Self-Completed Bipartite Graph Learning for Fast Incomplete Multi-View Clustering

Xiaojia Zhao, Qiangqiang Shen, Yongyong Chen, *Member, IEEE*, Yongsheng Liang, Junxin Chen, *Senior Member, IEEE*, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*— Incomplete multi-view clustering (IMVC), excavating diversity and consistency from multiple incomplete views, has aroused widespread research enthusiasm. Nevertheless, most existing methods still encounter the following issues: 1) they generally concentrate on pair-wise instance correlation, which consumes at least a quadratic complexity and precludes them from applying at large scales; 2) they only concentrate on pair-wise instance relevance, whereas ignoring the discriminative correlation hidden across views. To overcome these drawbacks, we propose the Self-Completed Bipartite Graph Learning (SCBGL) method for fast IMVC, which adaptively learns a self-completed consensus bipartite graph with the guidance of global information. Specifically, SCBGL learns the consensus anchor matrix shared among diverse views and further constructs a consensus intra-view bipartite graph with missing instances to explore the diversity and complementarity underlying different views. Meanwhile, we concatenate all the multiple features with projection learning to learn global anchors that would be employed to construct an inter-view bipartite graph. Furthermore, SCBGL dexterously utilizes the abundant inter-view information to tutor the self-completion of the consensus intra-view bipartite graph. By devising an alternatively iterative strategy, we present an efficient algorithm, which enjoys a linear time complexity, to solve the proposed SCBGL model. Numerous experiments conducted on large-scale datasets substantiate the superior performance of the SCBGL beyond the state-of-the-arts.**

***Index Terms*— Incomplete multi-view clustering, bipartite graph learning, graph self-completion.**

Manuscript received 2 November 2022; revised 11 March 2023 and 2 June 2023; accepted 1 August 2023. Date of publication 7 August 2023; date of current version 5 April 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62106063, in part by the Guangdong Natural Science Foundation under Grant 2022A1515010819, in part by the Shenzhen Science and Technology Program under Grant RCBS20210609103708013, in part by the Humanities and Social Sciences Foundation of the Ministry of Education of China under Grant 22YJC630129, and in part by the Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies under Grant 2022B1212010005. This article was recommended by Associate Editor X. Zhang. *(Xiaojia Zhao and Qiangqiang Shen contributed equally to this work.) (Corresponding author: Yongyong Chen.)*

Xiaojia Zhao is with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong 518055, China.

Qiangqiang Shen and Yongsheng Liang are with the School of Electronics and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong 518055, China (e-mail: 1120810623@hit.edu.cn).

Yongyong Chen is with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong 518055, China, and also with the Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies, Shenzhen 518055, China (e-mail: YongyongChen.cn@gmail.com).

Junxin Chen is with the School of Software, Dalian University of Technology, Dalian 116621, China (e-mail: junxinchen@ieee.org).

Yicong Zhou is with the Department of Computer and Information Science, University of Macau, Macau, China (e-mail: yicongzhou@um.edu.mo).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCSVT.2023.3302326.

Digital Object Identifier 10.1109/TCSVT.2023.3302326

I. Introduction

MULTI-VIEW data, which could comprehensively describe data from multiple perspectives or heterogeneous features [1], are conducive to various downstream applications, such as clustering [2], [3], [4], [5], [6], human pose recovery [7], and image reranking [8]. Recently, tremendous research efforts are devoted to multi-view clustering (MVC) [9], [10], [11], [12], [13]. The essence of MVC is to fully excavate the inherent consistency and complementarity embedded in multi-view data by achieving one discriminative similarity matrix to explore the pair-wise correlation between two samples [14]. For example, Elhamifar and Vidal [13] and Liu et al. [12] imposed sparse and low-rank constraints on the similarity matrix, respectively, elegantly modeling the local and global underlying structure of the data. The work in [15] sought the lowest-rank representation by imposing low-rank constraints on the self-representation matrix. The work in [16] utilized a representation tensor to construct the similarity matrix and affinity matrix in one step. To handle the complex noises, Xing et al. [17] adopted kernel risk-sensitive loss to learn a low-rank similarity matrix.

A normal hypothesis for these above approaches is that all the views of multi-view data are complete, disregarding an inconspicuous but inevitable situation that incomplete multi-view data are ubiquitous in real-world scenarios [18]. For example, in the diagnosis of multimodal Alzheimer's disease, omissive disease tests would result in partially-available information due to economic reasons. Another example is that sensor fault would cause some missing features of the corresponding views. This incompleteness leads to a plummet in the performance of MVC or even execution failure [19], which makes incomplete multi-view clustering (IMVC) an extremely challenging task.

IMVC strives to take full advantage of the underlying information concealed in incomplete data to reveal the intrinsic distributions of the data. Recently proposed IMVC methods can be categorized into three strategies, namely, matrix factorization (MF)-based methods [20], [21], graph or kernel construction methods [22], [23], [24] and deep learning

methods [25], [26]. MF-based approaches focus on establishing a consistent low-dimensional representation across views. Representatively, PVC [20] incorporated the non-negative matrix factorization (NMF) technique to achieve a consensus latent subspace. Unfortunately, PVC only focuses on two-view data. For this consideration, Shao et al. [21] employed a co-regularized approach with $l_{2,1}$ regularizer to tackle arbitrary incomplete views. The second category of methods, striving to recover a unified similarity graph or kernel matrix, reveal the correlations of incomplete multi-view data. For example, Rai et al. [22] recovered the incomplete kernel matrix and then performed the kernel canonical correlation analysis for subsequent clustering. However, this approach demands at least one complete view, which limits its scalability and flexibility to handle complicated incomplete dataset. Furthermore, Wen et al. [24] integrated graph construction and consensus representation to exploit the geometric structures of incomplete data. By virtue of the powerful representation learning capabilities, deep learning-based methods focus on presuming missing items and extracting structured information. Lin et al. [25] constructed a unified deep neural framework to perform data recovery and consistency learning simultaneously. *Although these methods achieve a series of breakthroughs for IMVC, they suffer from enormous time and space consumption, which is prohibitive for processing large-scale datasets.*

As a promising way of alleviating time complexity, bipartite graph learning has attracted widespread attention in recent decades [27], [28], [29], [30]. Bipartite graph attempts to capture the intricate mechanisms by investigating the relationship between n raw samples and m anchors ($m \ll n$). Drawing the inspiration from this principle, the work in [29] constructed a small bipartite graph to efficiently compute the spectral embedding of the data. Yang et al. [31] utilized anchors to construct a sub-bipartite graph and then adopted the improved nonnegative and orthogonal factorization technology to obtain the clustering results. The work in [32] integrated anchor learning and correntropy learning to construct an efficient and robust clustering model. However, this method is single-view oriented. To enjoy the benefit of discriminative information from multi-view data, the work in [28] presented the first effort in handling large-scale MVC. Despite the remarkable performance of these approaches, *1) they only concentrate on pair-wise instance bipartite graphs while disregarding the global information across views; 2) they adopt a heuristic sampling strategy (e.g. k-means or random sampling) with fixed anchors, which is full of randomness and occasionality. Thus, anchor learning and subsequent bipartite graph learning are separated, which may lead to unstable clustering performance.*

To address these aforementioned deficiencies, we propose a novel self-completed bipartite graph learning (SCBGL) framework for fast IMVC, which seamlessly incorporates inter-view and intra-view bipartite graphs in a reciprocal way. Specifically, as shown in Fig. 1 (a), SCBGL leverages consensus anchors to construct an intra-view bipartite graph, which is contributed by all views, to preserve the heterogeneity of multi-view data. Meanwhile, by concatenating all heterogeneous features with projection learning, we derive an inter-view feature matrix that maximizes preserving the original information of multi-view data. Subsequently, SCBGL constructs an inter-view bipartite graph associated with global anchors, which contains plentiful information and elegantly encodes the complementary information, as displayed in Fig. 1 (b). Meanwhile, the informative inter-view bipartite graph can further tutor the self-completion of the consensus intra-view bipartite graph. The contributions of our proposed SCBGL mainly include:

- We unify incomplete bipartite graph self-completed learning, dynamic anchor learning and projection learning into a unified framework, which can efficiently tackle scalability IMVC in linear time respecting to instance numbers.
- Different from the previous works, SCBGL concurrently constructs both the intra-view and inter-view bipartite graphs to sufficiently explore the comprehensive characteristics hidden in the incomplete multi-view dataset. To the best of our knowledge, SCBGL is a pioneering work that simultaneously learns the intra-view and inter-view correlations for large-scale IMVC task.
- We devise an alternative optimization algorithm for the resultant objective. Validation experiments performed on large-scale datasets demonstrate the efficiency and competitiveness of SCBGL.

We organize the remainder of this paper as follows. Section II briefly introduces some related works and backgrounds. Section III elaborates our model SCBGL and the corresponding optimization process. The effectiveness and efficiency of our model SCBGL are valuated in Section IV. Finally, Section V offers the conclusion of this paper.

II. Preliminaries

A. Incomplete Multi-View Clustering

IMVC has inspired a surge of passion in the computer vision community. To date, a variety of IMVC methods have been presented, which can be classified into three main classes: matrix factorization (MF) [20], [21], [33], [34], graph or kernel construction [22], [23], [24], [35] and deep learning [25], [26]. Among these approaches, the MF technique has become popular in IMVC owing to its intuitive interpretability and promising performance. The MF-based methods are devoted to deriving a common low-dimensional clustering representation by decomposition technology. As a pioneer work, PVC [20] took advantage of NMF and partial alignment information to capture a consistent representation. Drawing support from weighted matrix factorization (WMF) and $l_{2,1}$ regularization, Shao et al. [21] and Hu and Chen [33] generated a consensus matrix for multi-view data, which is robust to noises and outliers. The work in [36] developed a graph-regularized MF model to capture the local underlying information. Nevertheless, almost all these approaches encounter costly computational and storage complexities. Considering the above issue, Shao et al. [37] processed IMVC in an online fashion and tackled the incomplete multi-view data chunk-by-chunk. However, this model still suffers from drawbacks in normalizing data matrices and handling missing instances. By means of regularized matrix factorization (RMF) and WMF, the work

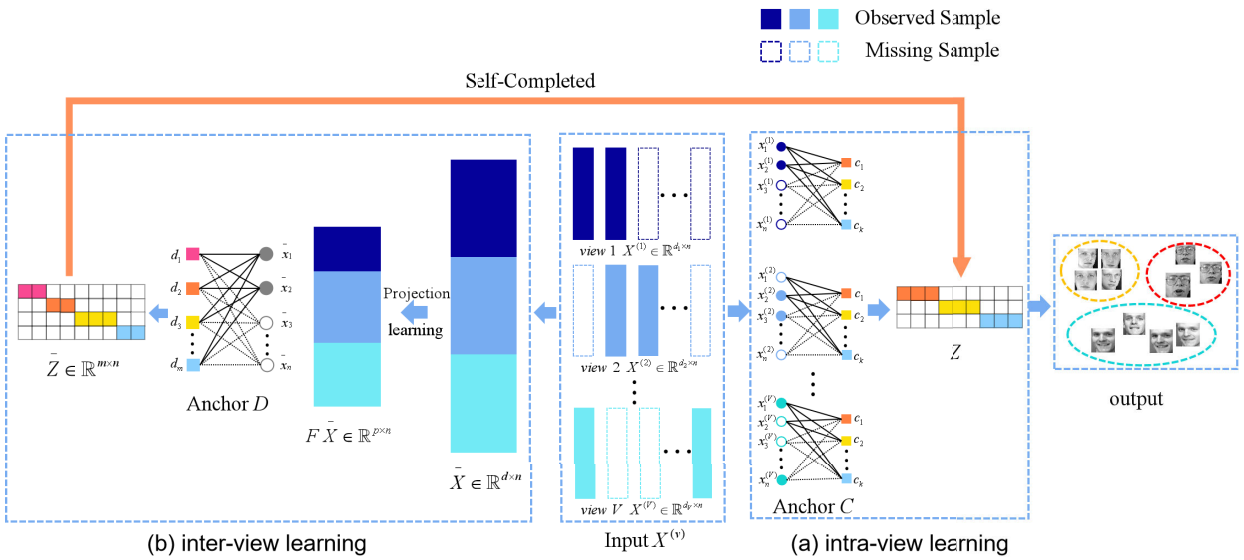


Fig. 1. The framework of the proposed SCBGL method. Given an arbitrary multi-view dataset with missing views $\{\boldsymbol{X}^{(v)} \in \mathbb{R}^{d_v \times n}\}_{v=1}^{V}$, for (a) intra-view learning, we dynamically learn consensus anchor matrix $\boldsymbol{C}$ and their corresponding intra-view bipartite graph $\boldsymbol{Z}$ to explore the diversity of incomplete multi-view datasets in the latent subspace. In (b) inter-view learning, we concatenate all heterogeneous features to get a global feature matrix $\widehat{\boldsymbol{X}}$ and then map it to a latent space with a projection matrix $\boldsymbol{F}$. Moreover, SCBGL adaptively construct global anchors $\boldsymbol{R}$ and an inter-view bipartite graph $\widehat{\boldsymbol{Z}}$ to excavate the global correlation cross views. To effectively recover the consensus intra-view bipartite graph with missing views, the informative inter-view bipartite graph further tutor the self-completion of the consensus intra-view bipartite graph.

