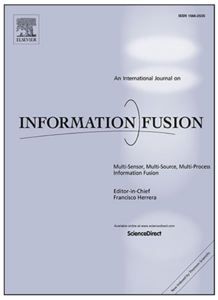
Journal Pre-proof



Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm

Haonan Xin, Zhezheng Hao, Zhensheng Sun, Rong Wang,
Zongcheng Miao, Feiping Nie

PII: S1566-2535(24)00261-6
DOI: https://doi.org/10.1016/j.inffus.2024.102483
Reference: INFFUS 102483

To appear in: *Information Fusion*

Received date : 15 January 2024
Revised date : 13 May 2024
Accepted date : 16 May 2024

Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm

Haonan Xin[a], Zhezheng Hao[a,b], Zhensheng Sun[c], Rong Wang[a,*], Zongcheng Miao[a,*] and Feiping Nie[a]

[a]School of Artifcial Intelligence, OPtics and ElectroNics (iOPEN), Northwestern Polytechnical University, Xi'an, Shaanxi 710072, P.R. China
[b]School of Cybersecurity, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, P.R. China
[c]Zhijian Laboratory, Rocket Force University of Engineering, Xi'an, Shaanxi 710025, P.R. China

ARTICLE INFO

Keywords:
graph sparsity
high-order graph
constrained $l_{1,2}$-norm
multi-view clustering

ABSTRACT

The graph-based multi-view clustering algorithms achieve decent clustering performance by consensus graph learning of the first-order graphs from different views. However, the first-order graphs are often sparse, lacking essential must-link information, which leads to suboptimal consensus graph. While high-order graphs can address this issue, a two-step strategy involving the selection of a fixed number of high-order graphs followed by their fusion may result in information loss or redundancy, restricting the exploration of high-order information. To address these challenges, we propose Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm (MoMvGC), which mitigates the impact of graph sparsity on multi-view clustering. By innovatively designing constrained $l_{1,2}$-norm, the model ingeniously integrates the selection of multi-order graphs and corresponding weight learning into a unified framework. Furthermore, MoMvGC not only enable sparse selection of multi-order graphs but also simultaneous selection of views. Afterwards, we design an efficient alternative optimization algorithm to solve the optimization problems in MoMvGC. The proposed model achieves state-of-the-art clustering performance on nine real-world datasets, with particularly notable improvements observed on the MSRC dataset, where the clustering accuracy is increased by 5.24% compared to the best baseline. Comprehensive experiments demonstrate the effectiveness and superiority of our model.

1. Introduction

With the advancements in multimedia and internet technologies, data often exhibits diverse characteristics that can be captured by different sensors or represented using multiple sets of features, commonly referred to as "views" [12]. The viewpoints from various media sources and the representations from distinct modalities concerning the same event can be regarded as different views. Harnessing insightful information from diverse views can foster the analysis and exploration of data, which is known as multi-view learning [45]. As a vital branch of multi-view learning and an extension of single-view clustering, multi-view clustering [40] has been extensively researched and used in Web ranking [27], community detection [41], and other areas [25, 5, 31].

Within the realm of multi-view clustering, graph-based multi-view clustering methods [20, 38] exhibit remarkable results compared to other multi-view clustering methods, such as matrix factorization methods [44], subspace learning-based methods [40, 13] and multiple kernel-based methods [34].The general workflow of graph-based multi-view clustering entails the construction of a dedicated graph for each view, followed by the integration of these graphs through techniques such as graph fusion [28, 9] or weighted aggregation [20, 22], ultimately yielding a consensus graph.

In the last decade, many graph-based multi-view clustering models are proposed by learning consensus embeddings or learning consensus graph [22, 39, 6, 14, 46, 26, 36, 42, 29]. And these algorithms have demonstrated impressive clustering capabilities. However, there remain several crucial aspects that require improvement. A prevalent procedure in such algorithms involves the generation of graphs from multi-view data, and these graphs is sparse. Sparse graphs mean that the edge density between nodes within clusters and nodes between clusters is small, and there is little must-link information, which greatly limits clustering performance. Fortunately, in the context of single-view clustering, AOPL proposed by Wu et al. successfully illustrates the effectiveness of high-order graph to solve the sparsity problem [37]. Inspired by this, we hope to extend this idea to multi-view clustering to further improve clustering performance by leveraging the complementarity and the consistency [15] of multi-view data.

However, the two-step strategy of selecting a fixed number of multi-order graphs and then performing graph fusion adopted by Wu et al., not only increases a hyperparameter, but may also cause information loss or redundancy, which hinders the mining of high-order information. Moreover, not every view is suitable for the consensus graph learning. Especially for views that are filled with noise, utilizing the corresponding multi-order graphs for graph clustering may not only fail to improve clustering accuracy but also result in a degradation of clustering performance. Therefore, in the context of multi-view and multi-order graph for graph clustering, adaptive selection is required not only for the multi-order graphs but also for the views so as to achieve effective consensus graph learning. In order to solve the above problems, we innovatively propose Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm (MoMvGC), which can fully mine and utilize multi-order

*Corresponding author
Email: wangrong07@tsinghua.org.cn (R. Wang);
miaozongcheng@nwpu.edu.cn (Z. Miao)

graph information from different views. Specifically, our contributions are summarized below:

- We extends the idea of high-order graphs to multi-view clustering. The MoMvGC model can fully excavate the rich structural information hidden in the high-order graphs, which eliminates the influence of graph sparsity on clustering performance.
- By leveraging the innovatively proposed constrained $l_{1,2}$-norm, the MoMvGC model integrates multi-order graph selection and weight learning into a unified framework, avoiding the impact on clustering accuracy caused by information loss or redundancy in the two-step strategy. Additionally, the MoMvGC model enables simultaneous selection of different views and multi-order graphs, thereby mitigating the negative influence of noise from both views and high-order graphs on the consensus graph learning.
- We have designed a comprehensive optimization framework for the proposed MoMvGC model. We conducted extensive and diverse experiments on nine real-world datasets. The model achieved SOTA clustering performance, providing compelling evidence for the effectiveness and superiority of our model.

Notations. For the matrix $\boldsymbol{E} \in \mathbb{R}^{p\times q}$, $\boldsymbol{E}^i$ and $\boldsymbol{E}_j$ are denotes as the i-th row and j-th column of $\boldsymbol{E}$, respectively. E_{ij} is the j-th element of $\boldsymbol{E}^i$. The Frobenius norm of $\boldsymbol{E}$ is $\|\boldsymbol{E}\|_F = \sqrt{\mathrm{Tr}(\mathbf{E}^T\mathbf{E})}$. The $l_{1,2}$-norm of $\boldsymbol{E}$ is $\|\boldsymbol{E}\|_{1,2} = \sqrt{\sum_{i=1}^{p}\|\boldsymbol{E}^i\|_1^2} = \sqrt{\sum_{i=1}^{p}(\sum_{j=1}^{q}|E_{ij}|)^2}$. All elements of the column vector $\mathbf{1}_m \in \mathbb{R}^{m\times 1}$ are one.

