplete multiview clustering (IMVC) optimally integrates complementary information within incomplete multiview data to improve clustering performance. Several one-step graph-based methods show great potential for IMVC. However, the low-rank structures of similarity graphs are neglected at the initialization stage of similarity graph construction. Moreover, further investigation into complementary information integration across incomplete multiple views is needed, particularly when considering the low-rank structures implied in high-dimensional multiview data. In this paper, we present one-step adaptive graph learning (OAGL) that adaptively performs spectral embedding fusion to achieve clustering assignments at the clustering indicator level. We first initiate affinity matrices corresponding to incomplete multiple views using spare representation under two constraints, i.e., the sparsity constraint on each affinity matrix corresponding to an incomplete view and the degree matrix of the affinity matrix approximating an identity matrix. This approach promotes exploring complementary information across incomplete multiple views. Subsequently, we perform an alignment of the spectral block-diagonal matrices among incomplete multiple views using low-rank tensor learning theory. This facilitates consistency information exploration across incomplete multiple views. Furthermore, we present an effective alternating iterative algorithm to solve the resulting optimization problem. Extensive experiments on benchmark datasets demonstrate that the proposed OAGL method outperforms several state-of-the-art approaches.**

***Index Terms*—Incomplete multiview clustering, adaptive graph learning, spectral embedding, tensor nuclear norm.**

I. Introduction

MULTIVIEW data usually come from diverse sources or are generated using various feature descriptors [1], [2], [3]. For instance, human facial images are captured by multicamera video surveillance systems, and user behavior records are collected across different social media platforms. These samples are frequently high-dimensional and are approximately drawn from a union of multiple low-dimensional subspaces [4], [5]. In real scenarios, multiview data frequently exhibit incompleteness due to the absence of instances for certain samples across partial views; this is referred to as incomplete multiview data. Incomplete multiview clustering (IMVC) aims to partition samples into distinct clusters by uncovering the intrinsic structures within incomplete multiview data.

IMVC has emerged as a hot topic in unsupervised learning, including subspace learning-based methods [6], [7], [8], [9], graph learning-based methods [10], [11], [12], [13], [14], [15], [16], multiple kernel-based methods [17], [18] and deep learning-based methods [19], [20], [21], [22], [23]. To reveal the intrinsic structures within multiview data, graph-based MVC methods are often combined with manifold learning techniques such as spectral clustering [12], [24] and kernel functions [17], [18]. By virtue of the advantage of effectively capturing the nonlinear structures of multiview data, graph-based IMVC methods have achieved encouraging clustering results. The main objective of graph-based methods is to learn a consensus representation by exploring the underlying graph structure from incomplete multiview data. Graph-based methods can effectively capture complementary information across multiple views.

Most of the existing graph-based IMVC methods can be roughly divided into two categories from the perspective of graph information fusion, including the fusion of affinity matrices [14], [16], [25] and the fusion of spectral embedding representations [11], [12], [13], [26]. For example, Tao et al. [10] presented an ensemble clustering model to exploit the higher-level information contained in the multiview data. Yin et al. [25] proposed a reconstruction strategy of incomplete views to learn a fused affinity matrix. Wen et al. [26] presented a consensus graph learning model that learns a fused spectral embedding representation from the graphs of multiple views. Chen et al. [11] applied spectral rotation fusion to learn a consensus spectral embedding representation from the spectral embedding graphs of multiple views. These methods typically consist of two successive stages. The fused affinity matrix or spectral embedding representation is first acquired from the similarity graphs of multiple views in the first stage. Then, one of the spectral clustering techniques or the k-means algorithm is applied to the fused affinity matrix or consensus spectral embedding representation in the second

Received 5 June 2024; revised 11 October 2024; accepted 15 February 2025. Date of publication 18 February 2025; date of current version 25 March 2025. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant U21B2040, Grant 62176171, and Grant 72374089, in part by the Fundamental Research Funds for the Central Universities under Grant CJ202303 and Grant SCU2023D008, and in part by Sichuan Science and Technology Planning Project under Grant 24NSFTD0130. Recommended for acceptance by X. Zhu. *(Corresponding authors: Zhu Wang; Chuanbin Liu.)*

Jie Chen and Xi Peng are with the College of Computer Science, Sichuan University, Chengdu 610065, China (e-mail: chenjie2010@scu.edu.cn; pengx.gm@gmail.com).

Hua Mao and Wai Lok Woo are with the Department of Computer and Information Sciences, Northumbria University, NE1 8ST Newcastle, U.K. (e-mail: hua.mao@northumbria.ac.uk; wailok.woo@northumbria.ac.uk).

Chuanbin Liu is with the Center for Scientific Research and Development in Higher Education Institutes, Ministry of Education, Beijing 100086, China (e-mail: liuchuanbingl@163.com).

Zhu Wang is with the Law School, Key Laboratory of Data Protection and Intelligent Management, Sichuan University, Chengdu 610065, China (e-mail: wangzhu@scu.edu.cn).

Digital Object Identifier 10.1109/TKDE.2025.3543696

stage, respectively. Although these methods have achieved significant improvements with regard to IMVC, they still have the following limitations. On the one hand, the clustering results of these methods may remain unstable since the k-means algorithm is involved in the second stage. This limits their practical application. On the other hand, the two separate stages may lead to a suboptimal IMVC result [27]. This inevitably affects clustering performance. As a result, integrating graph learning and the clustering process into a unified framework poses a great challenge for researchers.

Many efforts have recently been devoted to addressing one-step multiview clustering [24], [28], [29], [30], [31], [32]. For example, Zhang et al. [31] presented a one-step multiview subspace clustering (OMVSC) method that attempts to identify affinity matrices from multiple views under the F-norm constraint. However, this approach may be insufficient to characterize the subspace structures of high-dimensional multiview data using the F-norm constraint, especially for grossly corrupted original data. Tang et al. [24] presented a unified one-step multiview spectral clustering (UOMSC) method that integrates spectral embedding and k-means into a unified framework to obtain clustering results. In UOMSC, the inner product of a spectral embedding matrix and its transpose is considered a low-rank approximation of the similarity graph. However, the low-rank structure of the similarity graph is neglected at the initialization stage of graph construction. Consequently, IMVC still face significant challenges. First, when the affinity matrices of incomplete multiple views are initialized, the underlying structures of the similarity graphs deserve further investigation in high-dimensional multiview data. Second, how can the information fusion of similarity graphs be accomplished to explore consistency and complementary information across incomplete multiple views in a single-step approach?

To address the abovementioned issues, we present a one-step adaptive graph learning (OAGL) method to learn the clustering assignments for high-dimensional incomplete multiview data. We first construct affinity matrices that correspond to incomplete multiple views using the self-expressiveness property of the existing instances. In contrast to most previous IMVC studies, we utilize the block-diagonal structures implied in the affinity matrices to explore the complementary information across incomplete multiple views. The proposed OAGL method can be considered to be an incomplete multiview subspace clustering (IMVSC) technique. In addition, we introduce low-rank tensor learning theory to implement an alignment of the spectral block-diagonal matrices among incomplete multiple views, which facilitates the exploration of consistency information across incomplete multiple views. Furthermore, we employ spectral rotation to achieve clustering assignments during spectral embedding fusion. Finally, we present an effective alternating iterative algorithm to solve the resulting optimization problem. The major contributions of this paper are highlighted as follows.

- We present an end-to-end OAGL model that can adaptively perform spectral embedding fusion to achieve clustering assignments at the clustering indicator level.
- We exploit the self-expressiveness property of the existing instances to explore complementary information across incomplete multiple views under two additional constraints, i.e., the sparsity constraint on an affinity matrix corresponding to each incomplete view and the degree matrix of the affinity matrix approximating an identity matrix.
- We perform an alignment of the spectral block-diagonal matrices among incomplete multiple views using low-rank tensor learning theory, which is conducive to exploring the consistency information across incomplete multiple views.
- Extensive experiments show the effectiveness and advantages of the proposed OAGL method compared with several state-of-the-art methods.

The remainder of this paper is organized as follows. In Section II, we provide a brief review of the related work. Section III presents the proposed OAGL method in detail. Extensive experiments that are employed to validate the effectiveness of the proposed OAGL method are described in Section IV. Finally, we conclude this paper in Section V.

II. Related Work

A. Spectral Clustering

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \in \mathbb{R}^{d\times n}$ denote a single-view dataset, where d and n represent the dimensionality and the number of samples, respectively. The discrete binary cluster indicator matrix is denoted as $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n]^T \in \{0,1\}^{n\times c}$, where $\mathbf{y}_i \in \mathbb{R}^c$ is the clustering indicator vector, and c denotes the number of clusters. In particular, $y_{ij} = 1$ if $\mathbf{x}_i$ is assigned to the jth cluster, and 0 otherwise.

A weighted undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is represented by a set of nodes $\mathcal{V}$ and a set of edges $\mathcal{E}$, where each node v_i $(1 \le i \le n)$ in $\mathcal{V}$ corresponds to a sample in $\mathbf{X}$, and each edge e_{ij} in $\mathcal{E}$ represents a connection from node i to node j. The similarity graph can be described by a weighted adjacency matrix $\mathbf{W} \in \mathbb{R}^{n\times n}$, where W_{ij} denotes the similarity between $\mathbf{x}_i$ and $\mathbf{x}_j$. The degree $\mathbf{D} \in \mathbb{R}^{n\times n}$ of $\mathcal{G}$ is a diagonal matrix, whose diagonal elements are defined as

$$d_i = \sum_{j=1}^{n} W_{ij}. \tag{1}$$

The objective function of a graph normalized cut in spectral clustering [24], [33] can be formulated as

$$\begin{aligned} &\min_{\mathbf{Y}} tr\left(\left(\mathbf{Y}^T\mathbf{Y}\right)^{-\frac{1}{2}}\mathbf{Y}^T\mathbf{L}\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-\frac{1}{2}}\right) \\ &\text{s.t.} \quad \mathbf{Y} \in \{0,1\}^{n\times c},\ \mathbf{y}_i\mathbf{1}_c = 1 \end{aligned} \tag{2}$$

where $tr(\cdot)$ denotes the trace of the matrix and $\mathbf{L} \in \mathbb{R}^{n\times n}$ is a Laplacian matrix, i.e., $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}$. To avoid seeking the discrete solution of $\mathbf{Y}$, (2) is reformulated as

$$\min_{\mathbf{H}} tr\left(\mathbf{H}^T\mathbf{L}\mathbf{H}\right) \ \text{s.t.}\ \mathbf{H}^T\mathbf{H} = \mathbf{I}_c. \tag{3}$$

where $\mathbf{H} \in \mathbb{R}^{n\times c}$ represents the spectral embedding matrix. The optimal solution $\mathbf{H}$ in (3) consists of the eigenvectors corresponding to the c smallest eigenvalues of $\mathbf{L}$.

The goal of spectral rotation is to derive a discrete indicator matrix from $\mathbf{H}$ [34]. The objective function of spectral rotation can be formulated as follows:

$$\min_{\mathbf{Y},\mathbf{R}} \|\mathbf{HR}-\mathbf{Y}\|_F^2 \text{ s.t. } \mathbf{R}^T\mathbf{R}=\mathbf{I}_c, \mathbf{Y1}_c=\mathbf{1}_n. \tag{4}$$

where $\mathbf{R}\in\mathbb{R}^{c\times c}$ is considered a rotation matrix.