in [34] can efficiently and effectively deal with large-scale IMVC task. The graph or kernel construction strategies strive to establish a unified similarity graph or kernel matrix across views. In [35], a late fusion scheme was proposed to efficiently integrate the incomplete kernel representations. To uncover the incomplete data structures, Li et al. [38] iteratively constructed local incomplete graphs to pursue a consensus graph. Deep learning-based methods endeavor to extract high-level information and infer missing views via deep neural networks. For instance, the work in [26] devised a deep framework to integrate cross-view relevance transfer and fusion learning via graph networks and encoder.

B. Bipartite Graph Learning

Bipartite graph, which emerged as a powerful strategy to decline the computational complexity, is prevalent in efficiently processing large-scale datasets in MVC [28], [39], [40], [41]. The principal pursuit of bipartite graph learning is to choose a portion of representative anchors to generate a similarity graph to measure the relationship between anchors and original points. More formally, given a multi-view dataset $\{\boldsymbol{X}^{(v)} \in \mathbb{R}^{d_v \times n}\}_{v=1}^{V}$ with n samples corresponding to k groups, the general multi-view bipartite graph framework is concisely formulated as:

$$\min_{\boldsymbol{Z},\boldsymbol{Z}^{(v)}} \sum_{v=1}^{V} \|\boldsymbol{X}^{(v)} - \boldsymbol{C}^{(v)}\boldsymbol{Z}^{(v)}\|_F^2 + f(\boldsymbol{Z}^{(v)}, \boldsymbol{Z}). \tag{1}$$

In model (1), $\boldsymbol{C}^{(v)} \in \mathbb{R}^{d_v \times m}$ denotes anchors for the v-th view consisting of m points, where each column is treated as an anchor; $f(\boldsymbol{Z}^{(v)}, \boldsymbol{Z})$ encourages the bipartite graphs $\boldsymbol{Z}^{(v)} \in \mathbb{R}^{m \times n}$ learned from various views towards a consensus one $\boldsymbol{Z} \in \mathbb{R}^{m \times n}$; The consensus anchor graph mirrors the instance-to-anchor similarity relationship, which substantially declines the computational and storage burden from $\mathcal{O}(n^2)$ to $\mathcal{O}(mn)$. Li et al. [27] employed bipartite graphs to efficiently handle multi-view spectral clustering. Following this line, [28] made the first attempt to extend bipartite graph learning to the domain of multi-view subspace clustering. However, the anchor selection is isolated from the clustering optimization, which may adversely affect clustering performance. Consequently, [40] integrated anchor learning and graph construction, thus negotiating mutually to improve clustering performance. Subsequently, [39] imposed F-norm on the consistent bipartite graph. Despite the comparable performance of these methods, they are devised for complete MVC circumstances. To handle large-scale IMVC, Wang et al. [42] introduced the bipartite graph to IMVC by defining the incompleteness of the data. Unfortunately, this method only utilizes the complete part of the data while disregarding the hidden information of the missing instances. In the next section, inspired by bipartite graph learning, we propose a novel self-completion framework for IMVC, which simultaneously integrates inter-view and intra-view information.

III. The Proposed SCBGL Method

A. The Proposed Method

For large-scale incomplete data, there is plenty of redundancy and missing information, which makes it extremely challenging to deal with large-scale IMVC. As a pioneer work, [42] introduced bipartite graph learning to handle IMVC tasks. Mathematically, this algorithm can be concisely formulated as:

$$\min_{\alpha,\boldsymbol{Z},\boldsymbol{C},\boldsymbol{P}^{(v)}} \sum_{v=1}^{V} \alpha_v^2 \|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2 + \lambda\|\boldsymbol{Z}\|_F^2,$$
$$s.t.\ \alpha^\top \mathbf{1} = 1,\ {\boldsymbol{P}^{(v)}}^\top \boldsymbol{P}^{(v)} = \boldsymbol{I}_k,\ \boldsymbol{C}^\top \boldsymbol{C} = \boldsymbol{I}_m,$$

$$\boldsymbol{Z} \geq 0, \boldsymbol{Z}^{\top}\mathbf{1} = \mathbf{1}. \tag{2}$$

where $\boldsymbol{A}^{(v)} \in \mathbb{R}^{n\times n_v}$ denotes the completeness of $\boldsymbol{X}^{(v)}$, which implies that the v-th view possesses n_v existing samples, thus $\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)}$ stands for the complete part of $\boldsymbol{X}^{(v)}$; $\boldsymbol{P}^{(v)}$ illustrates the projection matrix for the v-th view, which projects diverse dimensional complete data into a common latent subspace; α_v is the weight of the v-th view. To attain a consistent bipartite graph $\boldsymbol{Z} \in \mathbb{R}^{m\times n}$, this method constructs consensus anchors $\boldsymbol{C}$ shared by all the views. However, this method only adopts available instances to execute subsequent clustering operations and overlooks the discriminant information of missing views, which may deteriorate clustering performance. As stated in [43] that view-wise weights may be ineffective and further increase the time and space overhead. Hence, we do not consider the weight coefficient α in our model. To overcome this issue, we propose the following self-completion model for fast IMVC, which seamlessly constructs intra-view and inter-view bipartite graphs to capture local and global discriminative information. Specifically, the inter-view bipartite graph can complete the missing instances of the intra-view bipartite graph by the following model.

$$\begin{aligned}\min_{\boldsymbol{P}^{(v)},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{F},\boldsymbol{R},\widehat{\boldsymbol{Z}},\boldsymbol{Q}} & \sum_{v=1}^{V}\|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2 \\ & + \lambda_1\|\boldsymbol{Z}\|_F^2 + \lambda_2\|\boldsymbol{F}\widehat{\boldsymbol{X}} - \boldsymbol{R}\widehat{\boldsymbol{Z}}\|_F^2 \\ & + \lambda_3\|\boldsymbol{Z} - \boldsymbol{Q}\widehat{\boldsymbol{Z}}\|_{2,1} \\ s.t.\ & \boldsymbol{P}^{(v)^{\top}}\boldsymbol{P}^{(v)} = \boldsymbol{I}_k, \boldsymbol{F}^{\top}\boldsymbol{F} = \boldsymbol{I}_d, \boldsymbol{R}^{\top}\boldsymbol{R} = \boldsymbol{I}_m, \\ & \boldsymbol{C}^{\top}\boldsymbol{C} = \boldsymbol{I}_k, \boldsymbol{Q}^{\top}\boldsymbol{Q} = \boldsymbol{I}_m, \boldsymbol{Z} \geq 0, \boldsymbol{Z}^{\top}\mathbf{1} = \mathbf{1}, \end{aligned} \tag{3}$$

To simplify the calculation and ensure the efficiency of our algorithm, we let $\boldsymbol{Q} = \boldsymbol{I}$. Therefore, the objective formula can be transformed into the following optimization problem:

$$\begin{aligned}\min_{\boldsymbol{P}^{(v)},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{F},\boldsymbol{R},\widehat{\boldsymbol{Z}}} & \sum_{v=1}^{V}\|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2 \\ & + \lambda_1\|\boldsymbol{Z}\|_F^2 + \lambda_2\|\boldsymbol{F}\widehat{\boldsymbol{X}} - \boldsymbol{R}\widehat{\boldsymbol{Z}}\|_F^2 + \lambda_3\|\boldsymbol{Z} - \widehat{\boldsymbol{Z}}\|_{2,1} \\ s.t.\ & \boldsymbol{P}^{(v)^{\top}}\boldsymbol{P}^{(v)} = \boldsymbol{I}_k, \boldsymbol{F}^{\top}\boldsymbol{F} = \boldsymbol{I}_d, \boldsymbol{R}^{\top}\boldsymbol{R} = \boldsymbol{I}_m, \\ & \boldsymbol{C}^{\top}\boldsymbol{C} = \boldsymbol{I}_k, \boldsymbol{Z} \geq 0, \boldsymbol{Z}^{\top}\mathbf{1} = \mathbf{1}. \end{aligned} \tag{4}$$

$$\begin{aligned}\min_{\boldsymbol{W}^{(v)},\boldsymbol{A},\boldsymbol{Z},\overline{\boldsymbol{D}},\overline{\boldsymbol{Z}},\boldsymbol{Q}} & \sum_{v=1}^{V}\|\boldsymbol{X}^{(v)} - \boldsymbol{W}^{(v)}\boldsymbol{A}\boldsymbol{Z}\|_F^2 + \tau\|\overline{\boldsymbol{X}} - \overline{\boldsymbol{D}\boldsymbol{Z}}\|_F^2 \\ & + \lambda_1(\|\boldsymbol{Z}\|_F^2 + \|\overline{\boldsymbol{Z}}\|_F^2 + \|\boldsymbol{Z} - \boldsymbol{Q}^{\top}\overline{\boldsymbol{Z}}\|_F) \\ & + \lambda_2\|\boldsymbol{Q}\|_{2,1}, \\ s.t.\ & \boldsymbol{W}^{(v)^{\top}}\boldsymbol{W}^{(v)} = \boldsymbol{I}_k, \boldsymbol{A}^{\top}\boldsymbol{A} = \boldsymbol{I}_k, \overline{\boldsymbol{D}}^{\top}\overline{\boldsymbol{D}} = \boldsymbol{I}_k, \\ & \boldsymbol{Z} \geq 0, \boldsymbol{Z}^{\top}\mathbf{1} = \mathbf{1}. \end{aligned} \tag{5}$$