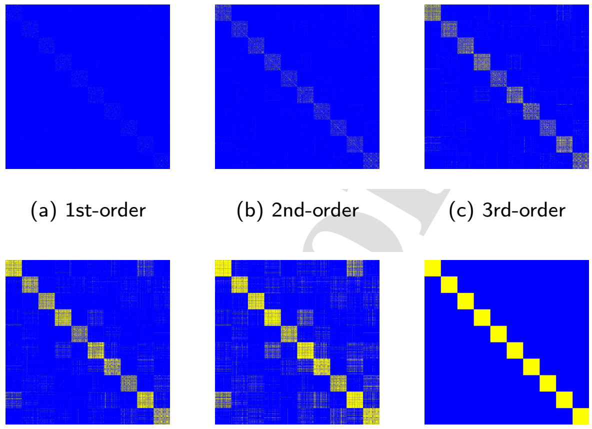
2. Related Work

2.1. Clustering with Adaptive Neighbors (CAN)

Given the data matrix $\boldsymbol{X} = [x_1, \cdots, x_n]^T \in \mathbb{R}^{n\times d}$, where n is the number of samples and d denotes the dimension of each sample. For any sample, the neighbors are defined as k samples closest to it in the Euclidean space. It can be considered that the probability of the i-th sample being adjacent to all other samples is denoted as $s_{ij}, j \in \{1, 2, \cdots, n\}$. A natural approach is that a larger $\|x_i - x_j\|_2^2$ corresponds to a smaller s_{ij}. So the objective of Clustering with Adaptive Neighbors (CAN) [23] is defined as

$$\min_{s_i} \sum_{i=1}^{n}\sum_{j=1}^{n}(\|x_i - x_j\|_2^2 s_{ij} + \sigma s_{ij}^2), \quad s.t.\ \forall i, \boldsymbol{s}^i\mathbf{1}_n = 1, \boldsymbol{s}_i \geq 0, \tag{1}$$

where $\boldsymbol{s}^i = [s_{i1}, s_{i2}, \cdots, s_{in}] \in \mathbb{R}^{1\times n}$, and the regular term $\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma s_{ij}^2$ is a uniform hypothesis, which ensures each data point has an equal likelihood of $\frac{1}{n}$ to be selected as the neighbor of $\boldsymbol{x}_i$. For the sake of making the learned similarity matrix $\boldsymbol{S}$ more suitable for clustering, a rank constraint about the Laplacian matrix is introduced on the basis of the



(a) 1st-order (b) 2nd-order (c) 3rd-order

(d) 4th-order (e) 5th-order (f) Ground truth

Figure 1: Visualization of high-order graphs and ground truth of Mfeat dataset.

problem (1), which allows $\boldsymbol{S}$ to be learned in the form of diagonal block, as follows

$$\min_{s_i} \sum_{i=1}^{n}\sum_{j=1}^{n}(\|x_i - x_j\|_2^2 s_{ij} + \sigma s_{ij}^2), \quad s.t.\ \forall i, \boldsymbol{s}^i\mathbf{1}_n = 1, \boldsymbol{s}_i \geq 0, rank(\boldsymbol{L}_s) = n - c, \tag{2}$$

where $\boldsymbol{L}_s = \boldsymbol{D}_s - \frac{\boldsymbol{S}^T+\boldsymbol{S}}{2}$, and $\boldsymbol{D}_s$ is a diagonal matrix where i-th diagonal value is $\sum_j(\frac{s_{ij}+s_{ji}}{2})$. After optimizing the problem (2) to obtain the optimal $\boldsymbol{S}$, the cluster labels are obtained directly without any post-processing.

2.2. High-order Graph

The concept of high-order graph is of great significance in the realm of graph learning. High-order graphs adhere to the principle that the neighbor of a neighbor is also a neighbor. LINE proposed by Tang et al. [30] integrates first-order graph and second-order graph for graph embedding learning, which excavates important structural information that is not easy to observe in first-order graph. Wu et al. expressed the concept of second-order graph in the form of matrix, and further proposed the definition of higher-order graphs [37]. Given the first-order graph $\boldsymbol{C} \in \mathbb{R}^{n\times n}$ i.e. the original graph, where n is the number of nodes, the k-th-order graph $\boldsymbol{C}^k$ is defined as

$$\boldsymbol{C}^k = \begin{cases} \boldsymbol{C}, & k = 1; \\ \boldsymbol{C}^{k-1}\boldsymbol{C}, & k > 1. \end{cases} \tag{3}$$

To clearly demonstrate the outstanding performance of high-order graphs, we have visualized the multi-order graphs of the Mfeat dataset, ranging from first-order to fifth-order graphs. In Figure 1, we observe that compared to the similarity graph corresponding to the ground truth, the first-order graph is significantly sparse and lacks essential must-link information. This largely restricts the performance of multi-view clustering methods. As the order of the graph increases, it becomes progressively denser. However, the number of cannot-link edges also increases, which can be regarded as a form of noise. So selecting graphs with too low or too high order is detrimental to graph learning. Notably, the third-order graph appears to be the most suitable for learning the consensus graph. Therefore, selecting appropriate multi-order graphs for graph fusion is of utmost importance for effective graph clustering.

2.3. Adaptive-order Proximity Learning (AOPL)

Dedicated to alleviating the impact of first-order graph sparsity on clustering performance, Wu et al. proposed a high-order graph clustering framework within a single-view context [38]. Specifically, the algorithm initially selects the optimal m multi-order graphs, followed by weight learning for these graphs and their fusion for consistent graph $\boldsymbol{S}$, which is essentially a two-step strategy. The objective function is defined as follows:

$$\begin{aligned} &\min_{\boldsymbol{S}} \sum_{k=1}^{K} q_k f(\|\boldsymbol{S}-\boldsymbol{C}_k\|_F^2), \\ &s.t.\ \boldsymbol{S}\mathbf{1}_n = \mathbf{1}_n, \boldsymbol{S} \geq 0, \boldsymbol{q}^T\mathbf{1}_K = m, \\ &\quad \boldsymbol{q} \in \{0,1\}^m, rank(\boldsymbol{L}_s) = n-c, \end{aligned} \tag{4}$$

where q_k is the boolean variable that determines whether to choose the k-th order graph $\boldsymbol{C}_k$, and $f(\|\boldsymbol{S}-\boldsymbol{C}_k\|_F^2)$ represents an approach to learn optimal consistent graph. Evidently, the hyperparameter m has a significant impact on the model, and manually adjusting m is highly subjective. When m is relatively large, it may lead to issues of information redundancy. When m is relatively small, it might result in insufficient exploration of high-order information, causing problems of information loss. Therefore, the adaptive selection of the number of multi-order graphs for graph fusion is very importance.

2.4. Related Work on Graph-based Multi-view Clustering

In recent years, the realm of graph-based multi-view clustering has witnessed a surge in activity, marked by the emergence of numerous high-quality algorithms. These algorithms fall into two main categories: one emphasizing the acquisition of consensus embedding and the other focusing on the acquisition of consensus graph. Effectively learning consensus embedding from diverse views involves judiciously amalgamating different representations to derive an embedding that best aligns with the real data distribution [39]. In addressing this challenge, Nie *et al*. proposed AMGL, an algorithm that adaptively learns a consensus spectral embedding from Laplacian matrices constructed across various views [21]. Furthermore, extending spectral clustering to multi-view scenarios, Hu *et al*. introduced a Re-Weighted framework [24] for adaptively learning a consistent non-negative embedding [7, 2]. Observing the differing relevance of embeddings across varied viewpoints in clustering tasks, Wang et al. formulated a globally-applied truncated weight allocation mechanism, thus amplifying the discriminative power of embeddings for clustering [33]. In response to the computational complexity entailed in embedding learning, He et al. proposed FAMvC, leveraging singular value decomposition (SVD) as an alternative to eigenvalue decomposition in spectral clustering [6], resulting in notable improvements in clustering performance.