B. One-Step Graph-Based MVC Techniques

Given a high-dimensional multiview data $\{\mathbf{X}_m^{(v)}\in\mathbb{R}^{d_v\times n}\}_{v=1}^{n_v}$ with n samples from n_v views, $\mathbf{X}_m^{(v)}$ denotes the vth view, and d_v represents the dimensionality of instances in the vth view.

To date, several representative one-step graph-based methods have been proposed for MVC [24], [31], [32]. For example, UOMSC jointly exploits the complementary information of view-specific graphs and spectral embedding matrices in a unified graph [24]. In UOMSC, $\mathbf{W}_x^{(v)}\in\mathbb{R}^{n\times n}$ and $\mathbf{W}_h^{(v)}\in\mathbb{R}^{n\times n}$ represent an original affinity matrix constructed from $\mathbf{X}_m^{(v)}$ and a low-rank spectral embedding matrix, respectively. The optimization problem of the UOMSC method is formulated as

$$\begin{aligned}&\max_{\mathbf{H},\mathbf{Y},\alpha_v,\beta_v} tr\left(\mathbf{H}^T\mathbf{W}\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-\frac{1}{2}}\right)\\ &\text{s.t. } \mathbf{H}^T\mathbf{H}=\mathbf{I}_c, \sum_{v=1}^{n_v}\alpha_v^2=1,\alpha_v\geq 0, \sum_{v=1}^{n_v}\beta_v^2=1,\beta_v\geq 0,\\ &\mathbf{W}=(1-\lambda)\sum_{v=1}^{n_v}\alpha_v\mathbf{W}_x^{(v)}+\lambda\sum_{v=1}^{n_v}\beta_v\mathbf{W}_h^{(v)}\end{aligned} \tag{5}$$

where λ, α_v and β_v are trade-off parameters.

One-step graph-based methods assume that all the views are complete [24], [31], [32]. However, each view may contain different percentages of missing instances in incomplete multiview data. Consequently, it is still an open problem to design a reasonable one-step graph-based scheme for IMVC.

C. Tensor Nuclear Norm Theory

Given two tensors $\mathcal{A}\in\mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathcal{B}\in\mathbb{R}^{n_2\times m\times n_3}$, the t-product of $\mathcal{A}$ and $\mathcal{B}$ is defined as

$$\mathcal{C}=\mathcal{A}*\mathcal{B}=fold\left(bcirc\left(\mathcal{A}\right)\cdot unfold\left(\mathcal{B}\right)\right) \tag{6}$$

where $\mathcal{C}\in\mathbb{R}^{n_1\times m\times n_3}$. The $unfold(\cdot)$ operator maps $\mathcal{B}$ to a two-dimensional matrix of size $(n_1n_3)\times n_2$, the $fold(\cdot)$ operator performs an inverse operation of $unfold(\cdot)$, and $bcirc(\cdot)$ maps a tensor to a block circulant matrix [35].

A tensor $\mathcal{X}\in\mathbb{R}^{n_1\times n_2\times n_3}$ is f-diagonal if each of its frontal slices is a diagonal matrix. The tensor singular value decomposition (t-SVD) of $\mathcal{X}$ is factorized as

$$\mathcal{X}=\mathcal{U}*\mathcal{S}*\mathcal{V}^T \tag{7}$$

where $\mathcal{U}\in\mathbb{R}^{n_1\times n_1\times n_3}$ and $\mathcal{V}\in\mathbb{R}^{n_2\times n_2\times n_3}$ are two orthogonal tensors and $\mathcal{S}\in\mathbb{R}^{n_1\times n_2\times n_3}$ is an f-diagonal tensor [35].

The tensor nuclear norm of the tensor $\mathcal{X}$ is defined as

$$\|\mathcal{X}\|_*=\sum_{i=1}^{r}S(i,i,1) \tag{8}$$

Algorithm 1: The t-SVT Operator [36].

Input: $\mathcal{A}\in\mathbb{R}^{n_1\times n_2\times n_3}$ and a parameter $\lambda>0$.
Computing $\bar{\mathcal{A}}=fft(\mathcal{A},[],3)$, where fft is a MATLAB command;
for $i=1,\ldots,\frac{\lceil n_3+1\rceil}{2}$ **do**
  /* Considering each frontal slice of $\bar{\mathcal{A}}$ */
  $[\mathbf{U},\mathbf{S},\mathbf{V}]=svd(\bar{\mathbf{A}}^{(i)})$;
  $\bar{\mathbf{W}}^{(i)}=\mathbf{U}(\mathbf{S}-\lambda)_+\mathbf{V}^T$;
end for
for $i=\frac{\lceil n_3+1\rceil}{2}+1,\ldots,n_3$ **do**
  $\bar{\mathbf{W}}^{(i)}=conj(\bar{\mathbf{W}}^{(n_3-i+2)})$;
end for
Output: $\mathcal{D}_\lambda(\mathcal{A})=ifft(\bar{\mathcal{W}},[],3)$, where $ifft$ is a MATLAB command.

where $S(i,i,1)$ $(1\leq i\leq r)$ denotes a singular value of the tensor $\mathcal{X}$, and r is the tubal rank of $\mathcal{X}$ [35], [36]. To seek a low-tubal-rank approximation of $\mathcal{A}$, the optimization problem can be formulated as

$$\min_{\mathcal{X}\in\mathbb{R}^{n_1\times n_2\times n_3}}\lambda\|\mathcal{X}\|_*+\frac{1}{2}\|\mathcal{X}-\mathcal{A}\|_F^2. \tag{9}$$

The optimization solution to (9) can be calculated according to a proximal operator of the matrix nuclear norm. The tensor-singular-value thresholding (t-SVT) operator is a proximal operator whose details are summarized in Algorithm 1 [36].

Tensor nuclear norm-based techniques have been introduced to capture higher-order information across multiple views. For example, Xie et al. [16] proposed a multiview clustering method based on tensor-singular value decomposition (t-SVD), which refines view-specific subspaces from different views. Similarly, Chen et al. [12] developed a multiview embedding matrix fusion model that leverages complementary information by identifying higher-order correlations among multiple views. These approaches have demonstrated effectiveness in multiview clustering.

III. One-Step Adaptive Graph Learning

In this section, we present an OAGL method that learns clustering assignments in a single step for IMVC, including the motivation of the proposed OAGL method, the details of the OAGL model and an optimization strategy of the overall objective function of OAGL.

A. Motivation

Given incomplete multiview data $\{\mathbf{X}^{(v)}\in\mathbb{R}^{d_v\times n}\}_{v=1}^{n_v}$ with c clusters, each view $\mathbf{X}^{(v)}$ contains n instances. As each view may have missing instances for some of the samples, the remainder of the instances are denoted as $\mathbf{X}_r^{(v)}\in\mathbb{R}^{d_v\times N_v}$, where N_v represents the number of existing instances in the vth view.

Without loss of generality, we assume that all instances in each view are rearranged according to the index order of the clusters, i.e., $\{\mathbf{X}^{(v)}=[\mathbf{X}_1^{(v)},\mathbf{X}_2^{(v)},\ldots\mathbf{X}_c{}^{(v)}]\}_{v=1}^{n_v}$. Considering an affinity matrix $\mathbf{W}^{(v)}\in\mathbb{R}^{n\times n}$ constructed from $\mathbf{X}^{(v)}$, we

can achieve an individual spectral embedding representation $\mathbf{H}^{(v)} \in \mathbb{R}^{n\times c}$ using (3). The inner product of $\mathbf{H}^{(v)}$ and its transpose is denoted as $\mathbf{H}^{(v)}(\mathbf{H}^{(v)})^T$, which contains a low-rank structure. It can be employed as an affinity matrix for measuring the relationships among samples. Ideally, it consists of c diagonal blocks. Therefore, $\mathbf{H}^{(v)}(\mathbf{H}^{(v)})^T$ is regarded as a spectral block-diagonal matrix.

When the degree matrix $\mathbf{D}^{(v)}$ of $\mathbf{W}^{(v)}$ satisfies an additional condition $\mathbf{D}^{(v)} = \mathbf{I}_n$ [24], (3) can be further reformulated as

$$\min_{\mathbf{H}^{(v)}} \left\| \mathbf{W}^{(v)} - \mathbf{H}^{(v)} \left(\mathbf{H}^{(v)}\right)^T \right\|_F^2 \text{ s.t. } \left(\mathbf{H}^{(v)}\right)^T \mathbf{H}^{(v)} = \mathbf{I}_c. \tag{10}$$

As indicated in (10), $\mathbf{W}^{(v)}$ approximates the spectral block-diagonal matrix under the F-norm constraint; this motivates us to pursue block-diagonal structures of the affinity matrix $\mathbf{W}^{(v)}$ constructed from $\mathbf{X}^{(v)}$. The task of integrating complementary information across incomplete multiple views is implemented by performing low-rank alignment of the spectral block-diagonal matrices among multiple views.

B. Proposed OAGL Model

High-dimensional multiview data are often considered to be a collection of samples approximately drawn from a union of multiple low-dimensional subspaces [4], [5]. To pursue block-diagonal structures of the affinity matrix $\mathbf{W}^{(v)}$, we exploit the self-expressiveness property of the existing instances to initialize the affinity matrix $\mathbf{W}^{(v)}$. Specifically, each existing data instance $\mathbf{x}_r^{(v)}$ in the vth incomplete view can be represented by a small number of other existing instances under the sparsity constraint, i.e.,

$$\begin{aligned} &\min_{\mathbf{w}_r^{(v)}} \left\| \mathbf{x}_r^{(v)} - \mathbf{X}_r^{(v)} \mathbf{w}_r^{(v)} \right\|_2^2 + \lambda \left\| \mathbf{w}_r^{(v)} \right\|_0 \\ &\text{s.t} \quad \sum_{i=1}^{N_v} \mathbf{w}_i^{(v)} = 1,\ w_{ir}^{(v)} = 0 \end{aligned} \tag{11}$$

where λ denotes a balance parameter, the sparse coefficients in $\mathbf{w}_r^{(v)} \in \mathbb{R}^{N_v}$ represent the relationship between $\mathbf{x}_r^{(v)}$ and the other existing instances in the vth incomplete view, and $\|\cdot\|_2$ indicates a specific regularization strategy for characterizing noise. Since (11) is a general NP-hard problem [37], it is reformulated as a convex relaxation of the l_1-norm, i.e.,

$$\begin{aligned} &\min_{\mathbf{w}_r^{(v)}} \left\| \mathbf{x}_r^{(v)} - \mathbf{X}_r^{(v)} \mathbf{w}_r^{(v)} \right\|_2^2 + \lambda \left\| \mathbf{w}_r^{(v)} \right\|_1 \\ &\text{s.t} \quad \sum_{i=1}^{N_v} \mathbf{w}_i^{(v)} = 1,\ w_{ir}^{(v)} = 0. \end{aligned} \tag{12}$$

The optimization problem in (12) can be efficiently solved using convex optimization tools [38], [39].