For the convenience of calculation, we introduce an auxiliary variable $\boldsymbol{E}$ into Eq. (5) and construct an unconstrained augmented Lagrangian function as:

$$\begin{aligned}\min_{\boldsymbol{P}^{(v)},\boldsymbol{C},\boldsymbol{E},\boldsymbol{Z},\boldsymbol{F},\boldsymbol{R},\widehat{\boldsymbol{Z}}} & \sum_{v=1}^{V}\|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2 \\ & + \lambda_1\|\boldsymbol{Z}\|_F^2 + \lambda_2\|\boldsymbol{F}\widehat{\boldsymbol{X}} - \boldsymbol{R}\widehat{\boldsymbol{Z}}\|_F^2 + \lambda_3\|\boldsymbol{E}\|_{2,1} \\ & + \frac{\mu}{2}\|\boldsymbol{Z} - \widehat{\boldsymbol{Z}} - \boldsymbol{E} + \frac{\boldsymbol{Y}}{\mu}\|_F^2 \\ s.t.\ & \boldsymbol{P}^{(v)^{\top}}\boldsymbol{P}^{(v)} = \boldsymbol{I}_k, \boldsymbol{F}^{\top}\boldsymbol{F} = \boldsymbol{I}_d, \boldsymbol{R}^{\top}\boldsymbol{R} = \boldsymbol{I}_m, \\ & \boldsymbol{C}^{\top}\boldsymbol{C} = \boldsymbol{I}_k, \boldsymbol{Z} \geq 0, \boldsymbol{Z}^{\top}\mathbf{1} = \mathbf{1}. \end{aligned} \tag{6}$$

where $\boldsymbol{Y} \in \mathbb{R}^{m\times n}$ and μ represent Lagrangian multiplier and penalty parameter, respectively; $\widehat{\boldsymbol{X}} \in \mathbb{R}^{d\times n}$ mirrors a global feature matrix by concatenating features of diverse views, as shown in Fig. 1 (b), which maximizes the preservation of the original incomplete information; $\boldsymbol{F} \in \mathbb{R}^{p\times d}$ ($d = \sum_{v=1}^{V} d_v$) denotes a mapping that succinctly projects the original data to latent spaces with distinct data distribution and less noise. To pursue a more consistent bipartite graph, we establish an intra-view bipartite graph $\boldsymbol{Z} \in \mathbb{R}^{k\times n}$ and an inter-view bipartite graph $\widehat{\boldsymbol{Z}} \in \mathbb{R}^{m\times n}$ with consensus anchors $\boldsymbol{C} \in \mathbb{R}^{k\times k}$ and global anchor $\boldsymbol{R} \in \mathbb{R}^{p\times m}$ ($m = k$) guidance, respectively. As the term $\lambda_3\|\boldsymbol{Z} - \widehat{\boldsymbol{Z}}\|_{2,1}^2$ shows, the inter-view bipartite graph $\widehat{\boldsymbol{Z}}$ further tutors the self-completion process of the intra-view bipartite graph $\boldsymbol{Z}$. Moreover, rather than using heuristic strategies to get fixed anchors [28], our model dynamically learns representative and discriminative features and points as anchors to capture the manifold structure. Finally, we apply SVD to $\boldsymbol{Z}$ to gain $\boldsymbol{U} \in \mathbb{R}^{n\times n}$ and then invoke k-means on it to attain clustering results. Upon this model, SCBGL can well capture the inter-view correlation within and across multiple views so as to perform efficiently and effectively for large-scale IMVC tasks.

B. Optimization

Following [44], [45], we solve the SCBGL model by the following iterative strategy, i.e., update all variables iteratively. Specifically, SCBGL decomposes the optimization process into the following seven subproblems:

Step 1 update $\boldsymbol{P}^{(v)}$: Fixing the variables that are irrelevant to $\boldsymbol{P}^{(v)}$, we can reformulate Eq. (5) as:

$$\begin{aligned}\min_{\boldsymbol{P}^{(v)}} & \sum_{v=1}^{V}\|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2, \\ s.t.\ & \boldsymbol{P}^{(v)^{\top}}\boldsymbol{P}^{(v)} = \boldsymbol{I}_k. \end{aligned} \tag{7}$$

Notice that each $\boldsymbol{P}^{(v)}$ is separable. Therefore, by removing irrelevant items, we can convert the above F-norm into the form of trace as:

$$\begin{aligned}\max_{\boldsymbol{P}^{(v)}} & tr(\boldsymbol{P}^{(v)^{\top}}(\boldsymbol{X}^{(v)} \otimes \boldsymbol{B}^{(v)})\boldsymbol{Z}^{\top}\boldsymbol{C}^{\top}), \\ s.t.\ & \boldsymbol{P}^{(v)^{\top}}\boldsymbol{P}^{(v)} = \boldsymbol{I}_k, \end{aligned} \tag{8}$$

for simplicity, we observe that $\boldsymbol{X}^{(v)} \otimes \boldsymbol{B}^{(v)} = \boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)}\boldsymbol{A}^{(v)^{\top}}$, where $\boldsymbol{B}^{(v)} = \mathbf{1}_{d_v}\boldsymbol{a}^{(v)} \in \mathbb{R}^{d_v\times n}$ and $\boldsymbol{a}^{(v)} = [\boldsymbol{a}_1^{(v)}, \cdots, \boldsymbol{a}_n^{(v)}]^{\top}$ with $\boldsymbol{a}_j^{(v)} = \sum_{l=1}^{n_v}\boldsymbol{A}_{j,l}^{(v)}$. Inspired by [46], we suppose the singular value decomposition of $(\boldsymbol{X}^{(v)} \otimes \boldsymbol{B}^{(v)})\boldsymbol{Z}^{\top}\boldsymbol{C}^{\top}$ is $\boldsymbol{U}_W\boldsymbol{S}_W\boldsymbol{V}_W^{\top}$, and the closed-form optimization solution of $\boldsymbol{P}^{(v)}$ is $\boldsymbol{U}_W\boldsymbol{V}_W^{\top}$.

When calculating $(\boldsymbol{X}^{(v)} \otimes \boldsymbol{B}^{(v)})\boldsymbol{Z}^{\top}\boldsymbol{C}^{\top}$, it costs $\mathcal{O}(d_v nk + d_v k^2)$ for each view. Accordingly, for the v sub-problem, update $\boldsymbol{P}$ needs $\mathcal{O}(dnk + dk^2)$. Besides, solving Eq. (8) costs

$\boldsymbol{P}$ needs $\mathcal{O}(dk^2)$. Therefore, obtaining $\boldsymbol{P}$ takes $\mathcal{O}(dnk+dk^2)$ per iteration.

Step 2 update $\boldsymbol{C}$: Removing the irrelevant terms in Eq. (5), consensus anchor matrix $\boldsymbol{C}$ can be settled by calculating:

$$\begin{aligned}&\min_{\boldsymbol{C}} \sum_{v=1}^{V} \|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2,\\ &s.t.\ \boldsymbol{C}^\top\boldsymbol{C} = \boldsymbol{I}_k.\end{aligned} \tag{9}$$

Similar with the optimization process of $\boldsymbol{P}^{(v)}$, we can get the equivalent transformation of Eq. (9) as:

$$\begin{aligned}&\max_{\boldsymbol{C}}\ tr\Big(\boldsymbol{C}^\top \sum_{v=1}^{V} \big(\boldsymbol{P}^{(v)^\top}(\boldsymbol{X}^{(v)} \otimes \boldsymbol{B}^{(v)})\big)\boldsymbol{Z}^\top\Big),\\ &s.t.\ \boldsymbol{C}^\top\boldsymbol{C} = \boldsymbol{I}_k.\end{aligned} \tag{10}$$

Likewise, by calculating SVD of $\sum_{v=1}^{V}\big(\boldsymbol{P}^{(v)^\top}(\boldsymbol{X}^{(v)} \otimes \boldsymbol{B}^{(v)})\big)\boldsymbol{Z}^\top$ as $\boldsymbol{U}_C\boldsymbol{S}_C\boldsymbol{V}_C^\top$, we can derive the closed-form solution as $\boldsymbol{C} = \boldsymbol{U}_C\boldsymbol{V}_C^\top$. Similar to $\boldsymbol{P}$, the total complexity of update $\boldsymbol{C}$ takes $\mathcal{O}(dnk + nk^2 + k^3)$.

Step 3 update $\boldsymbol{Z}$: To acquire the intra-view bipartite graph, we minimize Eq. (5) with respect to $\boldsymbol{Z}$ as:

$$\begin{aligned}\min_{\boldsymbol{Z}} \sum_{v=1}^{V} &\|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2 + \lambda_1\|\boldsymbol{Z}\|_F^2\\ &+ \frac{\mu}{2}\|\boldsymbol{Z} - \boldsymbol{G}\|_F^2\\ s.t.\ &\boldsymbol{Z} \geq 0, \boldsymbol{Z}^\top\boldsymbol{1} = \boldsymbol{1},\end{aligned} \tag{11}$$

where $\boldsymbol{G} = \widehat{\boldsymbol{Z}} + \boldsymbol{E} - \frac{\boldsymbol{Y}}{\mu}$. Employing the Lagrangian technique, we have the Lagrangian function of Eq. (11):

$$\begin{aligned}\mathcal{L}(\boldsymbol{Z}) = \min_{\boldsymbol{Z}} &\sum_{v=1}^{V} \|\boldsymbol{X}^{(v)}\boldsymbol{A}^{(v)} - \boldsymbol{P}^{(v)}\boldsymbol{C}\boldsymbol{Z}\boldsymbol{A}^{(v)}\|_F^2 + \lambda_1\|\boldsymbol{Z}\|_F^2\\ &+ \frac{\mu}{2}\|\boldsymbol{Z} - \boldsymbol{G}\|_F^2 + tr(\boldsymbol{y}^\top(\boldsymbol{Z}^\top\boldsymbol{1} - \boldsymbol{1})).\end{aligned} \tag{12}$$

Focusing on $\boldsymbol{Z}$, Eq. (12) can be equivalently transformed into:

$$\begin{aligned}\min_{\boldsymbol{Z}} &\sum_{v=1}^{V} \|\boldsymbol{Z} - \boldsymbol{M}\|_F^2 + tr(\boldsymbol{y}^\top(\boldsymbol{Z}^\top\boldsymbol{1} - \boldsymbol{1}))\\ s.t.\ &\boldsymbol{Z} \geq 0,\end{aligned} \tag{13}$$

where $\boldsymbol{M}_{:,j} = \frac{1}{\lambda_1+\lambda_3+\sum_{v=1}^{V}\boldsymbol{B}_{1,j}^{(v)}}\big(\sum_{v=1}^{V}\boldsymbol{N}_{:,j}\boldsymbol{B}_{1,j}^{(v)} + \lambda_3\boldsymbol{G}_{:,j}\big)$ with $\boldsymbol{N} = \boldsymbol{C}^\top\boldsymbol{P}^{(v)\top}\boldsymbol{X}^{(v)}$. $\boldsymbol{N}_{:,j}$ and $\boldsymbol{G}_{:,j}$ denote the j-th column of $\boldsymbol{N}$ and $\boldsymbol{G}$, respectively. Assuming $\boldsymbol{Z}_{:,j}$ represents the j-th column of $\boldsymbol{Z}$, we optimize $\boldsymbol{Z}$ column-by-column as:

$$\min_{\boldsymbol{Z}} \sum_{v=1}^{V} \|\boldsymbol{Z}_{:,j} - \boldsymbol{M}_{:,j}\|_F^2 + \boldsymbol{y}_j(\boldsymbol{Z}_{:,j}^\top\boldsymbol{1} - \boldsymbol{1}),\ s.t.\ \boldsymbol{Z} \geq 0. \tag{14}$$

Setting the derivative of (14) w.r.t. $\boldsymbol{Z}_{:,j}$ to zero, the optimal solution for $\boldsymbol{Z}_{:,j}$ can be yielded:

$$\boldsymbol{Z}_{:,j} = \max(\boldsymbol{M}_{:,j} - \frac{1}{2}\boldsymbol{y}_j\boldsymbol{1}, 0). \tag{15}$$

Together with the sum of each row of $\boldsymbol{Z}$ is 1, we have $\boldsymbol{y}_j = \frac{2(\sum_{i=1}^{k}\boldsymbol{M}_{i,j} - 1)}{k}$. For update $\boldsymbol{Z}$, it costs $\mathcal{O}(dnk)$.