In contrast to methods focusing on learning consensus embedding, approaches centered around learning consensus graphs often exhibit superior performance in multi-view clustering, fostering the development of numerous high-quality algorithms. Nie *et al*. proposed SwMC, which employs Laplacian rank constraints to learn a consensus graph [22]. Expanding upon this paradigm, a succession of related efforts [43, 32] has emerged, consistently showcasing commendable clustering accuracy.Recognizing the significance of initialization graphs for clustering performance, Tang *et al*. proposed CGD served as the first algorithm that enhances consensus graph quality via a diffusion process for learning improved graphs from different views [29]. Acknowledging that learning consensus graphs directly from the original data may lead to clustering performance instability, Zhao *et al*. proposed AONGR, a method that fuses spectral clustering and non-negative matrix decomposition into a unified framework for reconstructing the consensus graph [46]. While these models have achieved superior clustering performance, they predominantly prioritize consistency, overlooking the inherent diversity across views [10]. To address this limitation, Huang *et al*. proposed CDMGC, a representative work that decomposes the graph from each view into a consistent part and a divergent part, and then learns a high-quality consensus graph by fusing the consistent parts obtained from different views [10]. Furthermore, some multi-view clustering algorithms based on bipartite graphs have been proposed to address the issue of high computational complexity [26, 36, 14, 42]. These approaches typically generate a concise anchor set from the original data, serving to represent the overall structure of the sample set, which notably reduces the computational complexity, resulting in a linear reduction [17, 8, 4, 35].

However, the graph-based multi-view clustering algorithm mentioned above utilizes sparse graphs, as shown in Figure 1, thereby lacking a substantial amount of must-link information and consequently constraining clustering accuracy. So it is necessary to solve the problem for improving the clustering performance, which serves as the primary motivation of our research.

3. Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm (MoMvGC)

This section provides a detailed explanation of the Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm (MoMvGC) clustering model. We begin by presenting

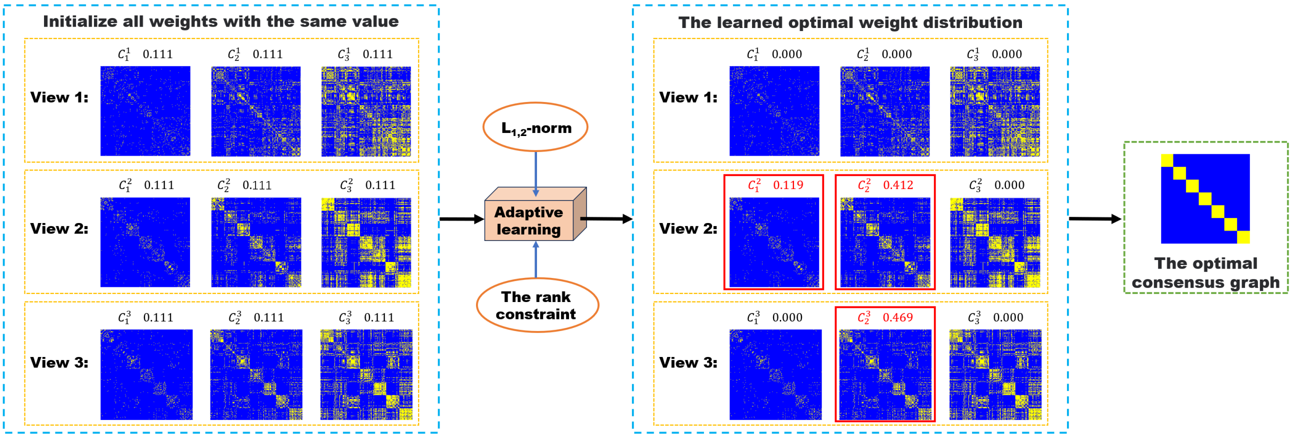


Figure 2: The framework of Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm.

the model along with its formulations, followed by the development of an efficient optimization algorithm.

3.1. The Proposed MoMvGC Model

Most graph-based multi-view clustering algorithms are susceptible to the influence of graph sparsity, which greatly hinders graph fusion. Inspired by AOPL, we introduce high-order graphs into multi-view clustering, aiming to fully exploit the rich structural information hidden in high-order graphs. Although the two-step strategy proposed by Wu et al. achieves decent clustering performance, the mandatory selection of a fixed number of multi-order graphs for graph fusion can potentially lead to information loss or redundancy, resulting in suboptimal consensus graphs. Apart from that, in real-world datasets, there are often certain views where the noise significantly outweighs the useful information. For such views, we have no choice but to discard them to avoid their negative impact on consensus graph learning. Therefore, it is imperative to not only select appropriate high-order graphs but also simultaneously choose the relevant views.

To address the aforementioned issues, we propose Multi-order Graph Clustering via Constrained $l_{1,2}$-norm (MoMvGC), which perfectly integrates the selection of multi-order graphs and corresponding weight learning into a unified framework, as shown in Figure 2. Specifically, the core idea is that the constrained $l_{1,2}$-norm, inspired by the feature selection model [18], is proposed innovatively to adaptively select multi-order graphs from different views. Simultaneously, the selected graphs are utilized for the consensus graph learning. The mutual enhancement between them enables the MoMvGC model to learn the optimal consensus graph.

Given multi-view data $\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_V]$ with V views, the first-order graph could be generated through the KNN algorithm [23], and then the K-order graphs for each view can be obtained by Eq. (3). $\boldsymbol{C}_k^v$ denotes k-th order graph of v-th view. $\boldsymbol{H} \in \mathbb{R}^{V\times K}$ is a coefficient matrix and H_{vk} represents the corresponding weight with respect to $\boldsymbol{C}_k^v$. The objective is to minimize the sum of squared errors between the consensus graph $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ and the multi-order graphs from all views, subject to the restriction of applying the constrained $l_{1,2}$-norm on the weight coefficients, which is defined as:

$$\begin{aligned} \min_{\boldsymbol{H},\boldsymbol{S}} & \sum_{v=1}^{V}\sum_{k=1}^{K} H_{vk}\|\boldsymbol{S}-\boldsymbol{C}_k^v\|_F^2 + \beta\|\boldsymbol{H}\|_{1,2}^2, \\ s.t.\ & \boldsymbol{S}\mathbf{1}_n = \mathbf{1}_n, \boldsymbol{S} \geq 0, \mathbf{1}_V^T\boldsymbol{H}\mathbf{1}_K = 1, \\ & \boldsymbol{H} \geq 0, rank(\boldsymbol{L}_s) = n - c, \end{aligned} \tag{5}$$

where β is the hyperparameter. On the one hand, the $l_{1,2}$-norm allows competition among different views, ensuring that each view has a fair chance to contribute. On the other hand, for the i-th view, it encourages sparsity in the i-th row of $\boldsymbol{H}$, meaning that non-zero weights tend to concentrate on few multi-order graphs. Hence, the adaptive learning of $\boldsymbol{H}$ ingeniously achieves the simultaneous selection of views and multi-order graphs. When β approaches 0, the graphs from only a few views participate in consensus graph learning. When β is sufficiently large, the selected multi-order graphs of all views contribute to graph fusion.