Suppose the instances are noise-free; then, we have

$$\left\| \mathbf{x}_r^{(v)} - \mathbf{X}_r^{(v)} \mathbf{w}_r^{(v)} \right\|_2^2 = 0. \tag{13}$$

This indicates that a set of existing instances is strictly drawn from a union of c independent subspaces of unknown dimensions. We set $\mathbf{W}_r^{(v)} = [\mathbf{w}_1^{(v)}, \mathbf{w}_2^{(v)}, \ldots, \mathbf{w}_{N_v}^{(v)}] \in \mathbb{R}^{N_v \times N_v}$. Ideally, $\mathbf{W}_r^{(v)}$ has c distinct diagonal blocks corresponding to the c subspaces. Each element $w_{ij}^{(v)}$ $(1 \le i \le N_v, 1 \le j \le N_v)$ in $\mathbf{W}_r^{(v)}$ represents the similarity between the ith and jth existing instances of the vth incomplete view. We fill each element of $\mathbf{W}_r^{(v)}$ into the corresponding position of $\mathbf{W}^{(v)}$. The similarities between the remaining pairs of the instances in $\mathbf{W}^{(v)}$ are set to zero. Thus, $\mathbf{W}^{(v)}$ remains as block-diagonal structures.

Due to noise and the incompleteness of the instances in high-dimensional multiview data, $\mathbf{W}^{(v)}$ may not contain accurate similarity among the instances of the vth incomplete view. Consequently, the clustering performance inevitably degenerates when we perform information fusion of the affinity matrices among multiple views. Thus, we introduce the third-order low-rank tensor to characterize low-rank structures of the spectral block-diagonal matrices among multiple views. Specifically, the third-order tensor $\mathcal{H} \in \mathbb{R}^{n\times n\times n_v}$ is first constructed by stacking $\{\mathbf{H}^{(v)}(\mathbf{H}^{(v)})^T\}_{v=1}^{n_v}$. Subsequently, we perform low-rank alignment of the spectral block-diagonal matrices to uncover a unified low-rank structure closely associated with spectral embedding representations. To implement information fusion of the spectral block-diagonal matrices across multiple views, we employ a tensor flip operation on $\mathcal{H}$ to obtain a surrogate third-order tensor $\mathcal{T} = \{\mathbf{H}^{(v)}(\mathbf{H}^{(v)})^T\}_{v=1}^{n_v} \in \mathbb{R}^{n\times n_v\times n}$, where first frontal slices of $\mathcal{T}$ are matrices of size $n \times n_v$. The purpose of introducing the tensor flip operation on $\mathcal{H}$ is to examine the correlations across multiple views [16].

To implement one-step adaptive graph learning, we perform spectral embedding fusion of multiple affinity matrices at the clustering indicator level. During spectral embedding fusion, we exploit spectral rotation to yield a binary clustering indicator matrix by considering the differences among multiple views. Therefore, the spectral embedding fusion at the clustering indicator level can be formulated as

$$\begin{aligned} &\min_{\mathbf{Y},\mathbf{R}^{(v)},\lambda^{(v)}} \sum_{v=1}^{n_v} \left\| \mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2} - \lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)} \right\|_F^2 \\ &\text{s.t.} \left(\mathbf{R}^{(v)}\right)^T \mathbf{R}^{(v)} = \mathbf{I}_c, \lambda^{(v)} > 0, \sum_{v=1}^{n_v} \left(\lambda^{(v)}\right)^2 = 1, \\ &\quad Y_{ij} \in \{0,1\},\ \mathbf{Y}_{i,:}\mathbf{1}_c = 1 \end{aligned} \tag{14}$$

where $\lambda^{(v)} \in \mathbb{R}$ is a weight factor that adaptively balances the importance of different views, $\mathbf{R}^{(v)} \in \mathbb{R}^{c\times c}$ represents the spectral rotation matrix, and $\mathbf{Y} \in \mathbb{R}^{n\times c}$ denotes the clustering indicator matrix. The proposed OAGL model unifies the adaptive spectral embedding fusion and the low-rank alignment of the spectral block-diagonal matrices, which can be implemented by spectral rotation and spectral embedding tensor learning, respectively. The final objective function of the OAGL can be formulated as

$$\min_{\mathbf{Y},\mathbf{H}^{(v)},\mathbf{R}^{(v)},\lambda^{(v)}} \|\mathcal{T}\|_* + \frac{\alpha}{2} \sum_{v=1}^{n_v} \left\| \mathbf{W}^{(v)} - \mathbf{H}^{(v)} \left(\mathbf{H}^{(v)}\right)^T \right\|_F^2$$

$$+\frac{\beta}{2}\sum_{v=1}^{n_v}\left\|\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}-\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}\right\|_F^2$$
$$\text{s.t.}\ \left(\mathbf{H}^{(v)}\right)^T\mathbf{H}^{(v)}=\mathbf{I}_c,\left(\mathbf{R}^{(v)}\right)^T\mathbf{R}^{(v)}=\mathbf{I}_c,\lambda^{(v)}>0,$$
$$\sum_{v=1}^{n_v}\left(\lambda^{(v)}\right)^2=1,Y_{ij}\in\{0,1\},\mathbf{Y}_{i,:}\mathbf{1}_c=1 \quad (15)$$

where α and β are balance parameters.

C. Optimization Strategy

We provide an alternating iterative algorithm based on the alternating direction method of multipliers (ADMM) framework [40] to efficiently solve the optimization problem in (15). An auxiliary tensor variable $\mathcal{G}\in\mathbb{R}^{n\times n_v\times n}$ is first introduced into (15). Then, this optimization problem can be converted into the following equivalent problem:

$$\min_{\mathbf{Y},\mathbf{H}^{(v)},\mathbf{R}^{(v)},\lambda^{(v)}}\|\mathcal{G}\|_*+\frac{\alpha}{2}\sum_{v=1}^{n_v}\left\|\mathbf{W}^{(v)}-\mathbf{H}^{(v)}\left(\mathbf{H}^{(v)}\right)^T\right\|_F^2$$
$$+\frac{\beta}{2}\sum_{v=1}^{n_v}\left\|\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}-\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}\right\|_F^2$$
$$\text{s.t.}\ \left(\mathbf{H}^{(v)}\right)^T\mathbf{H}^{(v)}=\mathbf{I}_c,\left(\mathbf{R}^{(v)}\right)^T\mathbf{R}^{(v)}=\mathbf{I}_c,\lambda^{(v)}>0,$$
$$\sum_{v=1}^{n_v}\left(\lambda^{(v)}\right)^2=1,Y_{ij}\in\{0,1\},\mathbf{Y}_{i,:}\mathbf{1}_c=1,\mathcal{G}=\mathcal{T}. \quad (16)$$

The augmented Lagrangian function in (16) is

$$\mathcal{L}\left(\mathcal{G},\mathbf{Y},\mathbf{H}^{(v)},\mathbf{R}^{(v)},\lambda^{(v)}\right)$$
$$=\|\mathcal{G}\|_*+\frac{\alpha}{2}\sum_{v=1}^{n_v}\left\|\mathbf{W}^{(v)}-\mathbf{H}^{(v)}\left(\mathbf{H}^{(v)}\right)^T\right\|_F^2$$
$$+\frac{\beta}{2}\sum_{v=1}^{n_v}\left\|\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}-\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}\right\|_F^2$$
$$+\langle\mathcal{S},\mathcal{T}-\mathcal{G}\rangle+\frac{\mu}{2}\|\mathcal{T}-\mathcal{G}\|_F^2 \quad (17)$$

where $\mathcal{S}\in\mathbb{R}^{n\times n_v\times n}$ is a Lagrange multiplier, and $\mu>0$ is an adaptive penalty factor. The augmented Lagrangian function of (17) can be equivalently converted into the following function as

$$\mathcal{L}\left(\mathcal{G},\mathbf{Y},\mathbf{H}^{(v)},\mathbf{R}^{(v)},\lambda^{(v)}\right)$$
$$=\|\mathcal{G}\|_*+\frac{\alpha}{2}\sum_{v=1}^{n_v}\left\|\mathbf{W}^{(v)}-\mathbf{H}^{(v)}\left(\mathbf{H}^{(v)}\right)^T\right\|_F^2$$
$$+\frac{\beta}{2}\sum_{v=1}^{n_v}\left\|\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}-\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}\right\|_F^2$$
$$+\langle\mathcal{S},\mathcal{T}-\mathcal{G}\rangle+\frac{\mu}{2}\|\mathcal{T}-\mathcal{G}\|_F^2. \quad (18)$$

There are five variables in (18): $\mathcal{G}$, $\mathbf{Y}$, $\mathbf{H}^{(v)}$, $\mathbf{R}^{(v)}$ and $\lambda^{(v)}$. We alternately update one variable until convergence while fixing the others. Thus, the optimization problem in (15) can be divided into five subproblems as follows:

$\mathbf{Y}$-subproblem: When fixing the other four variables $\mathcal{G}$, $\mathbf{H}^{(v)}$, $\mathbf{R}^{(v)}$ and $\lambda^{(v)}$, (18) can be rewritten as

$$\min_{\mathbf{Y}}\sum_{v=1}^{n_v}\left\|\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}-\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}\right\|_F^2$$
$$\text{s.t.}\ Y_{ij}\in\{0,1\},\ \mathbf{Y}_{i,:}\mathbf{1}_c=1. \quad (19)$$

The problem in (19) is equivalent to the following optimization problem:

$$\max_{\mathbf{Y}}tr\left(\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}\mathbf{Y}^T\mathbf{Q}\right)$$
$$\text{s.t.}\ Y_{ij}\in\{0,1\},\ \mathbf{Y}_{i,:}\mathbf{1}_c=1 \quad (20)$$

where $\mathbf{Q}=\sum_{v=1}^{n_v}\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}$. Eq (20) can be transformed to

$$\max_{\mathbf{Y}}\sum_{m=1}^{c}\frac{\sum_{i=1}^{n}y_{im}q_{im}}{\sqrt{\mathbf{y}_m^T\mathbf{y}_m}}\quad\text{s.t.}\quad y_{im}\in\{0,1\},\ \mathbf{Y}_{i,:}\mathbf{1}_c=1 \quad (21)$$

where $\mathbf{y}_m$ denotes the mth column of $\mathbf{Y}$ and q_{im} represents the entry in the ith row and mth column of $\mathbf{Q}$. Each row of $\mathbf{Y}$ can be solved using an incremental mechanism [24], [41], which is formulated as

$$\rho_{im}=\frac{\sum_{j=1}^{n}y_{jm}q_{jm}+q_{im}\left(1-y_{im}\right)}{\sqrt{\mathbf{y}_m^T\mathbf{y}_m+\left(1-y_{im}\right)}}-\frac{\sum_{j=1}^{n}y_{jm}q_{jm}-q_{im}y_{im}}{\sqrt{\mathbf{y}_m^T\mathbf{y}_m-y_{im}}}. \quad (22)$$

The clustering indicator r in $\mathbf{y}_i$, which represents the ith row of $\mathbf{Y}$, can be calculated as

$$r=\arg\max_{1\le m\le c}\rho_{im}. \quad (23)$$

$\mathbf{R}^{(v)}$-subproblem: With the other four variables $\mathcal{G}$, $\mathbf{Y}$, $\mathbf{H}^{(v)}$ and $\lambda^{(v)}$ being fixed, (18) can be simplified as

$$\min_{\mathbf{R}^{(v)}}\sum_{v=1}^{n_v}\left\|\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}-\lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)}\right\|_F^2$$
$$\text{s.t.}\ \left(\mathbf{R}^{(v)}\right)^T\mathbf{R}^{(v)}=\mathbf{I}_c \quad (24)$$