Step 4 update $\boldsymbol{F}$: With other variables fixed, the objective function in terms of $\boldsymbol{F}$ is reduced to:

$$\min_{\boldsymbol{F}} \|\boldsymbol{F}\widehat{\boldsymbol{X}} - \boldsymbol{R}\widehat{\boldsymbol{Z}}\|_F^2,\ s.t.\ \boldsymbol{F}^\top\boldsymbol{F} = \boldsymbol{I}_d. \tag{16}$$

Similar to the optimization of $\boldsymbol{P}^{(v)}$, expanding the F-norm norm by trace, we have:

$$\max_{\boldsymbol{F}}\ tr(\boldsymbol{F}^\top\boldsymbol{R}\widehat{\boldsymbol{Z}}\widehat{\boldsymbol{X}}^\top),\ s.t.\ \boldsymbol{F}^\top\boldsymbol{F} = \boldsymbol{I}_d. \tag{17}$$

Likewise, the closed-form solution for $\boldsymbol{F}$ is $\boldsymbol{U}_F\boldsymbol{V}_F^\top$, where $\boldsymbol{R}\widehat{\boldsymbol{Z}}\widehat{\boldsymbol{X}}^\top = \boldsymbol{U}_F\boldsymbol{S}_F\boldsymbol{V}_F^\top$. To obtain the optimal $\boldsymbol{F}$, it needs $\mathcal{O}(pmn + pnd + pd^2)$ to slove the problem in Eq. (17).

Step 5 update $\boldsymbol{R}$: Ignoring terms irrelevant of $\boldsymbol{R}$, the Eq. (5) w.r.t. $\boldsymbol{R}$ can be written as:

$$\min_{\boldsymbol{R}} \|\boldsymbol{F}\widehat{\boldsymbol{X}} - \boldsymbol{R}\widehat{\boldsymbol{Z}}\|_F^2,\ s.t.\ \boldsymbol{R}^\top\boldsymbol{R} = \boldsymbol{I}_m. \tag{18}$$

As well as updating $\boldsymbol{P}^{(v)}$, we can represent Eq. (18) as the equivalent optimization formula:

$$\max_{\boldsymbol{R}}\ tr(\boldsymbol{R}^\top\boldsymbol{F}\widehat{\boldsymbol{X}}\widehat{\boldsymbol{Z}}^\top),\ s.t.\ \boldsymbol{R}^\top\boldsymbol{R} = \boldsymbol{I}_m. \tag{19}$$

By performing a singular value decomposition, the optimization result for $\boldsymbol{R}$ is $\boldsymbol{U}_D\boldsymbol{V}_D^\top$, where $\boldsymbol{F}\widehat{\boldsymbol{X}}\widehat{\boldsymbol{Z}}^\top = \boldsymbol{U}_D\boldsymbol{S}_D\boldsymbol{V}_D^\top$. Congruously, computing the global anchors consumes $\mathcal{O}(pdn + pnm + pm^2)$.

Step 6 update $\widehat{\boldsymbol{Z}}$: By keeping other variables fixed, the process of updating inter-view bipartite graph $\widehat{\boldsymbol{Z}}$ can be formulated as:

$$\min_{\widehat{\boldsymbol{Z}}} \lambda_2\|\boldsymbol{F}\widehat{\boldsymbol{X}} - \boldsymbol{R}\widehat{\boldsymbol{Z}}\|_F^2 + \frac{\mu}{2}\|\boldsymbol{H} - \widehat{\boldsymbol{Z}}\|_F^2, \tag{20}$$

where $\boldsymbol{H} = \boldsymbol{Z} - \boldsymbol{E} + \frac{\boldsymbol{Y}}{\mu}$. Setting the derivative of Eq. (20) about $\widehat{\boldsymbol{Z}}$ to zero, the closed-form solution can be obtained by:

$$\begin{aligned}\widehat{\boldsymbol{Z}} &= (\lambda_2\boldsymbol{R}^\top\boldsymbol{R} + \frac{\mu}{2}\boldsymbol{I})^{-1}(\lambda_2\boldsymbol{R}^\top\boldsymbol{F}\widehat{\boldsymbol{X}} + \frac{\mu}{2}\boldsymbol{I}^\top\boldsymbol{H})\\ &= (\lambda_2\boldsymbol{R}^\top\boldsymbol{F}\widehat{\boldsymbol{X}} + \frac{\mu}{2}\boldsymbol{H})/(\lambda_2 + \frac{\mu}{2}).\end{aligned} \tag{21}$$

The last equation is based on the orthogonal constraints on variables $\boldsymbol{R}$. The optimization process of $\widehat{\boldsymbol{Z}}$ costs $\mathcal{O}(mpd + mdn)$ by calculating matrix multiplication.

Step 7 update $\boldsymbol{E}$: Fixing other variables, the optimization sub-problem about $\boldsymbol{E}$ is simplified as:

$$\min_{\boldsymbol{E}} \lambda_3\|\boldsymbol{E}\|_{2,1} + \frac{\mu}{2}\|\boldsymbol{Z} - \widehat{\boldsymbol{Z}} - \boldsymbol{E} + \frac{\boldsymbol{Y}}{\mu}\|_F^2. \tag{22}$$

Since the above formula has been investigated in [47] and [48], so we can achieve a closed-form solution with the shrinkage operator $\boldsymbol{S}_{\frac{\lambda_3}{\mu}}$:

$$\boldsymbol{E} = \boldsymbol{S}_{\frac{\lambda_3}{\mu}}(\boldsymbol{Z} - \widehat{\boldsymbol{Z}} + \frac{\boldsymbol{Y}}{\mu}). \tag{23}$$

The optimization process of $\boldsymbol{E}$ costs $O(mn)$ by calculating matrix addition.

Algorithm 1 TDASC Algorithm

Input: Multi-view data $\{\boldsymbol{X}^{(v)}\}_{v=1}^{V}$, $\rho = 1.1$, $\mu = 10^{-6}$, $max_{\mu} = 10^{10}$, parameter λ_1, λ_2, λ_3 and the number of clusters k.

Initialize: $\boldsymbol{P}^{(v)} = \boldsymbol{0}, \boldsymbol{C} = \boldsymbol{0}, \boldsymbol{Z} = \boldsymbol{0}, \boldsymbol{F} = \boldsymbol{0}, \boldsymbol{R} = \boldsymbol{0}, \boldsymbol{E} = \boldsymbol{0}, \widehat{\boldsymbol{Z}} = \boldsymbol{0}$.

1: **while** not converged **do**
2: Update view-specific projection matrix $\boldsymbol{P}^{(v)}$ by solving Eq. (8).
3: Update consensus anchors $\boldsymbol{C}$ by solving Eq. (10).
4: Update consistent bipartite graph $\boldsymbol{Z}$ by Eq. (15).
5: Update global projection matrix $\boldsymbol{F}$ by solving Eq. (17).
6: Update global anchor $\boldsymbol{R}$ by solving Eq. (19).
7: Update inter-view bipartite graph $\widehat{\boldsymbol{Z}}$ by solving Eq. (21).
8: Update an auxiliary variable $\boldsymbol{E}$ by solving Eq. (23).
9: Update the multiplier $\boldsymbol{Y}$: $\boldsymbol{Y} = \boldsymbol{Y} + \mu(\boldsymbol{Z} - \widehat{\boldsymbol{Z}} - \boldsymbol{E})$.
10: Update the parameter μ: $\mu = min(\rho\mu, max_{\mu})$.
11: **end while**
12: **return** $\boldsymbol{U}$ by performing SVD on $\boldsymbol{Z}$.

Output: Invoking k-means on $\boldsymbol{U}$ to attain final cluster labels.

TABLE I
SUMMARY OF ALL REAL-WORLD MULTI-VIEW DATABASES

Dataset	Sample	Class	View	Features
ORL	400	40	3	4096,3304,6750
NGs	500	5	3	200,200,200
MSRC	210	7	3	1302,512,256
WebKB	1051	2	2	2949,334
Caltech101-7	1474	7	6	48,40,254,1984,512,928
Caltech101-20	2386	20	6	48,40,254,1984,512,928
BDGP	2500	5	3	100,500,250
Reuters	18758	6	5	21531,24892,34251, 15506,11547
NUSWIDE	30000	31	5	65,226,145,74,129
Cifar10	50000	10	3	512,2048,1024
Cifar100	50000	100	3	512,2048,1024
MNIST	60000	10	3	342,1024,64

C. Complexity Analysis

We devise an alternating iterative optimization method to solve Eq. (5). The computational complexity of SCBGL contains seven parts. Specifically, for updating $\boldsymbol{P}^{(v)}$, the main complexity is matrix manipulation, which is $\mathcal{O}(dnk + dk^2)$. In step 2, it costs $\mathcal{O}(dnk + nk^2 + k^3)$ to update anchors consistent $\boldsymbol{C}$. To obtain an optimal $\boldsymbol{Z}$, it needs $\mathcal{O}(dnk)$. At each iteration, it consumes $\mathcal{O}(pmn + pnd + pd^2)$ to optimize $\boldsymbol{F}$ in a closed-form solution. As for update $\boldsymbol{R}$, it takes $\mathcal{O}(pdn + pnm + pm^2)$ operations. The optimization process of inter-view bipartite graph costs $\mathcal{O}(mpd + mdn)$ by calculating matrix multiplication. For the remaining step, it needs $\mathcal{O}(mn)$ to calculate $\boldsymbol{E}$. Obviously, these optimization sub-problems are all linearly growing complexity with n. Moreover, subsequent SVD operations on the consensus bipartite graph and the cluster results executed by k-means also satisfy linear time. Therefore, the time complexity of SCBGL is linear with the number of samples, indicating that SCBGL is commendable in tackling large-scale datasets.

IV. EXPERIMENTS

In this section, we estimate the competitiveness and efficiency of our proposed SCBGL on twelve incomplete datasets for IMVC.

A. Incomplete Multi-View Dataset

To verify the effectiveness of our proposed SCBGL, we perform experiments on twelve datasets distributed in six categories: face image, news story, generic object, web page, bioinformatic, and handwritten digit. Some information of these benchmark datasets are summarized in Table I.

ORL[1]**:** This face dataset possesses 400 images of 40 individuals. Following [16], intensity, LBP [49] and Gabor [50] features were extracted as three views with 4096, 3304, and 6750 dimensions, respectively.

NGs[2]**:** As a popular subset of 20Newsgroups, NGs derives from a wide variety of sources distributed in five categories. For this dataset, its three feature matrices with three different preprocessing were extracted to form multi-view dataset.

MSRC: MSRC is a generic object dataset described by three visual features, where images can be tagged with seven labels. Following [51], color moment, GIST, CENTRIST feature, and LBP visual features were selected as four views.

WebKB[3]**:** WebKB totally consists of 1051 web pages and hyperlink data from four university websites. The exploited multi-view WebKB dataset is depicted by two aspects, that is, the content and citation.

Caltech101-7 and Caltech101-20[4]**:** Caltech101 is a widely used generic object dataset with 101 categories. Following [27], we select the commonly used 7 classes with 1474 images and 20 classes with 2386 instances, respectively, as the experiment multi-view datasets.

BDGP[5]**:** It is an extensively used bioinformatic dataset, including 2500 Drosophila embryo samples of five categories.

Reuters[6]**:** Reuters is a repository of documents with news articles, containing 18758 documents over six labels. Each document is characterized in five different languages, that is, 21531-D English, 24892-D French, 34251-D German, 15506-D Italian, 11547-D Spanish.

NUSWIDE[7]**:** NUSWIDE is also a generic object database with 30,000 instances, which encodes images with five types of low-level feature descriptors.

Cifar10 and Cifar100[8]**:** They are widely used large-scale datasets, including 50, 000 tiny images corresponding to 10 and 100 categories, respectively. Moreover, they are 3-channel color RGB images with dimensions 512-D, 2048-D, and 1024-D.