In order to ensure the learning of high-quality $\boldsymbol{S}$ and prevent the scenario where only one graph is selected for each view, it is necessary to introduce two regularization terms on the basis of Eq. (5). So the final objective of our proposed MoMvGC model is as detailed below

$$\begin{aligned} \min_{\boldsymbol{H},\boldsymbol{S}} & \sum_{v=1}^{V}\sum_{k=1}^{K} H_{vk}\|\boldsymbol{S}-\boldsymbol{C}_k^v\|_F^2 + \mu\|\boldsymbol{S}\|_F^2 \\ & + \beta(\|\boldsymbol{H}\|_{1,2}^2 + \alpha\|\boldsymbol{H}\|_F^2), \\ s.t.\ & \boldsymbol{S}\mathbf{1}_n = \mathbf{1}_n, \boldsymbol{S} \geq 0, \mathbf{1}_V^T\boldsymbol{H}\mathbf{1}_K = 1, \\ & \boldsymbol{H} \geq 0, rank(\boldsymbol{L}_s) = n - c, \end{aligned} \tag{6}$$

where μ and α are the hyperparameters. α balances the interaction of two norms, which is set to 1 in our model. Compared to the APOL algorithm, our proposed MoMvGC model eliminates the need for manual selection of a specific number of multi-order graphs. Instead, it intricately integrates the selection of multi-order graphs with weight learning, owing to the introduction and flexible application of the constrained $l_{1,2}$-norm.

3.2. Optimization

Since $\boldsymbol{L}_s$ is a positive semi-definite matrix, there exist n non-negative eigenvalues: $0 \leq \sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$. The number of multiplicities of zero as an eigenvalue of $\boldsymbol{L}_s$ is equal to connected component count of graph S [19]. It indicates that data are already clustered into c clusters when the constraint $\sum_{i=1}^{c} \sigma_i = 0$. According to Ky Fan's theory [3], we have

$$\sum_{i=1}^{c} \sigma_i = \min_{\boldsymbol{F}^T\boldsymbol{F}=\boldsymbol{I}} \operatorname{Tr}(\boldsymbol{F}^T \boldsymbol{L}_s \boldsymbol{F}), \tag{7}$$

where $\boldsymbol{F} \in \mathbb{R}^{n\times c}$ denotes the graph embedding. Thus, the problem (6) can be reformulated as

$$\begin{aligned} \min_{\boldsymbol{H},\boldsymbol{S},\boldsymbol{F}} & \sum_{v=1}^{V}\sum_{k=1}^{K} H_{vk}\|\boldsymbol{S}-\boldsymbol{C}_k^v\|_F^2 + \mu\|\boldsymbol{S}\|_F^2 \\ & +\beta(\|\boldsymbol{H}\|_{1,2}^2 + \|\boldsymbol{H}\|_F^2) + 2\lambda \operatorname{Tr}(\boldsymbol{F}^T\boldsymbol{L}_s\boldsymbol{F}), \\ & s.t.\ \boldsymbol{S}\boldsymbol{1}_n = \boldsymbol{1}_n, \boldsymbol{S} \geq 0, \boldsymbol{1}_V^T\boldsymbol{H}\boldsymbol{1}_K = 1, \boldsymbol{H} \geq 0, \\ & \boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{I}, \boldsymbol{L}_s = \boldsymbol{D}_s - \frac{\boldsymbol{S}^T+\boldsymbol{S}}{2}. \end{aligned} \tag{8}$$

Step 1: Update $\boldsymbol{S}$ with fixed $\boldsymbol{H}$ and $\boldsymbol{F}$: Let $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{D}(i,j) = \lambda\|\boldsymbol{f}^i - \boldsymbol{f}^j\|_2^2$, where $\boldsymbol{f}^i \in \mathbb{R}^{1\times c}$ is i-th row of $\boldsymbol{F}$. The problem (8) can be simplified to

$$\begin{aligned} \min_{\boldsymbol{S}\boldsymbol{1}_n=\boldsymbol{1}_n,\boldsymbol{S}\geq 0} & \sum_{v=1}^{V}\sum_{k=1}^{K} H_{vk}(\operatorname{Tr}(\boldsymbol{S}^T\boldsymbol{S}) \\ & -2\operatorname{Tr}(\boldsymbol{S}^T\boldsymbol{C}_k^v)) + \operatorname{Tr}(\boldsymbol{S}^T\boldsymbol{D}). \end{aligned} \tag{9}$$

For any two rows of $\boldsymbol{S}$, the constraints applied to them are mutually unrelated. So the problem (9) can be further reduced to

$$\min_{\boldsymbol{S}^i\boldsymbol{1}_n=1,\boldsymbol{S}^i\geq 0} \frac{1}{2}\|\boldsymbol{S}^i + \frac{1}{2(\mu+1)}\boldsymbol{M}^i\|_2^2, \tag{10}$$

where $\boldsymbol{M} = \boldsymbol{D} - 2\sum_{v=1}^{V}\sum_{k=1}^{K} H_{vk}\boldsymbol{C}_k^v$. Using the Lagrange multiplier method for the problem (10), we have

$$\mathcal{L} = \frac{1}{2}\|\boldsymbol{S}^i + \frac{1}{2(\mu+1)}\boldsymbol{M}^i\|_2^2 - \eta_i(\boldsymbol{S}^i\boldsymbol{1}_n - 1) - \boldsymbol{S}^i\theta_i, \tag{11}$$

where η_i and θ_i are the lagrangian multiplier. The closed-form solution of $\boldsymbol{S}$ is obtained through the Karush Kuhn-Tucker (KKT) condition [1, 23].

Step 2: Update $\boldsymbol{H}$ with fixed $\boldsymbol{S}$ and $\boldsymbol{F}$: The problem (8) can be written as

$$\min_{\boldsymbol{1}_V^T\boldsymbol{H}\boldsymbol{1}_K=1,\boldsymbol{H}\geq 0} \sum_{v=1}^{V}\sum_{k=1}^{K} H_{vk}P_{vk} + \beta(\|\boldsymbol{H}\|_{1,2}^2 + \|\boldsymbol{H}\|_F^2), \tag{12}$$

where $\boldsymbol{P} \in \mathbb{R}^{V\times K}$ and $P_{vk} = \|\boldsymbol{S} - \boldsymbol{C}_k^v\|_F^2$. The problem (12) could be solved through the QP solver.

Algorithm 1 The procedure to solve problem (8)

Input: The order K, the first-order graphs, the number of cluster c, the parameters μ, β and λ

Output: The consensus graph $\boldsymbol{S}$

1: Obtain multi-order graphs of each view by Eq. (1). Initialize $H_{vk} = \frac{1}{VK}$ and $\boldsymbol{F}$ with spectral embedding of the graph $\boldsymbol{S} = \frac{1}{V}\sum_{v=1}^{V}\boldsymbol{C}_1^v$, respectively.
2: **while** not converge **do**
3: Update $\boldsymbol{S}$ by Eq. (11);
4: Update $\boldsymbol{H}$ by Eq. (15);
5: Update $\boldsymbol{F}$ by Eq. (16);
6: **end while**
7: **return** $\boldsymbol{S}$

Theorem 1. *For the matrix $\boldsymbol{Z} \in \mathbb{R}^{r\times s}$ where $\boldsymbol{Z} \geq 0$, $\|\boldsymbol{Z}\|_{1,2}^2 = \tilde{\boldsymbol{Z}}^T\boldsymbol{J}\tilde{\boldsymbol{Z}}$. $\tilde{\boldsymbol{Z}} = vec(\boldsymbol{Z}) = [\boldsymbol{Z}_{11}, \ldots, \boldsymbol{Z}_{1s}, \ldots, \boldsymbol{Z}_{r1}, \ldots, \boldsymbol{Z}_{rs}]^T$, where $vec(\cdot)$ is a matrix-vector operator. $\boldsymbol{J}$ is a block diagonal matrix consisting of r identical matrices $\boldsymbol{T} \in \mathbb{R}^{s\times s}$, whose elements are all 1.*