The problem in (24) is equivalent to

$$\max_{\mathbf{R}^{(v)}}\left(\mathbf{R}^{(v)}\right)^T\left(\mathbf{H}^{(v)}\right)^T\mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2}$$
$$\text{s.t.}\ \left(\mathbf{R}^{(v)}\right)^T\mathbf{R}^{(v)}=\mathbf{I}_c. \quad (25)$$

The economy SVD of the matrix $\mathbf{A}^{(v)}$ is $\mathbf{A}^{(v)}=\mathbf{U}_c\mathbf{\Sigma}_c\mathbf{V}_c^T$, where $\mathbf{A}^{(v)}=(\mathbf{H}^{(v)})^T\mathbf{Y}(\mathbf{Y}^T\mathbf{Y})^{-1/2}$. The problem in (24) has a closed-form solution [34], i.e.,

$$\mathbf{R}^{(v)}=\mathbf{U}_c\mathbf{V}_c^T. \quad (26)$$

$\mathbf{H}^{(v)}$-subproblem: When the other four variables $\mathcal{G}$, $\mathbf{Y}$, $\mathbf{R}^{(v)}$ and $\lambda^{(v)}$ are fixed, the problem in (18) can be rewritten as

$$\begin{aligned}\min_{\mathbf{H}^{(v)}} \alpha \sum_{v=1}^{n_v} \left\| \mathbf{W}^{(v)} - \mathbf{H}^{(v)} \left(\mathbf{H}^{(v)}\right)^T \right\|_F^2 \\ + \beta \sum_{v=1}^{n_v} \left\| \mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2} - \lambda^{(v)}\mathbf{H}^{(v)}\mathbf{R}^{(v)} \right\|_F^2 \\ + \mu \left\| \mathcal{G} - \left(\mathcal{T} + \frac{\mathcal{S}}{\mu}\right) \right\|_F^2 \\ \text{s.t.} \quad \left(\mathbf{H}^{(v)}\right)^T \mathbf{H}^{(v)} = \mathbf{I}_c \end{aligned} \tag{27}$$

Setting $\mathcal{M} = \mathcal{G} - \frac{\mathcal{S}}{\mu}$ in (27), we have

$$\max_{\mathbf{H}^{(v)}} tr\left(\left(\mathbf{H}^{(v)}\right)^T \mathbf{B}^{(v)}\mathbf{H}^{(v)}\right) + tr\left(\left(\mathbf{H}^{(v)}\right)^T \mathbf{C}^{(v)}\right) \tag{28}$$

where $\mathbf{B}^{(v)} = \alpha\mathbf{W}^{(v)} + \frac{\mu}{2}(\mathcal{M}^{(v)} + (\mathcal{M}^{(v)})^T)$ and $\mathbf{C}^{(v)} = \beta\lambda^{(v)}\mathbf{Y}(\mathbf{Y}^T\mathbf{Y})^{-1/2}(\mathbf{R}^{(v)})^T$. The problem in (28) can be relaxed to

$$\max_{\mathbf{H}^{(v)}} tr\left(\left(\mathbf{H}^{(v)}\right)^T \mathbf{P}^{(v)}\right) \text{ s.t. } \left(\mathbf{H}^{(v)}\right)^T \mathbf{H}^{(v)} = \mathbf{I}_c \tag{29}$$

where $\mathbf{P}^{(v)} = \mathbf{B}^{(v)}\mathbf{H}^{(v)} + \mathbf{C}^{(v)}$. The problem in (29) can be solved by the generalized power iteration (GPI) algorithm [42], where the optimal solution is obtained by iteratively updating $\mathbf{H}^{(v)}$ after an initialization of $\mathbf{H}^{(v)}$ is provided.

$\mathcal{G}$-subproblem: By fixing the other four variables $\mathbf{Y}$, $\mathbf{H}^{(v)}$, $\mathbf{R}^{(v)}$ and $\lambda^{(v)}$, the problem in (18) can be converted to

$$\min_{\mathcal{G}} \|\mathcal{G}\|_* + \frac{\mu}{2}\left\| \mathcal{G} - \left(\mathcal{T} + \frac{\mathcal{S}}{\mu}\right)\right\|_F^2. \tag{30}$$

The problem in (30) can be solved by Algorithm 1 [36].

$\lambda^{(v)}$-subproblem: When the other four variables $\mathcal{G}$, $\mathbf{Y}$, $\mathbf{R}^{(v)}$ and $\mathbf{H}^{(v)}$ are fixed, we have

$$\begin{aligned}\max_{\lambda^{(v)}} tr\left(\sum_{v=1}^{n_v} \lambda^{(v)} \left(\mathbf{H}^{(v)}\right)^T \mathbf{Y}\left(\mathbf{Y}^T\mathbf{Y}\right)^{-1/2} \left(\mathbf{R}^{(v)}\right)^T\right) \\ \text{s.t.} \quad \sum_{v=1}^{n_v} \left(\lambda^{(v)}\right)^2 = 1,\ \lambda^{(v)} > 0 \end{aligned} \tag{31}$$

The closed-form solution of the problem in (31) can be updated as

$$\lambda^{(v)} = \frac{\delta^{(v)}}{\sqrt{\sum_{v=1}^{n_v}\left(\delta^{(v)}\right)^2}} \tag{32}$$

where $\delta^{(v)} = tr((\mathbf{H}^{(v)})^T\mathbf{Y}(\mathbf{Y}^T\mathbf{Y})^{-1/2}(\mathbf{R}^{(v)})^T)$ [24].

In addition, the Lagrange multiplier $\mathcal{S}$ and penalty parameter μ are updated as

$$\mathcal{S}^{(v)} \leftarrow \mathcal{S}^{(v)} + \mu\left(\mathcal{G} - \mathcal{T}\right) \tag{33}$$

$$\mu \leftarrow \min\left(\rho\mu, \mu_{\max}\right) \tag{34}$$

where ρ and $\mu_{\max}$ are constants.

Algorithm 2: Solving the Problem in (15) Using the ADMM Framework.

Input: Data matrices $\{\mathbf{H}^{(v)}\}_{v=1}^{n_v}$ and $\{\mathbf{W}^{(v)}\}_{v=1}^{n_v}$ and parameters α and β.

initialize: $\rho = 1.2$, $\mu = 10^{-4}$, $\mu_{\max} = 10^6$, $\varepsilon = 10^{-6}$, $t = 1$ and $maxIters = 200$;
while not converged **do**
  Update $\mathbf{Y}$ via (19);
  Update $\{\mathbf{R}^{(v)}\}_{v=1}^{n_v}$ via (24);
  Update $\{\mathbf{H}^{(v)}\}_{v=1}^{n_v}$ via (27);
  Update $\mathcal{G}$ via (30);
  Update $\lambda^{(v)}$ via (31);
  Update the Lagrange multiplier $\mathcal{S}$ via (33);
  Update the penalty parameter μ via (34);
  Check the following convergence condition,
  $\|\mathcal{G} - \mathcal{T}\|_{\max} < \varepsilon$;
  if $t > maxIters$ or $converged$ **then**
    break;
  end if
  $t \leftarrow t + 1$;
end while
Output: $\mathbf{Y}$.

The iterative optimization procedure stops when the convergence condition is satisfied, i.e., $\|\mathcal{G} - \mathcal{T}\|_{\max} < \varepsilon$, where ε is a small constant, e.g., $\varepsilon = 1e^{-6}$. We summarize the entire procedure of the optimization for solving the objective function in (15) in Algorithm 2.

D. Theoretical Analysis

1) Why Sparse Representation is Needed in IMVSC? Sparse representations jointly recover subspace structures of samples in high-dimensional multiview data. In addition to the missing instances in incomplete multiple views, the existing instances are often corrupted by noise. As a result, the instances do not lie perfectly in a union of subspaces. For each existing data instance in an incomplete view, the nonzero elements of the sparse representation, referred to as sparse coefficients, correspond to a few of the other existing instances from the same subspace. Sparse coefficients encode the memberships of the existing instances, which can be employed in spectral embedding fusion. Such construction of affinity matrices can reduce the negative influence of spectral embedding fusion when a fraction of the instances is missing in incomplete multiple views.

In sparse representation, the additional constraint $\sum_{i=1}^{N_v} \mathbf{w}_i^{(v)} = 1$ is introduced in (12). This approach offers two advantages for IMVSC. On the one hand, this constraint requires that the sum of the sparse coefficients for each existing data instance equals 1. This ensures that $\mathbf{D}^{(v)} = \mathbf{I}_n$ is implied in (15). On the other hand, the sparse coefficients corresponding to different instances across incomplete multiple views are uniformly constrained to the same scale. Consequently, spectral embedding fusion can effectively reduce redundant information from the original incomplete multiview data in a single step.

TABLE I
STATISTICS OF THE EXPERIMENTAL DATASETS

Dataset	Clusters	Views	Samples
MSRC-v1	7	5	210
BBC	5	4	685
Flower17	17	7	1,360
Handwritten	10	6	2,000
NUSwide	12	6	2,400
Scene-15	15	3	4,485
Caltech101	101	6	8,677

2) Convergence Analysis: It is difficult to theoretically prove the convergence of the proposed OAGL method under the ADMM framework [40]. Fortunately, the optimization problem in (15) can be divided into five subproblems. Each subproblem can achieve an optimal solution. The auxiliary tensor variable $\mathcal{G} \in \mathbb{R}^{n\times n_v\times n}$ is introduced into the optimization problem in (15) with the constraint $\mathcal{G} = \mathcal{T}$. Consequently, the proposed OAGL method can find a locally optimal solution when the convergence condition $\|\mathcal{G} - \mathcal{T}\|_{\max} < \varepsilon$ is reached during the iterations, where $\|\cdot\|_{\max}$ denotes the maximum value of all the elements in a matrix. In addition, the proposed OAGL method converges well in practice, which can be validated in experiments.

3) Computational Complexity Analysis: In Algorithm 2, the computational complexity of the proposed OAGL method consists of five critical steps. First, the computational complexity of updating $\mathbf{Y}$ is $\mathcal{O}(nc)$. Second, the computational complexity of calculating the SVD of $\mathbf{A}^{(v)}$ is $\mathcal{O}(n_v c^3)$ for solving $\mathbf{R}^{(v)}$. Next, the computational complexity of calculating the SVD of $\mathbf{P}^{(v)}$ is $\mathcal{O}(t_1 n_v n^3)$, where t_1 is the number of iterations of the GPI algorithm. Then, the computational complexity of updating $\mathcal{G}$ is $\mathcal{O}(n_v n^2 \log(n) + n_v^2 n^2)$. Finally, the computational complexity of updating $\lambda^{(v)}$ is $\mathcal{O}(n_v c^3)$. Consequently, the computational complexity of Algorithm 2 is $\mathcal{O}(t_2(nc + 2n_v c^3 + t_1 n_v n^3 + n_v n^2 \log(n) + n_v^2 n^2))$, where t_2 is the number of iterations of Algorithm 2. The final computational complexity of Algorithm 2 is $\mathcal{O}(tn^3)$ if $n_v \ll n$ and $c \ll n$, where $t = t_1 t_2$.

IV. EXPERIMENTS

In this section, we conduct a series of experiments to validate the effectiveness of the proposed OAGL method. The source code for OAGL is written in MATLAB 2022b. All of the experiments are conducted on a Windows 11 platform with an Intel Core i7-10700 CPU and 32 GB of RAM. The source code for the OAGL is available online at https://codeocean.com/capsule/7420051/tree/v1.