[1]http://www.uk.research.att.com/facedatabase.html
[2]https://lig-membres.imag.fr/grimal/data.html
[3]http://yallara.cs.rmit.edu.au/ phmartin/WebKB/
[4]http://www.vision.caltech.edu/Image Datasets/Caltech101/
[5]https://www.fruitfly.org/
[6]https://archive.ics.uci.edu/ml/datasets/reuters-21578+text+categorization+collection
[7]https://lms.comp.nus.edu.sg/wp-content/uploads/2019/research/nuswide/NUS-WIDE.html
[8]http://www.cs.toronto.edu/ kriz/cifar.html

TABLE II
CLUSTERING PERFORMANCE (ACC%±STD%) ON TWELVE BENCHMARK DATASETS

Dataset	BSV	MIC	APMC	DAIMC	FLSD	IMSR	V^{3}H	IMVC-CBG	PSIMVC-PG	SCBGL
ORL	33.45±1.81	39.05±1.20	68.50±1.94	67.45±0.43	40.35±2.67	67.00±0.00	67.50±2.22	67.59±1.35	59.96±0.70	**83.10±0.17**
NGs	41.13±2.03	20.90±0.17	41.20±0.00	58.33±0.62	83.00±0.00	80.00±0.00	71.69±0.00	84.42±0.23	87.00±0.00	**90.40±0.00**
MSRC	25.81±0.92	46.00±1.26	37.14±0.00	63.17±0.76	54.38±0.62	51.33±0.00	68.27±1.64	64.14±1.33	63.81±0.00	**79.23±0.26**
WebKB	57.08±2.39	72.81±0.27	65.84±0.00	82.17±0.05	62.13±0.00	79.54±0.00	81.18±0.00	44.94±0.00	70.92±0.10	**86.20±0.00**
Caltech101-7	54.63±0.07	37.33±1.97	N/A	42.10±3.72	55.08±0.22	34.53±0.00	N/A	62.32±1.03	49.80±0.03	**72.08±0.09**
Caltech101-20	39.16±0.35	26.53±1.55	N/A	45.68±2.10	41.78±1.18	42.37±0.00	N/A	48.23±1.57	44.83±1.11	**54.98±0.10**
BDGP	34.96±1.06	25.90±0.62	33.88±0.00	35.70±0.16	42.41±0.06	37.00±0.00	32.18±0.00	42.40±0.00	42.90±0.20	**50.07±0.65**
NUSWIDE	12.05±0.03	N/A	N/A	**13.79±0.37**	N/A	N/A	N/A	12.53±0.11	11.61±0.09	13.22±0.10
Reuters	N/A	N/A	N/A	N/A	N/A	N/A	N/A	44.94±0.00	46.05±0.00	**48.06±0.00**
Cifar10	N/A	N/A	N/A	90.81±0.45	N/A	N/A	N/A	91.74±0.51	95.94±0.00	**96.47±0.00**
Cifar100	N/A	N/A	N/A	89.71±1.00	N/A	N/A	N/A	93.09±1.18	97.01±0.00	**99.05±0.53**
MNIST	N/A	N/A	N/A	97.57±0.31	N/A	N/A	N/A	98.10±0.06	98.24±0.00	**98.33±0.00**

TABLE III
CLUSTERING PERFORMANCE (NMI%±STD%) ON TWELVE BENCHMARK DATASETS

Dataset	BSV	MIC	APMC	DAIMC	FLSD	IMSR	V^{3}H	IMVC-CBG	PSIMVC-PG	SCBGL
ORL	54.11±1.91	55.80±0.82	80.57±0.93	83.22±0.29	57.79±1.99	79.67±0.00	79.98±0.82	83.15±0.42	77.00±0.42	**91.29±0.05**
NGs	20.24±1.37	2.42±0.26	24.42±0.00	39.32±0.47	60.89±0.00	57.46±0.00	52.97±0.00	69.62±0.00	66.23±0.00	**76.16±0.00**
MSRC	21.85±0.68	33.61±0.18	24.57±0.56	53.67±0.59	46.45±1.13	39.17±0.01	58.29±1.00	56.39±1.93	51.62±0.00	**67.38±0.23**
WebKB	1.80±0.86	6.97±0.00	3.71±0.00	17.66±0.10	1.26±0.00	19.07±0.00	33.09±0.00	29.95±0.00	10.17±0.08	**37.89±0.00**
Caltech101-7	15.93±0.60	24.8±1.13	N/A	45.45±2.12	37.20±0.35	40.30±0.00	N/A	44.60±0.04	46.22±0.05	**47.24±0.13**
Caltech101-20	25.26±0.66	30.02±1.31	N/A	55.56±1.18	50.90±0.70	50.47±0.00	N/A	50.45±1.17	50.25±0.40	**57.21±0.03**
BDGP	12.88±0.94	4.49±0.47	10.12±0.00	11.75±0.14	18.50±0.09	14.80±0.00	10.58±0.00	19.06±0.15	18.93±0.10	**27.01±0.00**
NUSWIDE	2.68±0.03	N/A	N/A	**11.83±0.36**	N/A	N/A	N/A	10.42 ±0.03	9.39±0.07	10.47±0.10
Reuters	N/A	N/A	N/A	N/A	N/A	N/A	N/A	**29.95±0.00**	29.25±0.00	29.47±0.00
Cifar10	N/A	N/A	N/A	90.47±0.55	N/A	N/A	N/A	90.57±0.00	90.45±0.00	**91.05±0.00**
Cifar100	N/A	N/A	N/A	98.26±0.16	N/A	N/A	N/A	98.63±0.29	99.32±0.00	**99.81±0.08**
MNIST	N/A	N/A	N/A	93.89±0.53	N/A	N/A	N/A	94.94±0.13	95.25±0.00	**95.35±0.00**

MNIST[9]**:** The MNIST database is a large-scale handwritten digit dataset that can be distributed into 10 mutually exclusive groups. It is comprised of 60,000 digital images, containing numbers from 0 to 9.

Following [42], we generate incomplete versions of the datasets mentioned above. Specifically, on the premise that each sample exists in at least one view, we haphazardly delete $100p\%$ samples in each view to form an incomplete multi-view dataset with missing rates $p \in [0.1 : 0.2 : 0.9]$.

B. Evaluation Metric

We employ four commonly-used evaluation metrics to investigate the superiority of SCBGL, namely, Accuracy (ACC), Normalized Mutual Information (NMI), Purity and Fscore [42]. Different evaluation metrics have different emphases, but they all satisfy the property that the larger the value, the better the performance.

C. Compared Methods

We compare SCBGL with the following approaches: As a baseline method, **Best Single-view Spectral Clustering (BSV)** [52] fills missing views with zeros and reports the best performance of all views. Besides this, several state-of-the-art IMVC approaches are adopted as competitors: **MIC** [21] utilizes weighted NMF and $l_{2,1}$-norm regularization to learn a consensus latent representation; **APMC** [53] uses the Gaussian kernel-based similarity to capture non-linear relations; **FLSD** [36] develops a general graph-regularized framework with adaptive weighted learning; Inspired by genetics, **V^{3}H** [54] is a pioneering work to excavate both the consistent and unique information of incomplete datasets; **DAIMC** [33] yields a consistent latent feature matrix with a respective weight matrix and $l_{2,1}$-norm regularized regression; **IMSR** [55] adapts the self-representation subspace strategy and data imputation learning to handle incomplete data; **IMVC-CBG** [42] introduces consensus anchors to learn a common bipartite graph, which can flexibly and efficiently handle large-scale IMVC; **PSIMVC-PG** [56] presents a parameter-free and scalable IMVC framework to flexibly complete the prototype graph. The comparison methods BSV, APMC and PSIMVC-PG were parameter-free; the trade-off parameters α, β of MIC were selected from $[10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}]$; in FLSD, the candidate parameters λ_1 and λ_2 were tuned from the sets of $[10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}]$ and $[10^{0}, 10^{1}, 10^{2}, 10^{3}, 10^{4}]$, respectively; the ranges of α, β, γ in V^{3}H were defined as $[10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}]$; the two hyperparameters α, β in DAIMC were optimized from $[10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}, 10^{3}]$; λ, γ in IMSR were empirically selected from $[2^{-10}, 2^{-5}, 1, 2^{5}, 2^{10}]$; the λ and anchors numbers in IMVC-CBG vary from $[10^{-3}, 10^{-2}, 10^{0}, 10^{1}]$ and [k,2k,3k,5k] (k is the number of clusters), respectively.

D. Experimental Setting

There are three parameters λ_1, λ_2 and λ_3 in our SCBGL model. To find the optimal parameters, in the experiment, the parameters λ_1 and λ_2, λ_3 are tuned from the sets of $[10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}]$ and $[10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}]$, respectively. For simplicity, we set the number of global anchors parameter m equals to

[9]http://yann.lecun.com/exdb/mnist/

TABLE IV

CLUSTERING PERFORMANCE (PURITY%±STD%) ON TWELVE BENCHMARK DATASETS

Dataset	BSV	MIC	APMC	DAIMC	FLSD	IMSR	V^{3}H	IMVC-CBG	PSIMVC-PG	SCBGL
ORL	37.97±1.35	45.20±1.87	71.77±1.83	71.65±0.43	43.30±1.77	70.25±0.00	70.93±1.52	70.86±0.69	62.56±0.58	**85.50±0.11**
NGs	43.15±1.51	22.90±0.02	43.60±0.00	58.77±0.55	83.00±0.00	80.00±0.00	71.94±0.00	84.42±0.11	87.00±0.00	**90.40±0.00**
MSRC	26.86±0.85	53.62±2.43	24.58±0.56	65.79±0.61	58.19±1.62	52.48±0.02	71.52±0.93	67.31±1.31	63.81±0.00	**79.24±0.26**
WebKB	78.12±0.00	72.81±0.27	81.63±0.00	82.57 ±0.41	78.12±0.00	79.54±0.00	84.80±0.00	51.86±0.00	**92.29±0.00**	86.20±0.00
Caltech101-7	64.70±0.61	71.58±1.24	N/A	81.12±1.08	80.03±0.27	84.06±0.00	N/A	81.87±0.00	81.96±0.03	**84.33±0.06**
Caltech101-20	47.78±0.72	54.31±1.24	N/A	74.92±0.93	71.08±0.63	72.59±0.00	N/A	69.63±0.73	69.86±0.35	**75.36±0.05**
BDGP	36.75±0.89	34.32±0.28	44.96±0.00	36.50±0.96	44.73±0.06	37.56±0.00	40.32±0.05	42.93±0.00	44.33±0.20	**50.25±0.04**
NUSWIDE	13.72±0.04	N/A	N/A	**23.41±0.63**	N/A	N/A	N/A	22.17±0.19	20.79±0.12	22.63±0.17
Reuters	N/A	N/A	N/A	N/A	N/A	N/A	N/A	51.86±0.00	53.21±0.00	**54.38±0.00**
Cifar10	N/A	N/A	N/A	95.81±0.45	N/A	N/A	N/A	91.74±0.20	95.94±0.00	**96.47±0.00**
Cifar100	N/A	N/A	N/A	92.60±0.54	N/A	N/A	N/A	94.98±0.82	97.77±0.00	**99.17±0.44**
MNIST	N/A	N/A	N/A	97.57±0.31	N/A	N/A	N/A	98.10±0.16	98.24±0.00	**98.33±0.00**