Proof. Given a vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$, we have $(x_1 + x_2 + \cdots + x_n)^2 = \boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}$, where $\boldsymbol{G} \in \mathbb{R}^{n\times n}$ and its elements are all 1. Similarly, for the matrix $\boldsymbol{Z}$, we have

$$\|\boldsymbol{Z}\|_{1,2}^2 = \sum_{i=1}^{r}\|\boldsymbol{Z}^i\|_1^2 = \sum_{i=1}^{r}\boldsymbol{Z}^i\boldsymbol{T}\boldsymbol{Z}^{iT} = \tilde{\boldsymbol{Z}}^T\boldsymbol{J}\tilde{\boldsymbol{Z}}, \tag{13}$$

where $\boldsymbol{Z}^i$ is i-th row of $\boldsymbol{Z}$. Eq. (13) satisfies the second equality when $\boldsymbol{Z} \geq 0$. Taking $\boldsymbol{Z} \in \mathbb{R}^{2\times 2}$ as an example, there is

$$\boldsymbol{J} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{14}$$

□

Let $\boldsymbol{U} \in \mathbb{R}^{VK\times VK}$ be a block diagonal matrix composed of V identical matrices with dimensions K by K like $\boldsymbol{J}$. $\tilde{\boldsymbol{H}} = vec(\boldsymbol{H})$ and $\tilde{\boldsymbol{P}} = vec(\boldsymbol{P})$. So the problem (12) could be converted to

$$\min_{\boldsymbol{1}_V^T\tilde{\boldsymbol{H}}\boldsymbol{1}_K=1,\tilde{\boldsymbol{H}}\geq 0} \beta\tilde{\boldsymbol{H}}^T\boldsymbol{R}\tilde{\boldsymbol{H}} + \tilde{\boldsymbol{P}}^T\tilde{\boldsymbol{H}}, \tag{15}$$

where $\boldsymbol{R} = \boldsymbol{U} + \boldsymbol{I}$ and $\boldsymbol{I} \in \mathbb{R}^{VK\times VK}$ is an identity matrix. Obviously, the problem (15) can be solved by the QP solver.

Step 3: Update $\boldsymbol{F}$ with fixed $\boldsymbol{S}$ and $\boldsymbol{H}$: Based on the above analysis, the problem (8) becomes

$$\min_{\boldsymbol{F}^T\boldsymbol{F}=\boldsymbol{I}} \operatorname{Tr}(\boldsymbol{F}^T\boldsymbol{L}_s\boldsymbol{F}), \tag{16}$$

so the optimal $\boldsymbol{F}$ is the eigenvector corresponding to the first c minimum eigenvalues of $\boldsymbol{L}_s$.

3.3. Complexity Analysis

Time complexity The MoMvGC model necessitates the construction of $V \times (K-1)$ high-order graphs from V initial graphs. The construction of each graph requires a complexity of $\mathcal{O}(n|\mathcal{E}|)$, where $|\mathcal{E}|$ denotes the number of edges. Due to the sparsity of initial graphs, $|\mathcal{E}|$ and n are comparable in magnitude. Then the time complexity for this stage is $\mathcal{O}(V(K-1)n|\mathcal{E}|)$. In the process of updating $\boldsymbol{S}$, first computing $\boldsymbol{D}$ requires $\mathcal{O}(cn^2)$, and then computing $\boldsymbol{S}$ requires $\mathcal{O}(n^2)$. Updating $\boldsymbol{H}$ only requires a quadratic programming, which is independent of n and can be negligible. The computation of the optimal $\boldsymbol{F}$ requires a complexity of $\mathcal{O}(cn^2)$. Overall, the MoMvGC model exhibits a quadratic time complexity, placing it on par with other graph-based multi-view clustering algorithms.

Space complexity During the learning procedure, the major memory costs of our model are matrices $\boldsymbol{S} \in \mathbb{R}^{n\times n}$, $\boldsymbol{H} \in \mathbb{R}^{VK\times VK}$ and $\boldsymbol{F} \in \mathbb{R}^{n\times c}$. So the space complexity of MoMvGC is $\mathcal{O}(n^2 + V^2K^2 + cn)$. Since $c \ll n$, $K \ll n$ and $V \ll n$, the overall space complexity of our model is approximately $\mathcal{O}(n^2)$.

4. Experiments

In this section, we employ real-world datasets to verify the clustering performance of MoMvGC from different aspects such as clustering performance, sensitivity analysis, and convergence analysis.

4.1. Experimental Settings

4.1.1. Dataset Descriptions

We utilize nine real-world benchmark datasets to assess the quality and effectiveness of our clustering model. *a*) The face dataset *Yale*[1] contains 15 people or 15 classes, each with 11 face images with different expressions, postures and lighting, for a total of 165 images. *b*) *3Sources*[2] comprises news articles gathered from three distinct sources: BBC, The Guardian, and Reuters. It includes a total of 169 news items, each reported by all three media outlets, and classified into six categories. *c*) *WebKB*[3] is a dataset that contains webpages characterized by two aspects: content and links. These webpages are gathered from four universities, with a specific focus on the data from the University of Wisconsin. *d*) *MSRC*[4] encompasses 210 images capturing objects across 7 distinct classes and is presented from 5 different viewpoints. The classes include airplane, bicycle, building, car, cow, face, and tree. The dataset is characterized by 5 views, each represented by a unique feature set: 24-D color moment (CM), 576-D histogram of oriented gradient (HOG), 512-D GIST, 256 local binary pattern (LBP), and 254-D centrist (CENT). *e*) *YaleB*[5] comprises a total of 2,414 face images taken from 38 individuals. Each person is represented by 65 images captured under varying illuminations. In our experiments, we focus on a subset of the dataset, specifically the first 10 individuals, which amounts to a total of 650 images. *f*) *COIL-20*[6] is a commonly used dataset in the fields of image processing and pattern recognition. This dataset comprises images of 20 objects with 3 views, with each object having 72 images captured under various angles and lighting conditions. *g*) Object recognition dataset *Caltech101*[7] has 101 categories, whose number of views is 6. *h*) The *Mfeat*[8] dataset, sourced from the UCI repository, is a collection of handwritten digits (0-9). This dataset comprises 2000 samples, with each sample characterized by six different types of features. *i*) The dataset *Scene*[9] comprises a total of 2688 images divided into 8 distinct groups. For each image, we extract four distinct feature vectors, namely a 512-D GIST feature vector, a 432-D color moment feature vector, a 256-D HOG feature vector, and a 48-D local binary pattern (LBP) feature vector.

4.1.2. Comparison Models

We assess the performance of our MoMvGC model by benchmarking it against nine existing graph-based multi-view clustering algorithms. We acquired these implementations either from the authors' official websites or by direct communication with the authors. Specifically, the comparison models and the corresponding brief descriptions are listed as follows.

- SwMC [22] is a classical graph-based multi-view clustering algorithm that learns a consistent structured graph by imposing Laplacian rank constraints. The computational complexity is $O(cn^2)$.
- MVGL [43] obtains a global consistent structured graph by adaptively learning the optimal weight of each sample in different views. The computational complexity is $O(cn^2)$.
- GMC [32] gets the optimal graph fusion by simultaneously learning the optimal similarity matrix for each view and adaptively weighting each view. The computational complexity is $O(cn^2)$.
- CGD [28] is the first time to use the fusion process to learn an improved graph for each view, easing the dependence on high-quality initial graphs. The computational complexity is $O(n^3)$.
- CoMSC [16] utilize eigendecomposition to acquire a robust data representation characterized by low redundancy, which could help obtain better clustering results. The computational complexity is $O(cn^2)$.
- SMVSC [26] unifies anchor learning and graph construction within a single optimization framework, making the learned anchors are capable of more

[1] http://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html
[2] http://mlg.ucd.ie/datasets/3sources.html
[3] http://lig-membres.imag.fr/grimal/data.html
[4] https://www.microsoft.com/en-us/research/project/
[5] http://vision.ucsd.edu/leekc/ExtYaleDatabase/ExtYaleB.html
[6] http://cs.columbia.edu/CAVE/software/softlib/
[7] http://www.vision.caltech.edu/datasets/
[8] https://archive.ics.uci.edu/ml/datasets/Multiple+Features
[9] https://figshare.com/articles/dataset/15-Scene_Image_Dataset/7007177

Table 1
Clustering performance comparison. Best results are in bold, and the second best results are underlined.