A. Experimental Settings

1) Datasets: Seven complete benchmark multiview datasets are employed to evaluate the proposed OAGL method. The statistics of the datasets used are summarized in Table I. Details of these multiview datasets are provided as follows:

- **MSRC-v1 dataset**:[1] This dataset contains 210 images for 7 classes of scene recognition. Each image is characterized by five distinct feature sets.
- **BBC dataset** [43]: This dataset contains 685 documents in 5 categories collected from the BBC news website. Each document is represented by four views.
- **Flower17 dataset** [44]: This dataset contains 1,360 flower images that belong to 17 categories. Each image includes seven views.
- **Handwritten dataset** [45]: This dataset consists of 2,000 handwritten images of digits 0 to 9. Each image is described by six different features.
- **NUSwide dataset** [46]: This dataset consists of 2,400 images in 12 categories. Each image includes six views.
- **Outdoor Scene (O-Scene) dataset** [47]: This dataset contains 2, 688 images in 8 categories. Each image is characterized by four distinct features.
- **Caltech101 dataset** [48]: This dataset consists of 8,677 images of objects in 101 classes. Each object has approximately 30-800 images. In particular, the background category has been removed.

2) Comparison Methods: We compare OAGL with several state-of-the-art methods, including consensus spectral rotation fusion (CSRF) [11], graph structure refining (GSR) [49], UOMSC [24], efficient and effective one-step multiview clustering (E^{2}OMVC) [29], adaptive graph learning and spectral rotation (AGLSR) [28] and spectral embedded clustering (SEC) [30]. These comparison methods can be divided into two categories, including one-step and two-stage methods. Two-stage graph-based methods typically involve two successive stages, including learning a consensus representation from incomplete multiview data, and applying k-means to the consensus representation, e.g., CSRF and GSR. The remainder of the comparison methods accomplish the IMVC task in a single step. The construction of $\{\mathbf{W}^{(v)}\}_{v=1}^{n_v}$ is a preprocessing step in the proposed OAGL method. In the experiments, an adaptive neighbor graph learning (ANGL) method [50] is used to initialize individual sparse affinity matrices for incomplete multiview data. In addition, we employ an additional baseline for comparison. First, the affinity matrices corresponding to incomplete multiple views can be obtained using sparse representation, which is also used in OAGL. Then, we aggregate all the affinity matrices into an accumulated affinity matrix. Finally, we apply a standard spectral clustering (SC) method to the accumulated affinity matrix, e.g., NCut [33]. The baseline is regarded as the SC method. Additionally, the missing entries in the affinity matrices constructed by the four one-step MVC methods are filled with zeros. The source codes for the other competing methods were provided by their respective authors.

3) Evaluation Metrics: Three widely used clustering metrics are employed to evaluate the clustering performance of all the algorithms, namely, the clustering accuracy (ACC), the normalized mutual information (NMI) and the adjusted rand index (ARI) [51], [52]. For each clustering metric, a higher value indicates superior clustering performance.

[1]https://www.microsoft.com/en-us/research/project/image-understanding/

TABLE II
CLUSTERING RESULTS (MEAN ± STD.) OF DIFFERENT METHODS ON SEVEN MULTIVIEW DATASETS WITH VARIOUS MISSING DATA RATIOS

Datasets	Methods	ACC				NMI				ARI			
		0	10%	30%	50%	0	10%	30%	50%	0	10%	30%	50%
MSRC-v1	SC	86.19±0.00	88.57±0.00	80.95±0.39	68.81±0.32	76.45±0.00	79.18±0.00	67.26±0.51	54.62±0.2	72.04±0.00	75.82±0.00	59.64±0.56	46.95±0.21
	CSRF	93.33±0.00	90.95±0.00	82.86±0.00	73.48±0.23	85.8±0.00	82.38±0.00	72.05±0.00	57.69±0.25	84.77±0.00	79.74±0.00	61.92±0.00	48.84±0.35
	GSR	99.52±0.00	99.05±0.00	97.14±0.00	96.1±0.3	98.92±0.00	98.19±0.00	94.38±0.00	92.66±0.07	98.88±0.00	97.79±0.00	93.52±0.00	91.35±0.07
	UOMSC	74.76±0	58.57±0	49.52±0	24.29±0	70.4±0	53.48±0	36.43±0	6.57±0	61.65±0	43.02±0	28.11±0	1.2±0.00
	E^2OMVC	88.57±0	38.1±0	38.1±0	37.14±0	79.35±0	23.7±0	26.17±0	20.96±0	75.51±0	12.77±0	14.19±0	74.7±0
	AGLSR	91.90±0	92.86±0	90.95±0	92.38±0	84.35±0	85.09±0	82.57±0	84.4±0	82.07±0	83.81±0	80.03±0	82.93±0
	SEC	89.05±0	91.43±0	87.62±0	73.81±0	79.34±0	82.82±0	75.99±0	57.48±0	76.25±0	80.47±0	73.66±0	49.55±0
	OAGL	**100±0**	**100±0**	**100±0**	**99.52±0**	**100±0**	**100±0**	**100±0**	**98.92±0**	**100±0**	**100±0**	**100±0**	**98.88±0**
BBC	SC	90.55±0.00	90.09±0.00	83.01±0.08	75.61±0.05	77.9±0.00	74.22±0.00	66.14±0.18	52.13±0.04	81.2±0.00	78.05±0.00	68.69±0.19	54.26±0.06
	CSRF	91.82±0.00	92.26±0.00	85.69±0.00	77.9±0.34	79.65±0.00	79.4±0.00	69.14±0.00	54.59±0.89	82.75±0.00	82.92±0.00	70.69±0.00	56.24±1.16
	GSR	98.69±0.00	98.39±0.00	97.81±0.00	94.74±0.00	96.38±0.00	95.46±0.00	94.26±0.00	85.25±0.00	96.9±0.00	96.26±0.00	95.1±0.00	88.11±0.00
	UOMSC	45.11±0	34.31±0	33.43±0	24.38±0	31.23±0	13.37±0	1.97±0	1.29±0	6.83±0	8.51±0	0.3±0	0.14±0
	E^2OMVC	85.11±0	57.81±0	47.45±0	37.23±0	71.93±0	35.41±0	21.54±0	14.41±0	72.56±0	33.96±0	19.95±0	5.73±0
	AGLSR	91.68±0	90.51±0	89.2±0	92.55±0	78.94±0	76.47±0	74.75±0	79.3±0	81.74±0	79.71±0	76.26±0	83.51±0
	SEC	89.2±0	94.6±0	91.39±0	80.15±0	73.54±0	84.36±0	76.05±0	56.2±0	77.61±0	88.5±0	81.21±0	60.09±0
	OAGL	**99.42±0**	**99.42±0**	**99.42±0**	**98.39±0**	**97.97±0**	**97.71±0**	**97.87±0**	**94.6±0**	**98.55±0**	**98.56±0**	**98.64±0**	**96.36±0**
Flower17	SC	59.44±0.86	57.8±0.62	54.45±0.27	43.85±0.78	56.42±0.45	54.65±0.32	48.26±0.22	40.5±0.18	41.05±0.63	35.95±0.46	31.88±0.28	22.51±0.48
	CSRF	61.06±0.35	59.85±0.67	55.96±0.24	45±1.19	60.54±0.47	58.47±0.23	52.49±0.14	43.07±0.28	41.46±0.38	36.62±0.54	32.92±0.34	24.15±0.69
	GSR	95.07±0.02	94.97±0.06	94.92±0.05	92.71±2.32	92.1±0.04	92.15±0.08	91.77±0.1	91.84±0.89	89.88±0.05	89.69±0.12	89.58±0.11	87.04±2.17
	UOMSC	54.93±0	20.44±0	11.47±0	11.54±0	56.24±0	19.64±0	4.96±0	4.58±0	31.87±0	10.26±0	3.6±0	3.2±0
	E^2OMVC	61.03±0	25.66±0	21.03±0	16.18±0	59.69±0	25.78±0	23.86±0	19.66±0	37.74±0	9.22±0	6.21±0	2.54±0
	AGLSR	55.22±0	55.88±0	52.35±0	55.51±0	53.52±0	53.86±0	53.68±0	52.64±0	36.27±0	37.96±0	36.82±0	36.23±0
	SEC	51.03±0	59.56±0	53.31±0	41.03±0	49.84±0	57.43±0	48.81±0	37.36±0	33.11±0	42.21±0	34.48±0	23.64±0
	OAGL	**96.03±0**	**96.03±0**	**96.32±0**	**95.22±0**	**94.54±0**	**94.4±0**	**94.88±0**	**93.67±0**	**91.89±0**	**91.84±0**	**92.52±0**	**90.51±0**
Handwritten	SC	85.01±0.02	87.3±0.00	86.38±0.04	83.75±0.02	84.77±0.03	87.17±0.00	85.35±0.03	76.48±0.02	78.3±0.04	82.44±0.00	80.87±0.05	74.99±0.03
	CSRF	97.75±0.00	97.65±0.00	95.55±0.00	88.71±0.02	94.77±0.00	94.38±0.00	90.39±0.00	78.16±0.03	95.07±0.00	94.84±0.00	90.26±0.00	76.32±0.03
	GSR	99.9±0.00	99.9±0.00	99.35±0.00	99.35±0.00	99.73±0.00	99.73±0.00	98.47±0.00	98.28±0.00	99.78±0.00	99.78±0.00	98.56±0.00	98.56±0.00
	UOMSC	98.1±0	79.2±0	45.7±0	32.05±0	95.61±0	86.91±0	52.86±0	25.69±0	95.78±0	76.33±0	39±0	8.34±0
	E^2OMVC	97.2±0	44.65±0	29.15±0	22.95±0	93.71±0	35.75±0	24.4±0	14.45±0	93.9±0	23.07±0	12.19±0	3.27±0
	AGLSR	88.3±0	88.1±0	87.7±0	87.75±0	89.59±0	90.38±0	88.83±0	88.3±0	84.77±0	85.04±0	83.33±0	83.46±0
	SEC	97.25±0	98.65±0	96.2±0	89.9±0	93.73±0	96.97±0	92.01±0	79.91±0	94±0	97.03±0	91.79±0	79.03±0
	OAGL	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**	**100±0**
NUSwide	SC	29.7±0.13	27.03±0.11	24.48±0.16	21.36±0.18	16.65±0.1	14.34±0.07	12.47±0.05	10.2±0.23	8.47±0.05	7.69±0.06	5.16±0.06	4.88±0.18
	CSRF	31.7±0.19	28.62±0.02	26.17±0.33	23.91±0.09	19.43±0.54	17.56±0.02	14.02±0.28	11.5±0.17	10.48±0.31	8.2±0.01	5.84±0.36	5.32±0.1
	GSR	92.09±0.07	91.68±0.08	85.21±0.11	68.49±2.9	85.51±0.11	84.75±0.1	79.03±0.16	73.47±1.00	84.06±0.14	83.49±0.14	74.48±0.21	58.19±2.07
	UOMSC	24.08±0	24.38±0	20.04±0	18.46±0	13.84±0	15.06±0	10.52±0	6.87±0	4.88±0	5.98±0	3.33±0	2.68±0
	E^2OMVC	29.88±0	19.33±0	16.75±0	14.71±0	19.14±0	11.43±0	7.79±0	8.42±0	8.95±0	3.39±0	2.17±0	1.2±0.00
	AGLSR	28.75±0	30.83±0	29.96±0	28.96±0	16±0	15.23±0	16.55±0	17.49±0	9.61±0	10.5±0	9.39±0	9.71±0
	SEC	32.08±0	35.04±0	28.67±0	22.63±0	16.52±0	19.35±0	14.43±0	9.25±0	11.26±0	15.51±0	11.28±0	6.44±0
	OAGL	**95.88±0**	**94.08±0**	**91.08±0**	**85.33±0**	**93.06±0**	**89.06±0**	**84.39±0**	**75.71±0**	**91.14±0**	**87.79±0**	**81.81±0**	**71.7±0**
O-Scene	SC	65.51±0	64.29±0	60.57±0	53.85±0	55.17±0	50.02±0	42.92±0	34.67±0	45.11±0	37.42±0	31.85±0	23.92±0
	CSRF	71.69±0.01	66.6±0.02	62.39±0	55.61±0.02	57.51±0.02	52.19±0	44.06±0.01	36.37±0.01	47.51±0.02	40.08±0.02	36.28±0.01	26.64±0.01
	GSR	97.28±0	96.84±0	95.46±0	95.24±0	93.66±0	92.56±0	89.66±0	89±0	93.63±0	92.66±0	89.58±0	89.38±0
	UOMSC	70.99±0	40.66±0	23.88±0	19.23±0	61.99±0	38.54±0	18.92±0	6.05±0	49.69±0	30.77±0	13.05±0	2.04±0
	E^2OMVC	62.91±0	37.87±0	31.62±0	24.81±0	52.95±0	26.95±0	20.79±0	13.19±0	40.2±0	18.28±0	9.41±0	2.26±0
	AGLSR	66.67±0	67.15±0	63.1±0	66.26±0	53.29±0	54.32±0	55.77±0	56.96±0	42.47±0	44.33±0	45.39±0	45.87±0
	SEC	66.48±0	68.45±0	56.18±0	47.62±0	52.34±0	55.41±0	36.7±0	26.51±0	47.19±0	49.59±0	30.84±0	21.08±0
	OAGL	**97.77±0**	**97.66±0**	**97.21±0**	**95.42±0**	**94.99±0**	**95.12±0**	**93.87±0**	**90.74±0**	**95.23±0**	**95.07±0**	**94±0**	**90.28±0**
Caltech101	SC	27.38±1.18	25.49±1	24.54±0.81	23.13±0.76	49.33±0.27	47.11±0.2	46.25±0.19	42.67±0.18	18.67±2.27	17.2±1.71	17±1.67	14.55±0.92
	CSRF	28.81±0.74	28.24±0.58	27.87±0.63	24.87±0.79	50.59±0.23	50.12±0.27	48.63±0.22	42.76±0.25	19.54±1.29	19.72±1.31	18.36±1.21	16.79±1.14
	GSR	-	-	-	-	-	-	-	-	-	-	-	-
	UOMSC	21.52±0	20.89±0	14.53±0	10.03±0	48.14±0	43.58±0	32.58±0	23.97±0	4.89±0	5.59±0	3.85±0	2.27±0
	E^2OMVC	28.42±0	10.73±0	11.23±0	10.11±0	49.34±0	19.61±0	18.94±0	17.28±0	8.14±0	0.95±0	1.18±0	0.65±0
	AGLSR	32.11±0	31.14±0	29.6±0	28.12±0	48.8±0	49.86±0	48.67±0	49.04±0	16.21±0	22.54±0	21.4±0	19.39±0
	SEC	28.58±0	27.79±0	27.37±0	17.53±0	50.29±0	48.63±0	46.4±0	35.11±0	19.33±0	19.67±0	19.7±0	10.21±0
	OAGL	**47.22±0**	**50.78±0**	**47.67±0**	**46.41±0**	**82.56±0**	**82.65±0**	**83.9±0**	**76.47±0**	**31.39±0**	**35.04±0**	**37.52±0**	**25.34±0**