TABLE V

CLUSTERING PERFORMANCE (FSCORE%±STD%) ON TWELVE BENCHMARK DATASETS

Dataset	BSV	MIC	APMC	DAIMC	FLSD	IMSR	V^{3}H	IMVC-CBG	PSIMVC-PG	SCBGL
ORL	11.15±1.59	15.08±0.88	54.59±1.49	56.42±0.72	23.17±1.93	54.95±0.00	55.08±2.01	51.38±0.23	45.98±1.04	**76.42±0.16**
NGs	32.39±1.08	32.95±0.02	33.74±0.00	44.64±0.34	69.83±0.00	66.83±0.00	57.34±0.00	71.15±0.18	76.31±0.00	**80.83±0.00**
MSRC	25.95±0.06	32.39±1.05	26.21±1.18	49.99±0.66	42.74±1.23	34.24±0.03	55.19±1.36	52.06±1.58	43.24±0.00	**64.83±0.29**
WebKB	60.55±1.28	68.24±0.29	66.83±0.00	79.46±0.43	63.85±0.00	73.40±0.00	77.39±0.00	39.72±0.00	65.21±0.09	**80.67±0.00**
Caltech101-7	56.03±0.02	37.74±1.35	N/A	49.44±2.93	55.63±0.03	45.05±0.00	N/A	60.61±1.27	52.48±0.05	**69.75±0.10**
Caltech101-20	32.34±0.32	23.78±1.38	N/A	40.34±2.41	38.07±1.61	32.92±0.00	N/A	42.25±1.76	39.01±1.45	**45.94±0.16**
BDGP	28.76±0.61	30.05±0.00	30.24±0.00	28.85±0.06	32.13±0.04	29.44±0.00	31.05±0.04	34.39±0.15	32.33±0.13	**37.35±0.03**
NUSWIDE	**10.95±0.00**	N/A	N/A	8.58±0.19	N/A	N/A	N/A	7.70±0.03	7.12±0.06	8.05±0.05
Reuters	N/A	N/A	N/A	N/A	N/A	N/A	N/A	39.72±0.00	39.71±0.00	**40.97±0.00**
Cifar10	N/A	N/A	N/A	92.16±0.68	N/A	N/A	N/A	92.60±0.22	92.13±0.00	**93.16±0.00**
Cifar100	N/A	N/A	N/A	90.82±0.94	N/A	N/A	N/A	90.87±2.97	97.12±0.00	**99.06±0.05**
MNIST	N/A	N/A	N/A	95.28±0.57	N/A	N/A	N/A	96.29±0.28	96.55±0.00	**96.72±0.00**

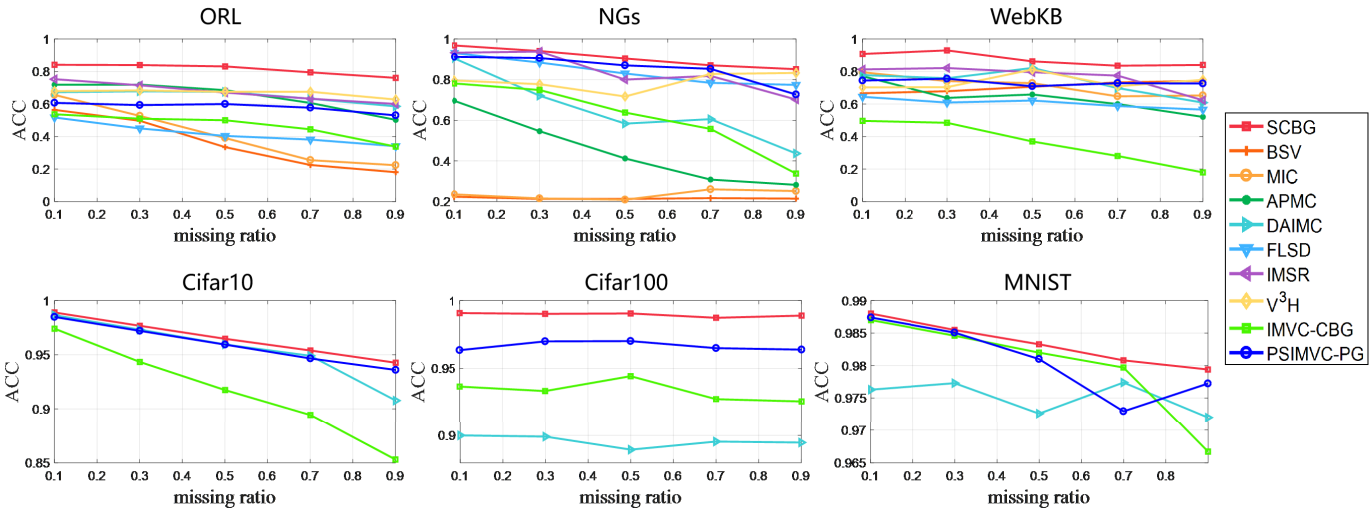


Fig. 2. The clustering performances of ACC on six incomplete datasets with different missing rates.

TABLE VI

CLUSTERING ACCURACY (%) ↑ AND RUNNING TIME (S) ↓ OF THE PROPOSED METHOD AND FSMSC. THE BOLD FORM MARKS THE BEST PERFORMANCE

datasets	running time (s)		ACC (%)		NMI (%)		Purity (%)		Fscore (%)	
	FSMSC	SCBGL	FSMSC	SCBGL	FSMSC	SCBGL	FSMSC	SCBGL	FSMSC	SCBGL
RGB-D	104.51	**1.42**	**44.93**	44.65	**37.82**	36.72	**55.83**	54.93	33.70	**34.27**
SUNRGBD	1299.11	**17.45**	14.18	**20.45**	9.26	**23.63**	16.67	**36.43**	12.08	**13.04**
Reuters	N/A	**580.11**	N/A	**48.06**	N/A	**31.06**	N/A	**54.51**	N/A	**41.32**
Cifar10	2124.03	**46.89**	99.15	**99.39**	97.66	**98.29**	99.15	**99.39**	98.32	**98.79**
Cifar100	5969.82	**92.30**	91.31	**99.50**	98.60	**99.88**	93.98	**99.57**	92.78	**99.50**
MNIST	2703.10	**39.95**	**99.07**	98.88	**97.11**	96.64	**99.07**	98.88	**98.13**	97.79

the number of intra-view anchors parameter k. Besides, we set the feature selection parameter of global anchor $p = 200$ to select representative and robust features. As for comparison algorithms, we follow the parameter settings of the public codes and employ the traversal strategy for their best results. To avoid randomness, we perform each competitor twenty

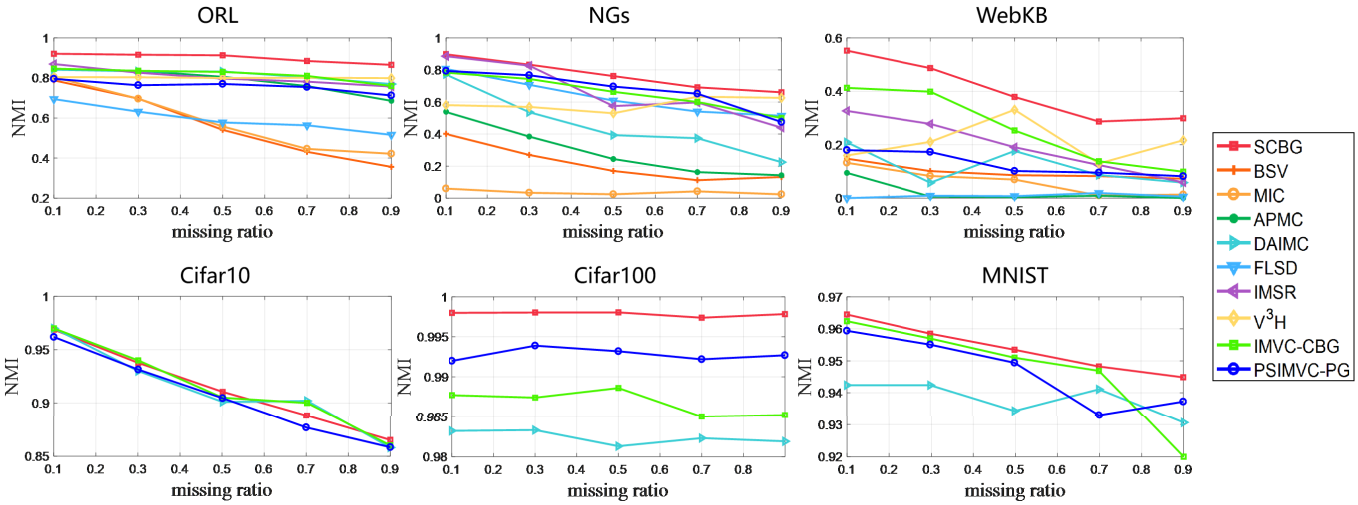


Fig. 3. The clustering performances of NMI on six incomplete datasets with different missing rates.

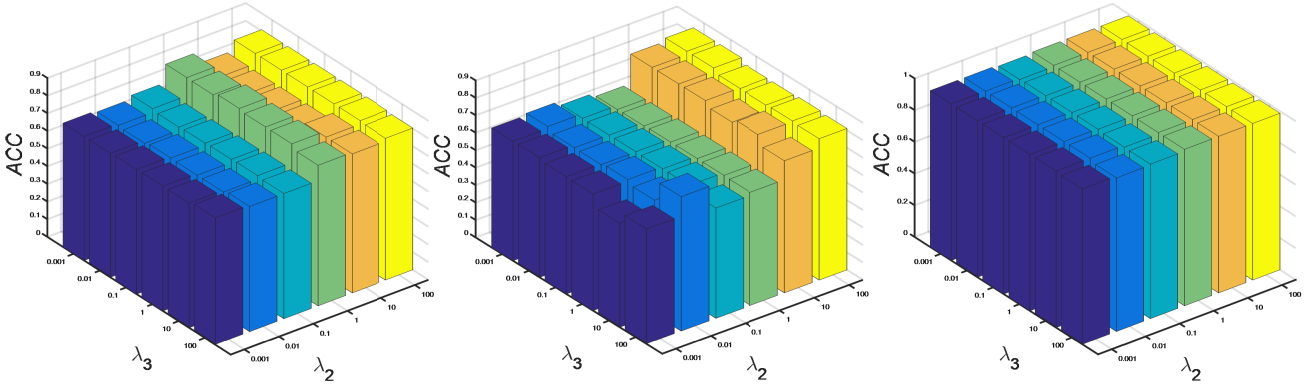


Fig. 4. Parameters tuning in terms of ACC on three incomplete datasets.

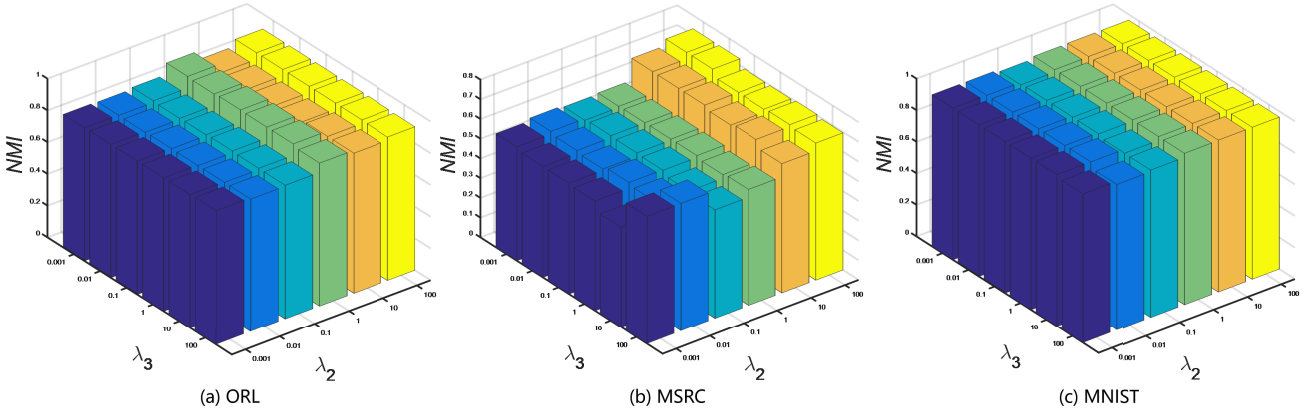


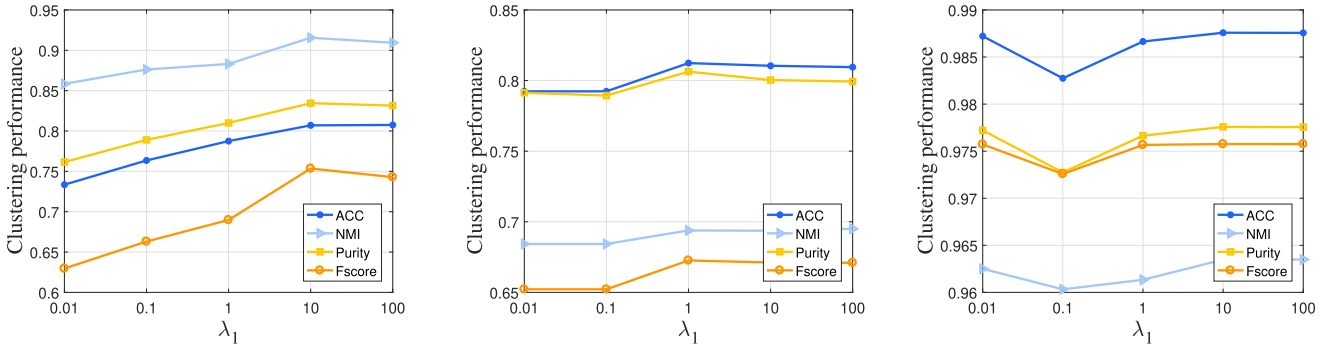
Fig. 5. Parameters tuning in terms of NMI on three incomplete datasets.

times and record the average value and corresponding standard deviations under the optimal parameters.