Dataset	Metric	SwMC	MVGL	GMC	CGD	CoMSC	SMVSC	CDMGC	MSCTM	UDBGL	MoMvGC
Yale	ACC	$0.6606_{\pm 0.00}$	$0.6000_{\pm 0.00}$	$0.654_{\pm 0.00}$	$0.456_{\pm 0.03}$	$0.2897_{\pm 0.00}$	$0.4939_{\pm 0.02}$	$0.6527_{\pm 0.00}$	$0.6000_{\pm 0.00}$	$0.5454_{\pm 0.00}$	**$0.7333_{\pm 0.00}$**
	NMI	$0.6687_{\pm 0.00}$	$0.6363_{\pm 0.00}$	$0.6735_{\pm 0.00}$	$0.5348_{\pm 0.00}$	$0.3555_{\pm 0.01}$	$0.5374_{\pm 0.02}$	$0.6793_{\pm 0.00}$	$0.5820_{\pm 0.00}$	$0.6125_{\pm 0.00}$	**$0.6969_{\pm 0.00}$**
	ARI	$0.4159_{\pm 0.00}$	$0.4146_{\pm 0.00}$	$0.4410_{\pm 0.00}$	$0.3008_{\pm 0.04}$	$0.0718_{\pm 0.00}$	$0.3017_{\pm 0.02}$	$0.4283_{\pm 0.00}$	$0.3393_{\pm 0.00}$	$0.3761_{\pm 0.00}$	**$0.4775_{\pm 0.00}$**
	F-Score	$0.4568_{\pm 0.00}$	$0.4538_{\pm 0.00}$	$0.4800_{\pm 0.00}$	$0.3567_{\pm 0.04}$	$0.1321_{\pm 0.00}$	$0.3506_{\pm 0.02}$	$0.4679_{\pm 0.00}$	$0.3839_{\pm 0.00}$	$0.4182_{\pm 0.00}$	**$0.5127_{\pm 0.00}$**
3Sources	ACC	$0.6686_{\pm 0.00}$	$0.7514_{\pm 0.00}$	$0.6923_{\pm 0.00}$	**$0.7869_{\pm 0.00}$**	$0.5479_{\pm 0.02}$	$0.3082_{\pm 0.01}$	$0.3538_{\pm 0.00}$	$0.5266_{\pm 0.00}$	$0.4082_{\pm 0.00}$	$0.7573_{\pm 0.00}$
	NMI	$0.5382_{\pm 0.00}$	$0.5990_{\pm 0.00}$	$0.5479_{\pm 0.00}$	**$0.6985_{\pm 0.00}$**	$0.3719_{\pm 0.01}$	$0.0832_{\pm 0.00}$	$0.0676_{\pm 0.00}$	$0.4343_{\pm 0.00}$	$0.1995_{\pm 0.00}$	$0.6626_{\pm 0.00}$
	ARI	$0.4168_{\pm 0.00}$	$0.5244_{\pm 0.00}$	$0.4431_{\pm 0.00}$	**$0.6282_{\pm 0.00}$**	$0.2836_{\pm 0.02}$	$0.0341_{\pm 0.00}$	$0.0128_{\pm 0.00}$	$0.2605_{\pm 0.00}$	$0.1325_{\pm 0.00}$	$0.5699_{\pm 0.00}$
	F-Score	$0.5813_{\pm 0.00}$	$0.6573_{\pm 0.00}$	$0.6046_{\pm 0.00}$	**$0.7234_{\pm 0.00}$**	$0.4797_{\pm 0.02}$	$0.2342_{\pm 0.01}$	$0.3566_{\pm 0.00}$	$0.4802_{\pm 0.00}$	$0.3213_{\pm 0.00}$	$0.6764_{\pm 0.00}$
WebKB	ACC	$0.0197_{\pm 0.00}$	$0.5172_{\pm 0.00}$	$0.7586_{\pm 0.00}$	$0.7635_{\pm 0.00}$	$0.5468_{\pm 0.00}$	$0.5985_{\pm 0.01}$	$0.0197_{\pm 0.00}$	$0.7931_{\pm 0.00}$	$0.6945_{\pm 0.00}$	**$0.8177_{\pm 0.00}$**
	NMI	$0.2024_{\pm 0.00}$	$0.0614_{\pm 0.00}$	$0.3634_{\pm 0.00}$	$0.3269_{\pm 0.00}$	$0.0384_{\pm 0.00}$	$0.3072_{\pm 0.04}$	$0.2024_{\pm 0.00}$	**$0.5076_{\pm 0.00}$**	$0.1679_{\pm 0.00}$	$0.4799_{\pm 0.00}$
	ARI	$0.000_{\pm 0.00}$	$0.0269_{\pm 0.00}$	$0.4207_{\pm 0.00}$	$0.4702_{\pm 0.00}$	$0.0359_{\pm 0.00}$	$0.3148_{\pm 0.06}$	$0.000_{\pm 0.00}$	$0.5755_{\pm 0.00}$	$0.2507_{\pm 0.00}$	**$0.6122_{\pm 0.00}$**
	F-Score	$0.000_{\pm 0.00}$	$0.5288_{\pm 0.00}$	$0.6894_{\pm 0.00}$	$0.7001_{\pm 0.00}$	$0.5627_{\pm 0.00}$	$0.5548_{\pm 0.00}$	$0.000_{\pm 0.00}$	$0.7648_{\pm 0.00}$	$0.6095_{\pm 0.00}$	**$0.7855_{\pm 0.00}$**
MSRC	ACC	$0.7666_{\pm 0.00}$	$0.8714_{\pm 0.00}$	$0.7476_{\pm 0.00}$	$0.8238_{\pm 0.00}$	$0.4123_{\pm 0.00}$	$0.8190_{\pm 0.00}$	$0.7452_{\pm 0.01}$	$0.3523_{\pm 0.00}$	$0.7904_{\pm 0.00}$	**$0.9238_{\pm 0.00}$**
	NMI	$0.7537_{\pm 0.00}$	$0.7731_{\pm 0.00}$	$0.7421_{\pm 0.00}$	$0.7314_{\pm 0.00}$	$0.3463_{\pm 0.00}$	$0.7118_{\pm 0.00}$	$0.6879_{\pm 0.00}$	$0.2724_{\pm 0.00}$	$0.6722_{\pm 0.00}$	**$0.8535_{\pm 0.00}$**
	ARI	$0.6662_{\pm 0.00}$	$0.7152_{\pm 0.00}$	$0.6399_{\pm 0.00}$	$0.6641_{\pm 0.00}$	$0.2090_{\pm 0.00}$	$0.6486_{\pm 0.00}$	$0.5635_{\pm 0.02}$	$0.1456_{\pm 0.00}$	$0.6242_{\pm 0.00}$	**$0.8308_{\pm 0.00}$**
	F-Score	$0.7185_{\pm 0.00}$	$0.7560_{\pm 0.00}$	$0.6968_{\pm 0.00}$	$0.7123_{\pm 0.00}$	$0.3249_{\pm 0.00}$	$0.6987_{\pm 0.00}$	$0.6333_{\pm 0.01}$	$0.2845_{\pm 0.00}$	$0.6782_{\pm 0.00}$	**$0.8545_{\pm 0.00}$**
YaleB	ACC	$0.4784_{\pm 0.00}$	$0.3738_{\pm 0.00}$	$0.4338_{\pm 0.00}$	$0.3280_{\pm 0.00}$	$0.5647_{\pm 0.00}$	$0.3016_{\pm 0.01}$	$0.3963_{\pm 0.02}$	$0.4646_{\pm 0.00}$	$0.1630_{\pm 0.00}$	**$0.5769_{\pm 0.00}$**
	NMI	$0.4340_{\pm 0.00}$	$0.3391_{\pm 0.00}$	$0.4162_{\pm 0.00}$	$0.3085_{\pm 0.00}$	**$0.5433_{\pm 0.00}$**	$0.2282_{\pm 0.01}$	$0.3415_{\pm 0.02}$	$0.4479_{\pm 0.00}$	$0.0813_{\pm 0.00}$	$0.5395_{\pm 0.00}$
	ARI	$0.1830_{\pm 0.00}$	$0.0891_{\pm 0.00}$	$0.1571_{\pm 0.00}$	$0.1287_{\pm 0.00}$	**$0.3174_{\pm 0.00}$**	$0.0702_{\pm 0.00}$	$0.1161_{\pm 0.02}$	$0.2255_{\pm 0.00}$	$0.0194_{\pm 0.00}$	$0.3023_{\pm 0.00}$
	F-Score	$0.2899_{\pm 0.00}$	$0.2113_{\pm 0.00}$	$0.2651_{\pm 0.00}$	$0.2250_{\pm 0.00}$	**$0.3936_{\pm 0.00}$**	$0.1838_{\pm 0.00}$	$0.2439_{\pm 0.01}$	$0.3224_{\pm 0.00}$	$0.1191_{\pm 0.00}$	$0.3854_{\pm 0.00}$
COIL-20	ACC	$0.8541_{\pm 0.00}$	$0.7845_{\pm 0.00}$	$0.7909_{\pm 0.00}$	$0.7918_{\pm 0.00}$	$0.6091_{\pm 0.03}$	$0.5735_{\pm 0.03}$	$0.8666_{\pm 0.00}$	$0.8354_{\pm 0.00}$	$0.5784_{\pm 0.00}$	**$0.8986_{\pm 0.00}$**
	NMI	$0.9428_{\pm 0.00}$	$0.9130_{\pm 0.00}$	$0.9189_{\pm 0.00}$	$0.8786_{\pm 0.00}$	$0.7355_{\pm 0.01}$	$0.7078_{\pm 0.02}$	$0.9435_{\pm 0.00}$	$0.9196_{\pm 0.00}$	$0.7426_{\pm 0.00}$	**$0.9449_{\pm 0.00}$**
	ARI	$0.8318_{\pm 0.00}$	$0.7738_{\pm 0.00}$	$0.7819_{\pm 0.00}$	$0.7523_{\pm 0.00}$	$0.5210_{\pm 0.03}$	$0.4805_{\pm 0.03}$	$0.8376_{\pm 0.00}$	$0.7911_{\pm 0.00}$	$0.5296_{\pm 0.00}$	**$0.8587_{\pm 0.00}$**
	F-Score	$0.8410_{\pm 0.00}$	$0.7867_{\pm 0.00}$	$0.7942_{\pm 0.00}$	$0.7655_{\pm 0.00}$	$0.5482_{\pm 0.03}$	$0.5113_{\pm 0.03}$	$0.8465_{\pm 0.00}$	$0.8027_{\pm 0.00}$	$0.5568_{\pm 0.00}$	**$0.8666_{\pm 0.00}$**
Cal7	ACC	$0.7232_{\pm 0.00}$	$0.7367_{\pm 0.00}$	$0.6919_{\pm 0.00}$	$0.6791_{\pm 0.00}$	$0.4873_{\pm 0.00}$	$0.5487_{\pm 0.02}$	$0.8071_{\pm 0.00}$	$0.6933_{\pm 0.00}$	$0.5651_{\pm 0.00}$	**$0.8297_{\pm 0.00}$**
	NMI	$0.5569_{\pm 0.00}$	$0.5342_{\pm 0.00}$	**$0.6056_{\pm 0.00}$**	$0.5630_{\pm 0.00}$	$0.3508_{\pm 0.00}$	$0.4115_{\pm 0.05}$	$0.5233_{\pm 0.01}$	$0.5987_{\pm 0.00}$	$0.4699_{\pm 0.00}$	$0.4864_{\pm 0.00}$
	ARI	$0.4685_{\pm 0.00}$	$0.4778_{\pm 0.00}$	$0.5942_{\pm 0.00}$	$0.5480_{\pm 0.00}$	$0.2971_{\pm 0.00}$	$0.3532_{\pm 0.06}$	$0.5829_{\pm 0.00}$	$0.4874_{\pm 0.00}$	$0.4163_{\pm 0.00}$	**$0.6175_{\pm 0.00}$**
	F-Score	$0.6763_{\pm 0.00}$	$0.6884_{\pm 0.00}$	$0.7216_{\pm 0.00}$	$0.6831_{\pm 0.00}$	$0.4728_{\pm 0.00}$	$0.5363_{\pm 0.02}$	$0.7646_{\pm 0.00}$	$0.6703_{\pm 0.00}$	$0.5597_{\pm 0.00}$	**$0.7926_{\pm 0.00}$**
Mfeat	ACC	$0.8855_{\pm 0.00}$	$0.8550_{\pm 0.00}$	$0.8820_{\pm 0.00}$	$0.8545_{\pm 0.00}$	$0.6272_{\pm 0.00}$	$0.7474_{\pm 0.05}$	$0.8432_{\pm 0.01}$	$0.6980_{\pm 0.00}$	$0.7640_{\pm 0.00}$	**$0.8860_{\pm 0.00}$**
	NMI	$0.9039_{\pm 0.00}$	$0.8962_{\pm 0.00}$	$0.8940_{\pm 0.00}$	$0.8877_{\pm 0.00}$	$0.7037_{\pm 0.00}$	$0.7284_{\pm 0.02}$	$0.8823_{\pm 0.00}$	$0.7406_{\pm 0.00}$	$0.7259_{\pm 0.00}$	**$0.9041_{\pm 0.00}$**
	ARI	**$0.8641_{\pm 0.00}$**	$0.8373_{\pm 0.00}$	$0.8502_{\pm 0.00}$	$0.8328_{\pm 0.00}$	$0.5869_{\pm 0.00}$	$0.6662_{\pm 0.03}$	$0.8156_{\pm 0.00}$	$0.6328_{\pm 0.00}$	$0.6588_{\pm 0.00}$	$0.8611_{\pm 0.00}$
	F-Score	**$0.8788_{\pm 0.00}$**	$0.8547_{\pm 0.00}$	$0.8658_{\pm 0.00}$	$0.8507_{\pm 0.00}$	$0.6315_{\pm 0.00}$	$0.7037_{\pm 0.02}$	$0.8354_{\pm 0.00}$	$0.6738_{\pm 0.00}$	$0.6940_{\pm 0.00}$	$0.8753_{\pm 0.00}$
Scene	ACC	$0.4196_{\pm 0.00}$	$0.3251_{\pm 0.00}$	$0.3400_{\pm 0.00}$	$0.5871_{\pm 0.00}$	$0.6005_{\pm 0.00}$	$0.6077_{\pm 0.03}$	$0.2693_{\pm 0.06}$	$0.3846_{\pm 0.00}$	**$0.6945_{\pm 0.00}$**	$0.6525_{\pm 0.00}$
	NMI	$0.3436_{\pm 0.00}$	$0.2093_{\pm 0.00}$	$0.3141_{\pm 0.00}$	$0.4599_{\pm 0.00}$	$0.4712_{\pm 0.00}$	$0.4734_{\pm 0.03}$	$0.1463_{\pm 0.09}$	$0.2565_{\pm 0.00}$	$0.5300_{\pm 0.00}$	**$0.5375_{\pm 0.00}$**
	ARI	$0.1845_{\pm 0.00}$	$0.0727_{\pm 0.00}$	$0.1925_{\pm 0.00}$	$0.3757_{\pm 0.00}$	$0.3801_{\pm 0.00}$	$0.3887_{\pm 0.03}$	$0.0562_{\pm 0.07}$	$0.1013_{\pm 0.00}$	$0.4645_{\pm 0.00}$	**$0.4686_{\pm 0.00}$**
	F-Score	$0.3457_{\pm 0.00}$	$0.2680_{\pm 0.00}$	$0.3546_{\pm 0.00}$	$0.4676_{\pm 0.00}$	$0.4601_{\pm 0.00}$	$0.4815_{\pm 0.02}$	$0.2607_{\pm 0.05}$	$0.2865_{\pm 0.00}$	$0.5337_{\pm 0.00}$	**$0.5466_{\pm 0.00}$**