4) Parameter Settings: The proposed OAGL method has two parameters in (15), i.e., α and β. The values of the parameters were selected from $\{1e^{-3}, 5e^{-3}, 0.01, 0.05\}$ using a grid search strategy. We assumed that the true number of clusters is known for each dataset. For the competing methods involving the k-means algorithm, we repeated each experiment 10 times and reported the mean values and standard deviations. For a fair comparison, we reported the best clustering results achieved by tuning the parameters of these methods. The best and second-best clustering results are highlighted in bold and underlined, respectively.

We examined two specific scenarios for the multiview datasets. In the first scenario, all instances were available in each view. In the second scenario, we randomly removed varying percentages of instances from each view: 10%, 30%, and 50%. The ANGL method requires a parameter that defines the number of nearest neighbors for each instance, which is chosen from the ranges of $\{5, 10, 15, 20, 25\}$.

B. Performance Evaluation

The experimental results of all the competing methods are presented in Table II, where the symbol '-' indicates an out-of-memory issue for the corresponding algorithm. The proposed OAGL method consistently achieves the best clustering performance for all the datasets. For example, our method outperforms the second-best results by nearly 3.79%, 7.55% and 7.08% in

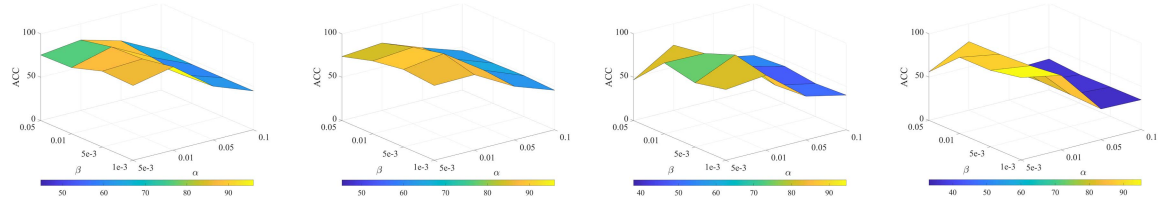

Fig. 1. The ACC of the OAGL method with different combinations of α and β on the Flower17 dataset for various missing ratios of instances: (a) $ratio = 0$, (b) $ratio = 10\%$, (c) $ratio = 30\%$, and (d) $ratio = 50\%$.

terms of the ACC, NMI and ARI, respectively, on the NUS-WIDE dataset with a missing data ratio of 0. Similarly, OAGL exhibits significant improvements in clustering performance for the other datasets. Therefore, this demonstrates the effectiveness and advantages of the OAGL for IMVC.

The clustering performance of all competing methods generally decreases as the missing data ratio gradually increases from 0 to 50%. The proposed OAGL method consistently outperforms the other competing methods. Compared with GSR, our method achieves improvements of approximately 3.79%, 3.4%, 5.87% and 16.84% on the NUS-WIDE dataset with missing data ratios of 0, 10%, 30% and 50%, respectively. The gap in the clustering performance between OAGL and GSR increases for higher missing data ratios, e.g., 30% and 50%, respectively. Therefore, this shows the robustness of the proposed OAGL method.

GSR achieves the second-best clustering results on all the datasets except the Caltech101 dataset. Unfortunately, this method suffers from an out-of-memory issue on the Caltech101 dataset. CSRF shows competitive clustering results on the datasets. The four one-step IMVC methods obtain comparable results with the complete multiview datasets. In addition, the clustering performance of UOMSC, E^{2}OMVC and SEC tends to deteriorate in most cases as the missing data ratio increases from 10% to 50% in the multiview datasets. In contrast, the clustering performance of AGLSR often remains relatively stable on several datasets, e.g., the MSRC-v1, BBC and Flower17 datasets, as the missing data ratio increases. This shows its robustness against incomplete views.

As a one-step IMVC method, OAGL consistently outperforms other one-step IMVC approaches, such as AGLSR, E^{2}OMVC, SEC, and UOMSC. All of these methods perform a spectral rotation operation on the spectral embedding matrix to obtain clustering results in a single step. In contrast, they employ different techniques to examine correlations across multiple views. OAGL introduces low-rank tensor learning to align spectral block-diagonal matrices, which can effectively reveal a unified low-rank structure within high-dimensional multiview data.

The advantages and effectiveness of the proposed OAGL method can be attributed to three primary reasons. First, spectral embedding fusion is adaptively performed to achieve clustering assignments in a one-step manner. This effectively mitigates the detrimental impact of noise and eliminates redundant information from multiple original views when contrasted with conducting information fusion at the original data level. Second, the sparsity constraint on an affinity matrix is introduced to improve

TABLE III
COMPUTATION TIMES (IN SECONDS) OF DIFFERENT METHODS ON SEVEN MULTIVIEW DATASETS WITH VARYING MISSING RATIOS OF INSTANCES

Datasets	Ratio	CSRF	GSR	UOMSC	E^{2}OMVC	AGLSR	SEC	OAGL
MSRC-v1	0	**0.09**	1.56	0.11	0.72	8.01	0.32	1.43
	10%	**0.12**	1.18	0.74	0.3	5.39	0.17	1.27
	30%	0.12	1.11	0.72	0.24	8.04	**0.11**	1.14
	50%	0.1	1.1	0.72	0.23	7.08	**0.08**	1.48
BBC	0	1.96	4.5	1.82	6.01	247	**0.69**	11.5
	10%	1.53	4.7	2.29	4.55	245	**0.57**	9.1
	30%	1.3	6.9	2.11	3.41	238	**0.62**	8.8
	50%	0.82	12.2	1.95	2.31	237	**0.63**	11.1
Flower17	0	**5.7**	56.9	5.9	274	92.1	9.3	68.8
	10%	**4.7**	55.3	9.1	34.7	634	10.4	68.2
	30%	**4.4**	50.6	8.8	30.9	542	5.9	59
	50%	5.9	50.9	10.1	31.6	548	**5.3**	62.7
Handwritten	0	19.4	134	**14.3**	23.6	861	27.9	132
	10%	19.1	73.9	32.4	**18.5**	789	28.9	123
	30%	20.6	77.2	36.9	**17.91**	484	30.9	118
	50%	13.1	72.1	37.9	**7**	247	32.1	105
NUSwide	0	27.2	122	25.9	**6.9**	389	29.1	185
	10%	26.3	133	67.4	**6.5**	388	19.4	189
	30%	24.6	115	53.8	**4**	390	17.9	183
	50%	22.4	101	62.6	**4.9**	387	16.2	184
O-Scene	0	24.1	151	14.3	**2.74**	355	11.2	168
	10%	25.7	136	59.4	**6.41**	354	9.92	161
	30%	22.6	128	72.4	**4.83**	351	12.8	163
	50%	26.5	140	71	**3.25**	355	10.8	145
Caltech101	0	854	-	627	**224**	8,722	559	9,902
	10%	784	-	1,948	**200**	9,001	375	9,159
	30%	1,057	-	2,062	**153**	9,254	310	9,186
	50%	1,200	-	3,086	**114**	9,802	286	7,894

the quality of the similarity graph corresponding to each incomplete view. The spectral block-diagonal matrices are constructed from the corresponding similarity graphs and projected into a union of lower-dimensional spaces. As a result, they preserve the intrinsic structures embedded in the incomplete views. Finally, an alignment of the spectral block-diagonal matrices among incomplete multiple views integrates the information from the incomplete views in a coherent manner, which is conducive to discovering the low-rank structures implied in similarity graphs.