E. Experimental Results and Analysis

We show the performance on twelve benchmark datasets in Table II-V, and highlight the best performance in boldface and mark the suboptimal performance in underlining. Besides, the symbol 'N/A' implies the algorithm encounters an out-of-memory problem. From Table II-V, we draw the following observations:

1) Our proposed SCBGL reaches the best performance under most circumstances except for the suboptimal performance on NUSWIDE, which demonstrates its superiority in large-scale incomplete datasets. For example, SCBGL captures around 14.6%, 3.4%, 10.96%, 4.03%, 9.76%, 6.75%,7.17%,



(a) ORL (b) MSRC (c) MNIST

Fig. 6. The effect of parameter λ_1 on ORL, MSRC and MNIST.

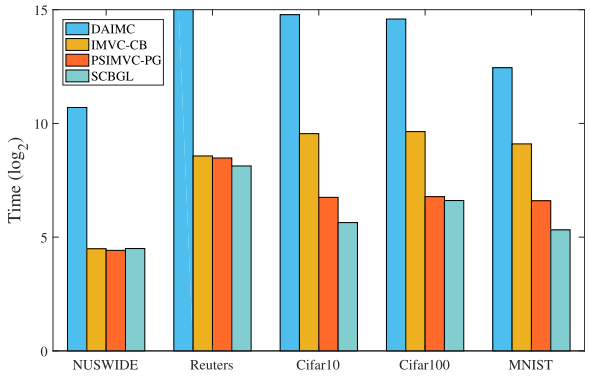


Fig. 7. Running time on five large-scale datasts.

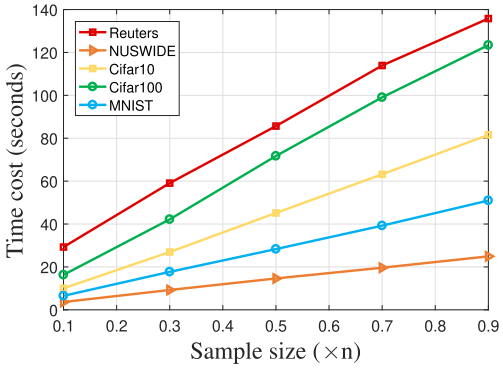
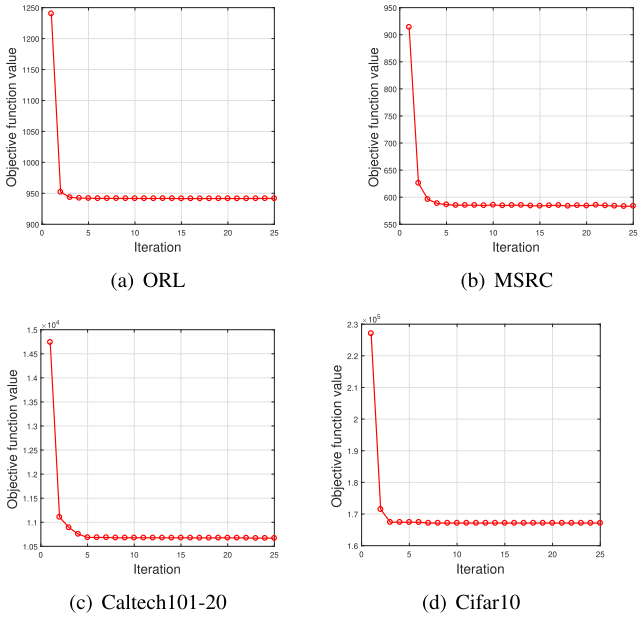


Fig. 8. Time cost on five large-scale datasets with different sample sizes.



(a) ORL (b) MSRC

(c) Caltech101-20 (d) Cifar10

Fig. 9. The objective function value of SCBGL on four incomplete datasets.

2.01%, 0.53%, 2.04% and 0.23% improvement in terms of ACC on all datasets except for NUSWIDE.

2) In comparison with the baseline method BSV, most IMVC methods have an overwhelming advantage, indicating that it is better to complete the missing features by diverse approaches instead of using zero or average value of all observed features.

3) As shown in Figs. 2-3, with the missing rate ranges from 0.1 to 0.9, the performance of these algorithms shows a monotonically decreasing trend.

4) Compared with IMVC-CBG, which makes the first effort to extend bipartite graph learning into IMVC domain and only focuses on inter-view consistency information, our method presents more capability in yielding remarkable clustering performance on incomplete datasets. This result may benefit from the fact that our method jointly explores the inter-view and intra-view information, which maximizes the preservation of the global correlation of multi-view data.

F. Model Discussion

1) Parameter Selection: Three trade-off parameters λ_1, λ_2 and λ_3 are employed in our model. For the balance parameters λ_2 and λ_3, we first fix parameter $\lambda_1 = 1$ and fine-tune them by searching the set $[0.001, 0.01, 0.1, 1, 10, 100]$. Figs. 4-5 show the performance of our algorithm varies with the parameters λ_2 and λ_3. We can observe that SCBGL is insensitive to λ_3. When λ_2 is greater than 1, our model can yield promising results, which indicates the effectiveness of inter-view bipartite graph learning. To illustrate the impact of consensus intra-view bipartite graph regularization term, we tune the parameter λ_1 from the sets of $[0.01, 0.1, 1, 10, 100]$. As shown in Fig.6, a lower value of λ_1 will lead to a poor clustering

result, and SCBGL enjoys more promising performance when $\lambda_1 \in [1, 100]$.

2) Running Time Analysis: Fig. 7 displays the running time of our SCBGL and competitors on five large-scale datasets ranging from 18758 to 60,000. We can clearly conclude that: 1) SCBGL takes substantially the shortest running time except for NUSWIDE, which implies SCBGL can remarkably decline the time and space complexity. 2) Due to the high dimensionality of Reuters, many experiments consume a long time to execute Reuters data and even suffer from memory shortage. In summary, SCBGL is both time and storage economical and is significantly superior to the other competitors.

Fig. 8 reports the time cost on five large-scale datasets with different sample size from 0.1 to 0.9 with 0.2 stepsize. Due to the high feature dimension of Reuters, its running time is high. To clearly reveal the conclusion of running time and sample size, we reduce the running time of Reuters by five. According to Fig. 8, the line of time spent is nearly straight, which indicate that the time cost increases essentially linearly with sample size. The results further verify that our proposed SCBGL can achieve a linear time complexity respecting to sample numbers.

3) Empirical Convergence: Fig. 9 displays the evolution of the objective function value on four incomplete datasets ORL, MSRC, Caltech101-20 and Cifar10. These convergence curves monotonically decrease with each iteration and converge to a stable point within five iterations, which indicates our method possesses a fast and stable convergence property.

G. Comparative Analysis With FSMSC

As a novel algorithm, FSMSC [57] also utilized anchors for bipartite graph learning, but there are still significant differences between FSMSC and our SCBGL: 1) FSMSC handles the complete multi-view clustering while our SCBGL method aims to solve the incomplete multiview clustering. This is the key difference between FSMSC and our SCBGL. 2) Although all of them construct the global feature $\widehat{\boldsymbol{X}}$, our SCBGL further considered the feature redundancy problem of $\widehat{\boldsymbol{X}}$, and hence our SCBGL method designed a projection matrix $\boldsymbol{F}$ to transform $\widehat{\boldsymbol{X}}$ into the latent feature, while FSMSC did not. Except for the redundancy reduction, another advantage of the projection matrix $\boldsymbol{F}$ is efficiency. Our SCBGL could flexibly handle the high-dimensional data and achieve faster clustering speed than FSMSC (See the compared clustering results in Table VI). 3) FSMSC introduced a mapping $\boldsymbol{Q}$ that fused the local representation and global representation by the Frobenius norm, while our SCBGL method could directly conduct the self-completed operation by regularizing the residual between the intra-view bipartite graph $\boldsymbol{Z}$ and the inter-view bipartite graph $\widehat{\boldsymbol{Z}}$ with $l_{2,1}$ norm.

Moreover, we also compared the proposed SCBGL with FSMSC on six multi-view datasets, and the results are summarized in Table VI. Please note that FSMSC aims to solve the complete multi-view clustering problem and cannot handle the incomplete multi-view clustering problem. We conducted experiments on complete datasets by setting $\boldsymbol{X}^v\boldsymbol{A}^v = \boldsymbol{X}^v$. According to Table VI, we can find that: 1) Our SCBGL method achieves a faster clustering speed than FSMSC on all the datasets. Especially, FSMSC would suffer from out-of-memory problem on Reuters, while our method still obtains promising performance on Reuters. This is because our method designed a projection matrix $\boldsymbol{F}$ for dimensionality reduction and fused the intra-view bipartite graph $\boldsymbol{Z}$ and the inter-view bipartite graph $\widehat{\boldsymbol{Z}}$ with $l_{2,1}$ norm for self-completion, which contributed to the efficiency and effectiveness of our method. 2) On most datasets, our SCBGL method achieves better clustering results in comparison with FSMSC. For example, our method improves 6.27%, 0.24%, and 8.19% with respect to ACC on SUNRGBD, Cifar10 and Cifar100, respectively. Although our method has inferior performance on RGB-D and MNIST datasets, the gap between our method and FSMSC is small. In summary, there exist significant differences between our work and FSMSC, further demonstrating the novelty and excellent performance of our approach.

V. Conclusion

In this paper, we proposed a novel fast IMVC method via constructing both intra-view and inter-view bipartite graphs simultaneously to perform dynamic anchor learning and self-completion of consistent bipartite graph. To the best of our knowledge, SCBGL makes the first attempt to address the large-scale IMVC by inter-view-tutord-intra-view learning. Specifically, SCBGL dynamically constructs the intra-view anchors and the corresponding consensus bipartite graph in the latent subspace. Meanwhile, to preserve global discriminative information, we concatenate all features and dynamically learn an inter-view bipartite graph with global anchors. Furthermore, we integrate these two bipartite graphs in a mutually reinforcing manner to recover the missing items embedded in consensus intra-view bipartite graph. Extensive experiments on benchmark incomplete datasets validate the efficiency and effectiveness of SCBGL over state-of-the-arts.

References

[1] Y. Wang, "Survey on deep multi-modal data analytics: Collaboration, rivalry, and fusion," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 17, no. 1s, pp. 1–25, Jan. 2021.

[2] P. Ji, T. Zhang, H. Li, M. Salzmann, and I. Reid, "Deep subspace clustering networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 30, 2017, pp. 24–33.

[3] S. Huang, Y. Ren, and Z. Xu, "Robust multi-view data clustering with multi-view capped-norm K-means," *Neurocomputing*, vol. 311, pp. 197–208, Oct. 2018.

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

[5] Y. Hu, Z. Song, B. Wang, J. Gao, Y. Sun, and B. Yin, "AKM^{3}C: Adaptive K-multiple-means for multi-view clustering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 11, pp. 4214–4226, Nov. 2021.

[6] Y. Wang, W. Zhang, L. Wu, X. Lin, M. Fang, and S. Pan, "Iterative views agreement: An iterative low-rank based structured optimization method to multi-view spectral clustering," in *Proc. Int. Joint Conf. Artif. Intell.*, 2016, pp. 2153–2159.