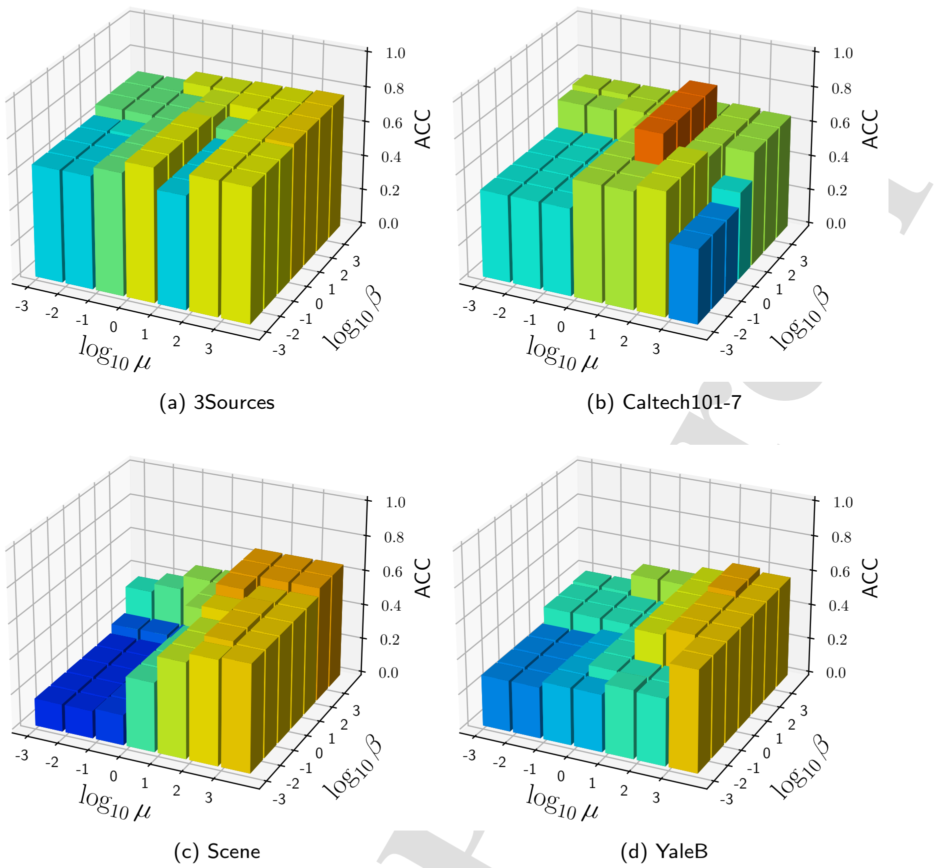
† It is worth noting that the MoMvGC results are based on the first-order graphs with five neighbors.