To illustrate the efficiency of the proposed OAGL method, we record the average running times of all the competing state-of-the-art algorithms on the six multiview datasets with varying missing data ratios. These average running times are presented in Table III. CSRF, UOMSC, E^{2}OMVC and SEC show the advantage of running time on all the datasets. The running times of AGLSR are higher than those of the other competing methods in most cases. Compared with GSR, OAGL obtains comparable running times on all the multiview datasets except for the Caltech101 dataset. In addition, the running times of the proposed OAGL method are higher than that of most competing

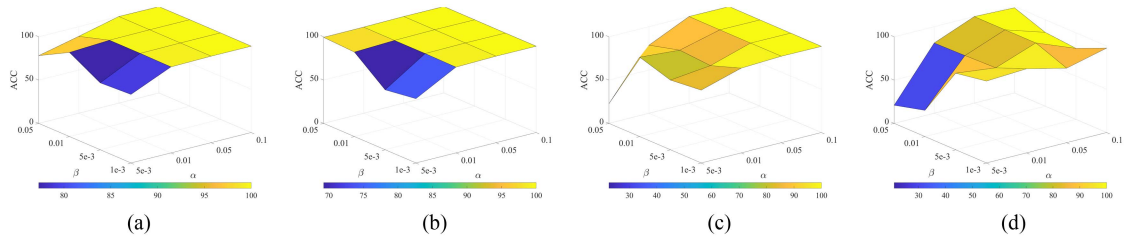

Fig. 2. The ACC of the OAGL method with different combinations of α and β on the Handwritten dataset for various missing ratios of instances: (a) $ratio = 0$, (b) $ratio = 10\%$, (c) $ratio = 30\%$, and (d) $ratio = 50\%$.

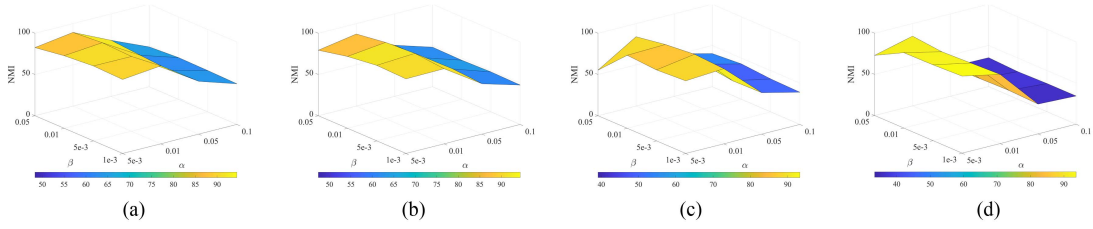

Fig. 3. The NMI of the OAGL method with different combinations of α and β on the Flower17 dataset for varying missing ratios of instances: (a) $ratio = 0$, (b) $ratio = 10\%$, (c) $ratio = 30\%$, and (d) $ratio = 50\%$.

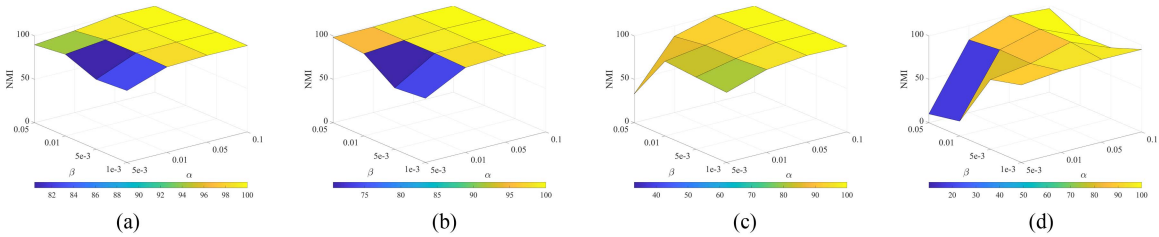

Fig. 4. The NMI of the OAGL method with different combinations of α and β on the Handwritten dataset for varying missing ratios of instances: (a) $ratio = 0$, (b) $ratio = 10\%$, (c) $ratio = 30\%$, and (d) $ratio = 50\%$.

methods. This is because OAGL involves low-rank tensor learning for the alignment of spectral block-diagonal matrices, which requires SVD during iterations. This leads to a relatively high computational cost. However, OAGL effectively explore correlations across multiple views using low-rank tensor learning. With the rapid advancement of computer hardware technology, computational cost may become a secondary concern.

C. Parameter Sensitivity Analysis

Due to space limitations, we utilize two multiview datasets to evaluate the parameter sensitivity of the OAGL, i.e., the Flower17 and Handwritten datasets. The proposed OAGL method has two parameters in (15), i.e., α and β. Additionally, the ANGL method requires a parameter, i.e., the number of nearest neighbor k.

The parameters α and β vary from $\{5e^{-3}, 0.01, 0.05, 0.1\}$ and $\{1e^{-3}, 5e^{-3}, 0.01, 0.05\}$, respectively. The number of nearest neighbors in the ANGL method is identical to that provided in Section IV-B. Figs. 1–4 show the clustering performances of OAGL on the Flower17 and Handitten datasets with different combinations of the two parameters α and β, respectively. We

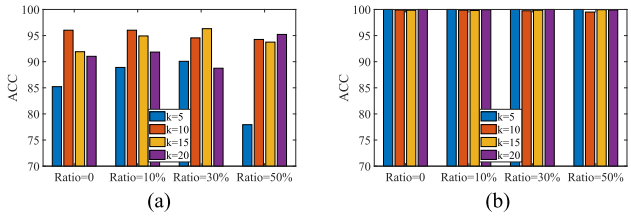

Fig. 5. Comparison of the ACC values of the OAGL method with different numbers of nearest neighbors k on the datasets: (a) Flower17 and (b) Handwritten.

can observe that OAGL obtains stable clustering results on the different datasets with relatively wide ranges of α and β. For example, OAGL can achieve satisfactory clustering results when α and β range from 0.01 to 0.1 and $1e^{-3}$ to 0.05 on the Handwritten dataset, respectively. In addition, α is more sensitive than β in the experiments. This is because the gap between each individual affinity matrix and the corresponding spectral block-diagonal matrix varies significantly across different datasets.

The number of nearest neighbors k in the ANGL method is chosen from $\{5, 10, 15, 20\}$. As shown in Fig. 5, the proposed

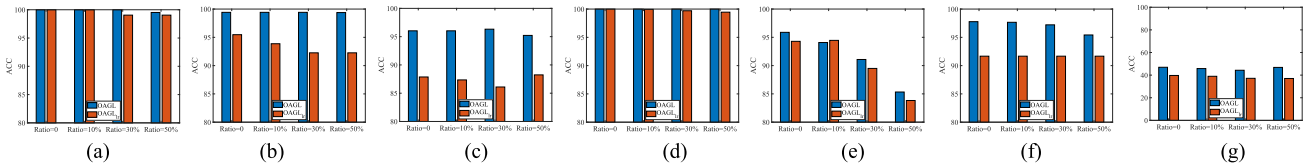

Fig. 6. Comparison of the ACC values of the OAGL method with those of OAGL$_{\mathbf{lr}}$ on all the datasets: (a) MSRCv1, (b) BBC, (c) Flower17, (d) Handwritten, (e) NUSwide, (f) O-Scene, and (g) Caltech101.

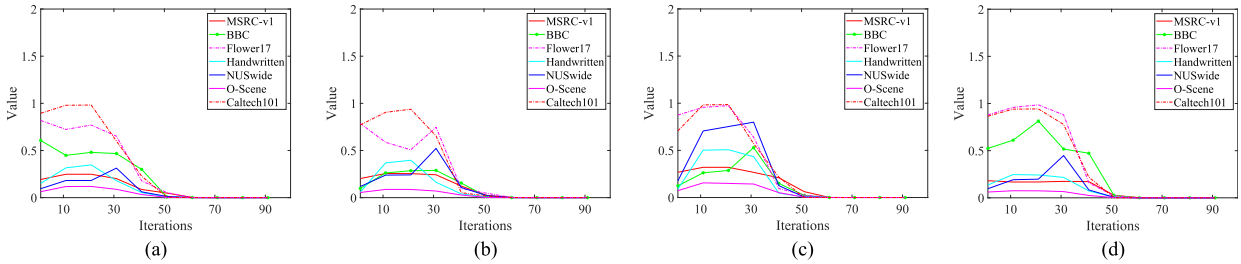

Fig. 7. Convergence results of the OAGL method on all datasets with varying missing ratios of instances: (a) $ratio = 0$, (b) $ratio = 10\%$, (c) $ratio = 30\%$, and (d) $ratio = 50\%$.

OAGL method achieves high ACC values across different k values on both the Flower17 and Handwritten datasets. The ACC remains relatively stable for $k = 10$ and $k = 15$. In contrast, the ACC is slightly degraded when $k = 5$ or $k = 20$. Additionally, OAGL consistently achieves satisfactory ACC values as the missing data ratio in each dataset gradually increases from 0 to 50%. These observations further demonstrate the robustness of the proposed OALG method in constructing the affinity matrices from similarity graphs.

D. Ablation Study

In this section, we empirically investigate low-rank structures implied in similarity graphs. In (15), $\mathbf{W}^{(v)}$ exhibits an approximate block-diagonal structure, whereas $\mathbf{H}^{(v)}\,(\mathbf{H}^{(v)})^T$ contains low-rank structures. Low-rank representation is a crucial subspace learning technique for exploring low-rank structures in high-dimensional data. We employed an adaptive low-rank (LR) representation method [51] to construct affinity matrices corresponding to incomplete multiple views. Thus, this variant is regarded as OAGL$_{\mathbf{lr}}$.

According to Fig. 6, the proposed OAGL method outperforms OAGL$_{\mathbf{lr}}$ on all the datasets in terms of ACC. OAGL and OAGL$_{\mathbf{lr}}$ obtained comparable ACCs on the MSRCv1, BBC, Handwritten and NUS-WIDE datasets, with a missing ratio of 0. This is because the affinity matrices produced by OAGL$_{\mathbf{lr}}$ also exhibit an approximate block-diagonal structure. OAGL generally maintains stable clustering performance as the missing ratio of the instances gradually increases. In contrast, the clustering performance of OAGL$_{\mathbf{lr}}$ does not remain stable. OAGL$_{\mathbf{lr}}$ cannot capture the intrinsic structure of incomplete multiple views because the missing instances are removed from incomplete multiple views. This illustrates the need for sparse representation in constructing affinity matrices in OAGL.

E. Convergence Analysis

To validate the convergence property of the proposed OAGL method, we conduct experiments on all the datasets with varying missing ratios of the instances. The parameters of the OAGL follow the settings in Section IV-B. Fig. 7 shows the curves representing the values of the convergence condition $\|\mathcal{G} - \mathcal{T}\|_{\max}$ of OAGL during the iterations. The convergence condition values of the proposed methods generally fluctuate slightly during the first 30 iterations, and then decrease slowly until convergence. Moreover, the number of iterations is less than 110 for all the datasets. This indicates that the proposed OAGL method can achieve a local optional solution during iterations. Consequently, Algorithm 2 is effective for solving (15).