[7] C. Hong, J. Yu, J. Wan, D. Tao, and M. Wang, "Multimodal deep autoencoder for human pose recovery," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5659–5670, Dec. 2015.

[8] J. Yu, Y. Rui, and D. Tao, "Click prediction for web image reranking using multimodal sparse coding," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2019–2032, May 2014.

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

[10] Y. Chen, X. Xiao, and Y. Zhou, "Multi-view clustering via simultaneously learning graph regularized low-rank tensor representation and affinity matrix," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, Jul. 2019, pp. 1348–1353.

[11] F. Nie, J. Li, and X. Li, "Self-weighted multiview clustering with multiple graphs," in *Proc. 26th Int. Joint Conf. Artif. Intell.*, Aug. 2017, pp. 2564–2570.

[12] 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.

[13] 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.

[14] Y. Chen, X. Xiao, and Y. Zhou, "Jointly learning kernel representation tensor and affinity matrix for multi-view clustering," *IEEE Trans. Multimedia*, vol. 22, no. 8, pp. 1985–1997, Aug. 2020.

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

[16] Y. Chen, X. Xiao, C. Peng, G. Lu, and Y. Zhou, "Low-rank tensor graph learning for multi-view subspace clustering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 1, pp. 92–104, Jan. 2022.

[17] L. Xing, B. Chen, J. Wang, S. Du, and J. Cao, "Robust high-order manifold constrained low rank representation for subspace clustering," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 2, pp. 533–545, Feb. 2021.

[18] M. Yang, Y. Li, Z. Huang, Z. Liu, P. Hu, and X. Peng, "Partially view-aligned representation learning with noise-robust contrastive loss," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 1134–1143.

[19] C. Xu, Z. Guan, W. Zhao, H. Wu, Y. Niu, and B. Ling, "Adversarial incomplete multi-view clustering," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Aug. 2019, pp. 3933–3939.

[20] S.-Y. Li, Y. Jiang, and Z.-H. Zhou, "Partial multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, vol. 28, no. 1, 2014, pp. 1968–1974.

[21] W. Shao, L. He, and P. S. Yu, "Multiple incomplete views clustering via weighted nonnegative matrix factorization with $L_{2,1}$ regularization," in *Proc. Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2015, pp. 318–334.

[22] P. Rai, A. Trivedi, H. Daumé III, and S. L. DuVall, "Multiview clustering with incomplete views," in *Proc. Mach. Learn. Social Comput. Workshop*. Citeseer, 2010, pp. 1–7.

[23] X. Liu et al., "Multiple kernel kk-means with incomplete kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 5, pp. 1191–1204, May 2020.

[24] J. Wen, Y. Xu, and H. Liu, "Incomplete multiview spectral clustering with adaptive graph learning," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1418–1429, Apr. 2020.

[25] Y. Lin, Y. Gou, Z. Liu, B. Li, J. Lv, and X. Peng, "Completer: Incomplete multi-view clustering via contrastive prediction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 11174–11183.

[26] Y. Wang, D. Chang, Z. Fu, J. Wen, and Y. Zhao, "Incomplete multiview clustering via cross-view relation transfer," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 1, pp. 367–378, Jan. 2023.

[27] Y. Li, F. Nie, H. Huang, and J. Huang, "Large-scale multi-view spectral clustering via bipartite graph," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 2750–2756.

[28] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, and Z. Xu, "Large-scale multi-view subspace clustering in linear time," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 4, 2020, pp. 4412–4419.

[29] D. Cai and X. Chen, "Large scale spectral clustering via landmark-based sparse representation," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1669–1680, Aug. 2015.

[30] B. Yang, X. Zhang, Z. Lin, F. Nie, B. Chen, and F. Wang, "Efficient and robust MultiView clustering with anchor graph regularization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 9, pp. 6200–6213, Sep. 2022.

[31] B. Yang, X. Zhang, F. Nie, F. Wang, W. Yu, and R. Wang, "Fast multi-view clustering via nonnegative and orthogonal factorization," *IEEE Trans. Image Process.*, vol. 30, pp. 2575–2586, 2021.

[32] B. Yang, X. Zhang, B. Chen, F. Nie, Z. Lin, and Z. Nan, "Efficient correntropy-based multi-view clustering with anchor graph embedding," *Neural Netw.*, vol. 146, pp. 290–302, Feb. 2022.

[33] M. Hu and S. Chen, "Doubly aligned incomplete multi-view clustering," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 2262–2268.

[34] M. Hu and S. Chen, "One-pass incomplete multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, 2019, pp. 3838–3845.

[35] X. Liu et al., "Late fusion incomplete multi-view clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 10, pp. 2410–2423, Oct. 2019.

[36] J. Wen, Z. Zhang, Z. Zhang, L. Fei, and M. Wang, "Generalized incomplete multiview clustering with flexible locality structure diffusion," *IEEE Trans. Cybern.*, vol. 51, no. 1, pp. 101–114, Jan. 2021.

[37] W. Shao, L. He, C.-t. Lu, and P. S. Yu, "Online multi-view clustering with incomplete views," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2016, pp. 1012–1017.

[38] L. Li, Z. Wan, and H. He, "Incomplete multi-view clustering with joint partition and graph learning," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 589–602, Jan. 2023.

[39] M. Sun et al., "Scalable multi-view subspace clustering with unified anchors," in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 3528–3536.

[40] S. Wang et al., "Fast parameter-free multi-view subspace clustering with consensus anchor guidance," *IEEE Trans. Image Process.*, vol. 31, pp. 556–568, 2022.

[41] X. Li, H. Zhang, R. Wang, and F. Nie, "Multiview clustering: A scalable and parameter-free bipartite graph fusion method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 1, pp. 330–344, Jan. 2022.

[42] S. Wang et al., "Highly-efficient incomplete largescale multiview clustering with consensus bipartite graph," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 9766–9775.

[43] S. Hu, Z. Lou, and Y. Ye, "View-wise versus cluster-wise weight: Which is better for multi-view clustering?" *IEEE Trans. Image Process.*, vol. 31, pp. 58–71, 2022.

[44] T. Xu, X. Kong, Q. Shen, Y. Chen, and Y. Zhou, "Deep and low-rank quaternion priors for color image processing," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 33, no. 7, pp. 3119–3132, Jul. 2023.

[45] Y. Chen, S. Wang, X. Xiao, Y. Liu, Z. Hua, and Y. Zhou, "Self-paced enhanced low-rank tensor kernelized multi-view subspace clustering," *IEEE Trans. Multimedia*, vol. 24, pp. 4054–4066, 2022.

[46] G. A. Watson, "The solution of orthogonal Procrustes problems for a family of orthogonally invariant norms," *Adv. Comput. Math.*, vol. 2, no. 4, pp. 393–405, Sep. 1994.

[47] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for ℓ_1-minimization: Methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, Jan. 2008.

[48] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," 2010, *arXiv:1009.5055*.

[49] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 32–39.

[50] M. Lades et al., "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Comput.*, vol. 42, no. 3, pp. 300–311, Mar. 1993.

[51] F. Nie, G. Cai, J. Li, and X. Li, "Auto-weighted multi-view learning for image clustering and semi-supervised classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1501–1511, Mar. 2018.

[52] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2002, pp. 849–856.

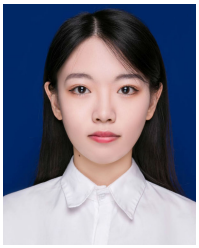
[53] J. Guo and J. Ye, "Anchors bring ease: An embarrassingly simple approach to partial multi-view clustering," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 118–125.

[54] X. Fang, Y. Hu, P. Zhou, and D. O. Wu, "V^{3}H: View variation and view heredity for incomplete multiview clustering," *IEEE Trans. Artif. Intell.*, vol. 1, no. 3, pp. 233–247, Dec. 2020.

[55] J. Liu et al., "Self-representation subspace clustering for incomplete multi-view data," in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 2726–2734.

[56] M. Li, S. Wang, X. Liu, and S. Liu, "Parameter-free and scalable incomplete multiview clustering with prototype graph," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 18, 2022, doi: 10.1109/TNNLS.2022.3173742.

[57] Z. Chen, X.-J. Wu, T. Xu, and J. Kittler, "Fast self-guided multi-view subspace clustering," *IEEE Trans. Image Process.*, early access, Mar. 29, 2023, doi: 10.1109/TIP.2023.3261746.



Xiaojia Zhao received the B.S. degree from the School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin, China, in 2021. She is currently pursuing the M.S. degree with the School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China. Her current research interests include subspace clustering and low-rank tensor methods for multi-view clustering.

Qiangqiang Shen received the B.S. and M.S. degrees from the Harbin Institute of Technology, Harbin, China, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree with the School of Electronics and Information Engineering, Harbin Institute of Technology, Shenzhen, China. His current research interests include pattern recognition and machine learning.

Yongyong Chen (Member, IEEE) received the B.S. and M.S. degrees from the Shandong University of Science and Technology, Qingdao, China, in 2014 and 2017, respectively, and the Ph.D. degree from the University of Macau, Macau, in 2020. He is currently an Assistant Professor with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China. He has authored or coauthored more than 50 research papers in top-tier journals and conferences, including IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING, and ACM MM. His current research interests include image processing, data mining, and computer vision.

Yongsheng Liang received the Ph.D. degree in mechanical engineering from the Harbin Institute of Technology, Harbin, China, in 1999. From 2002 to 2005, he was a Post-Doctoral Researcher of information and communication engineering with the Harbin Institute of Technology. Since 2017, he has been a Professor with the School of Electronics and Information Engineering, Harbin Institute of Technology, Shenzhen. He is the author of two books, more than 70 articles, and more than ten invention patents. His current research interests include video coding and transferring, neural networks, and deep learning.

Junxin Chen (Senior Member, IEEE) received the B.Sc. degree in communications engineering and the M.Sc. and Ph.D. degrees in communications and information systems from Northeastern University in 2007, 2009, and 2016, respectively. From 2019 to 2020, he was a Post-Doctoral Research Fellow with the Department of Computer and Information Science, University of Macau, under the UM Macau Talent Program (Class A). He was an Assistant Professor and an Associate Professor with the College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China. He is currently a Full Professor with the School of Software, Dalian University of Technology. He has authored/co-authored more than 70 scientific papers in international peer-reviewed journals and conferences, such as IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE INTERNET OF THINGS JOURNAL, IEEE NETWORK, and IEEE PHOTONICS JOURNAL. He has an H-index of 28 and a total of 2300 citations. His current research interests include the Internet of Medical Things, smart healthcare, artificial intelligence, and information security. He has received more than ten awards from Mainland China and Macau and has given more than 30 invited talks for universities and companies in China and Australia, such as the University of Macau and Deakin University. He is a Topic Editor of *Electronics*, the Leading Guest Editor of IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, and a Guest Editor of *Signal Processing: Image Communication*. He is a regular reviewer of extensive top-tier field journals.

Yicong Zhou (Senior Member, IEEE) received the B.S. degree in electrical engineering from Hunan University, Changsha, China, and the M.S. and Ph.D. degrees in electrical engineering from Tufts University, Medford, MA, USA.

He is currently a Professor with the Department of Computer and Information Science, University of Macau, Macau, China. His current research interests include image processing, computer vision, machine learning, and multimedia security.

Dr. Zhou is a fellow of the Society of Photo-Optical Instrumentation Engineers (SPIE) and was recognized as one of "World's Top 2% Scientists" on the Stanford University Releases List and one of "Highly Cited Researchers" in 2020. He received the Third Price of Macao Natural Science Award as a Sole Winner in 2020 and a co-recipient in 2014. He has been the Leading Co-Chair of Technical Committee on Cognitive Computing in the IEEE Systems, Man, and Cybernetics Society since 2015. He serves as an Associate Editor for IEEE TRANSACTIONS ON NEUTRAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and four other journals.