V. Conclusion

In this paper, we propose the OAGL method for IMVC, which can adaptively perform spectral embedding fusion to achieve clustering assignments at the clustering indicator level. We explore complementary and consistency information across incomplete multiple views from two perspectives. On the one hand, we employ sparse representation on the existing instances with incomplete multiple views to initiate each affinity matrix, and ensure that the degree matrix of the affinity matrix approximates an identity matrix. On the other hand, we exploit low-rank tensor learning theory to perform an alignment of the spectral block-diagonal matrices among incomplete multiple views. Compared with the existing graph-based IMVC methods, we consider the low-rank structures implied in similarity graphs, which can alleviate the negative impact of redundant information contained in incomplete multiview data. Additionally, an effective alternating iterative algorithm is presented to solve the resulting optimization problem. The experimental results on benchmark datasets demonstrate the advantages and effectiveness of the proposed OAGL method.

REFERENCES

[1] J. Wen et al., "A survey on incomplete multiview clustering," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 53, no. 2, pp. 1136–1149, Feb. 2023.

[2] C. Liu, Z. Wu, J. Wen, Y. Xu, and C. Huang, "Localized sparse incomplete multi-view clustering," *IEEE Trans. Multimedia.*, vol. 32, pp. 5539–5551, 2023.

[3] Z. Tao, H. Liu, H. Fu, and Y. Fu, "Multi-view saliency-guided clustering for image cosegmentation," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4634–4645, Sep. 2019.

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

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

[6] S. Deng, J. Wen, C. Liu, K. Yan, G. Xu, and Y. Xu, "Projective incomplete multi-view clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 8, pp. 10539–10551, Aug. 2024.

[7] S. Luo and X. Cao, "Multiview subspace dual clustering," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7425–7437, Dec. 2022.

[8] Z. Tao, J. Li, H. Fu, Y. Kong, and Y. Fu, "From ensemble clustering to subspace clustering: Cluster structure encoding," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 5, pp. 2670–2681, May 2023.

[9] Z. Yang, Q. Xu, W. Zhang, X. Cao, and Q. Huang, "Split multiplicative multi-view subspace clustering," *IEEE Trans. Image Process.*, vol. 28, no. 10, pp. 5147–5160, Oct. 2019.

[10] Z. Tao, H. Liu, S. Li, Z. Ding, and Y. Fu, "Marginalized multiview ensemble clustering," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 31, no. 2, pp. 600–611, Feb. 2020.

[11] J. Chen, H. Mao, D. Peng, C. Zhang, and X. Peng, "Multiview clustering by consensus spectral rotation fusion," *IEEE Trans. Image Process.*, vol. 32, pp. 5153–5166, Sep. 2023.

[12] J. Chen, Z. Wang, H. Mao, and X. Peng, "Low-rank tensor learning for incomplete multiview clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 11, pp. 11 556–11 569, Nov. 2023.

[13] J. Wen et al., "Adaptive graph completion based incomplete multi-view clustering," *IEEE Trans. Multimedia*, vol. 23, pp. 2493–2504, 2021.

[14] Y. Xie, W. Zhang, Y. Qu, L. Dai, and D. Tao, "Hyper-laplacian regularized multilinear multiview self-representations for clustering and semisupervised learning," *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 572–586, Feb. 2020.

[15] Z. Huang, J. T. Zhou, X. Peng, C. Zhang, H. Zhu, and J. Lv, "Multi-view spectral clustering network," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Macao, China, 2019, pp. 2563–2569.

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

[17] Y. Tang, Y. Xie, X. Yang, J. Niu, and W. Zhang, "Tensor multi-elastic kernel self-paced learning for time series clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 3, pp. 1223–1237, Mar. 2021.

[18] Y. Xie et al., "Robust kernelized multiview self-representation for subspace clustering," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 32, no. 2, pp. 868–881, Feb. 2020.

[19] Q. Wang, Z. Ding, Z. Tao, Q. Gao, and Y. Fu, "Partial multi-view clustering via consistent GAN," in *Proc. IEEE Int. Conf. Data Mining*, Singapore, 2018, pp. 1290–1295.

[20] J. Cheng, Q. Wang, Z. Tao, D. Xie, and Q. Gao, "Multi-view attribute graph convolution networks for clustering," in *Proc. 30th Int. Joint Conf. Artif. Intell.*, Montreal, Canada, 2021, pp. 2973–2979.

[21] J. Chen, H. Mao, W. L. Woo, and X. Peng, "Deep multiview clustering by contrasting cluster assignments," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Paris, France, 2023, pp. 16 752–16 761.

[22] Y. Xie et al., "Joint deep multi-view learning for image clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 11, pp. 3594–3606, May 2021.

[23] Y. Lin, Y. Gou, X. Liu, J. Bai, J. Lv, and X. Peng, "Dual contrastive prediction for incomplete multi-view representation learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4447–4461, Aug. 2012.

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

[25] J. Yin and S. Sun, "Incomplete multi-view clustering with reconstructed views," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 3, pp. 2671–2682, Mar. 2023.

[26] J. Wen et al., "Highly confident local structure based consensus graph learning for incomplete multi-view clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Vancouver, Canada, 2023, pp. 15 712–15 721.

[27] X. Zhu, S. Zhang, W. He, R. Hu, C. Lei, and P. Zhu, "One-step multi-view spectral clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 10, pp. 2022–2034, Oct. 2019.

[28] C. Tang, M. Wang, and K. Sun, "One-step multiview clustering via adaptive graph learning and spectral rotation," *IEEE Trans. Neural. Netw. Learn. Syst.*, early access, Apr. 10, 2024, doi: 10.1109/TNNLS.2024.3381223.

[29] J. Wang, C. Tang, Z. Wan, K. Sun, and A. Y. Zomaya, "Efficient and effective one-step multiview clustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 9, pp. 12224–12235, Sep. 2024.

[30] C. Zhang, J. Wei, B. Wang, Z. Li, C. Chen, and H. Li, "Robust spectral embedding completion based incomplete multi-view clustering," in *Proc. 31st ACM Int. Conf. Multimedia.*, Ottawa, Canada, 2023, pp. 300–308.

[31] P. Zhang, X. Li, S. Zhou, W. Zhao, E. Zhu, and Z. Cai, "Consensus one-step multi-view subspace clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 10, pp. 4676–4689, Oct. 2022.

[32] Y. Tang, Y. Xie, C. Zhang, Z. Zhang, and W. Zhang, "One-step multiview subspace segmentation via joint skinny tensor learning and latent clustering," *IEEE Trans. Cybern.*, vol. 52, no. 9, pp. 9179–9193, Sep. 2022.

[33] J. Shi, J. Malik, and S. Sastry, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 181–214, Aug. 2000.

[34] J. Huang, F. Nie, and H. Huang, "Spectral rotation versus k-means in spectral clustering," in *Proc. AAAI Conf. Artif. Intell.*, Bellevue, Washington, USA, 2013, pp. 431–437.

[35] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Its Appl.*, vol. 435, no. 3, pp. 641–658, Aug. 2011.

[36] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, Apr. 2020.

[37] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, Apr. 1995.

[38] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 629–654, Sep. 2008.

[39] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution," *Comm. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, Mar. 2006.

[40] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Mar. 2011.

[41] X. Chen, F. Nie, and M. Y. J. Z. Huang, "Scalable normalized cut with improved spectral rotation," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Melbourne, Australia, 2017, pp. 1518–1524.

[42] F. Nie, R. Zhang, and X. Li, "A generalized power iteration method for solving quadratic problem on the Stiefel manifold," *Sci. China Inf. Sci.*, vol. 60, pp. 1–10, May 2017.

[43] D. Greene and P. Cunningham, "Producing accurate interpretable clusters from high-dimensional data," in *Proc. Conf. Knowl. Discov. Databases*, Porto, Portugal, 2005, pp. 486–494.

[44] M. E. Nilsback and A. Zisserman, "A visual vocabulary for flower classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New York, USA, 2006, pp. 1447–1454.

[45] A. Asuncion and D. Newman, "UCI machine learning repository," University of California, Irvine, School of Information and Computer Sciences, 2007. [Online]. Available: http://archive.ics.uci.edu/ml/

[46] T. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng, "NUS-WIDE: A real-world web image database from national university of singapore," in *Proc. ACM Conf. Image Video Retrieval*, Santorini, Greece, 2009, pp. 486–494.

[47] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vis.*, vol. 42, no. 3, pp. 145–175, May 2001.

[48] D. Dua and C. Graff, "UCI machine learning repository," University of California, Irvine, School of Information and Computer Sciences, 2017. [Online]. Available: http://archive.ics.uci.edu/ml

[49] X. Li, M. Chen, C. Wang, and J. Lai, "Refining graph structure for incomplete multi-view clustering," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 35, no. 2, pp. 2300–2313, Feb. 2024.

[50] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, New York, USA, 2014, pp. 977–986.

[51] J. Chen, H. Mao, Z. Wang, and X. Zhang, "Low-rank representation with adaptive dictionary learning for subspace clustering," *Knowl. Based Syst.*, vol. 223, pp. 1–12, Jul. 2021.
[52] C. D. Manning, P. Raghavan, and H. Schútze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

Jie Chen (Member, IEEE) received the BSc degree in software engineering, the MSc and PhD degrees in computer science from Sichuan University, Chengdu, China, in 2005, 2008 and 2014, respectively. From 2008 to 2009, he was with Huawei Technologies Co., Ltd. as a software engineer. He is currently an associate professor with the College of Computer Science, Sichuan University, China. His current research interests include machine learning, Big Data analysis, and deep neural networks.

Hua Mao received the BS and MS degrees in computer science from the University of Electronic Science and Technology of China (UESTC) in 2006 and 2009, respectively and the PhD degree in computer science and engineering from Aalborg University, Denmark in 2013. She is currently a senior lecturer in Department of Computer and Information Sciences, Northumbria University, U.K. Her current research interests include deep neural networks and Big Data.

Wai Lok Woo (Senior Member, IEEE) received the BEng degree in electrical and electronics engineering and the MSc and PhD degrees in statistical machine learning from Newcastle University, U.K., respectively. He is currently a professor of Machine Learning with Northumbria University, U.K. Previously, he was the director of Research for the Newcastle Research and Innovation Institute, and the director of Operations for Newcastle University, U.K. His research interests include the mathematical theory and algorithms of machine learning and applications to deep learning, information processing, anomaly detection, digital health, and digital sustainability. His research is supported by the U.K. Research and Innovation.

Chuanbin Liu is currently an assistant research fellow with Center for Scientific Research and Development in Higher Education Institutes, Ministry of Education, China (CSRD). His research interests include policy evaluation and peer review. He has published in peer-reviewed international journals, including the *Information Sciences*, *Science China* (Technological Sciences), and *Journal of Cleaner Production*.

Zhu Wang received the BM, LLM, and LLD degrees in civil and commercial law from the Renmin University of China, China in 2003, 2006, and 2009, respectively. He is currently a professor with the Law of Sichuan University and Director of Key Laboratory of Sichuan Province for AI Empowered Governance in Smart Society (AegisS). His research interests are mainly in civil and commercial law and AI in law.

Xi Peng (Senior Member, IEEE) is currently a full professor with the College of Computer Science, Sichuan University. His current research interest includes machine intelligence and has authored more than 50 articles in these areas. He has served as an associate editor/guest editor for six journals, including *IEEE Transactions on SMC: Systems* and *IEEE Transactions on Neural Network And Learning Systems* and the area chair/senior program committee member for the conferences such as IJCAI, AAAI, and ICME.