accurately representing the latent data distribution. The computational complexity is $O(n)$.

- CDMGC [10] pays attention to the consistency and diversity of graphs from multi-views when learning the optimal consensus graph. The computational complexity is $O(cn^2)$.
- MSCTM [11] benefits from exploring the inherent data manifold by learning the topological relationships among data points, which seamlessly integrates multiple graphs from different views into a consensus graph. The computational complexity is $O(cn^2)$.
- UDBGL [4] integrates the construction of bipartite graphs through subspace learning with consensus graph learning into a unified framework, ensuring clustering performance while reducing computational complexity. The computational complexity is $O(n)$.

4.1.3. Evaluation Metrics and Hyperparameter Settings

We adopt widely used evaluation metrics including accuracy (ACC), normalized mutual information (NMI), ARI and F-Score to comprehensively assess the clustering performance. The closer the value of these metrics is to 1, the better the clustering performance of the corresponding algorithm. In the MoMvGC model, we use the grid search technique, setting both μ and β to vary in the range of $\{10^{-3}, 10^{-2}, 10^{-1}, \dots, 10^{3}\}$. And we conducted the experiments on MATLAB 2022a environment.



(a) 3Sources (b) Caltech101-7

(c) Scene (d) YaleB

Figure 3: Parameter sensitivity of MoMvGC w.r.t μ and β.

4.2. Result Analyses

4.2.1. Clustering Performance

Table 1 presents the results of four commonly adopted clustering performance metrics. Based on the aforementioned results, we can draw the following conclusions:

a) The result reveals that the proposed model consistently outperforms or closely competes with existing approaches across all nine datasets, as evidenced by most metrics employed. This indicates the superior performance of high-order graphs within the realm of graph-based multi-view clustering. Moreover, the clustering metrics of our model exhibit zero variance, indicating the strong stability of the MoMvGC model alongside its superior clustering performance.

b) For example, on MSRC dataset, our method surpasses the best comparison method (MVGL) by **+0.0524** ACC, **+0.0804** NMI, **+0.0985** F-Score and **+0.1156** ARI, which can be considered as a substantial improvement for clustering tasks. Specifically, through experiments, it can be found that the optimal result is obtained by the consensus learning of only multi-order graphs from the fourth view, which proves the necessity of simultaneous selection of multi-order graphs and views.

4.2.2. Parameter Sensitivity on μ and β

We plotted the sensitivity of the four datasets with respect to the parameters μ and β. As shown in Figure 3, the sensitivity of μ is relatively higher than that of β. μ could achieve coarse tuning by imposing constraints on consensus graph $\boldsymbol{S}$ while β has a fine-tuning effect. Taking the dataset Caltech101-7 as an example, when $\mu = 10$ and $\beta \in [10^{-1}, 10^{2}]$, the ACC obtained by MoMvGC is the best and the performance is stable. This fully demonstrates the advantages of simultaneous selection and weight learning of multi-order graphs.

4.2.3. Impact of The Number of Neighbors

The KNN algorithm [23] is used to generate first-order graph for each view as input of the MoMvGC model. The number of neighbors is a crucial parameter that determines the sparsity level of the graph. To further investigate the impact of high-order graphs on graph sparsity, we plotted the trend of ACC obtained by the MoMvGC model with varying numbers of neighbors, as shown in Figure 5. It can be observed that for most datasets: *a*) an insufficient or excessive number of neighbors will have a negative impact on the clustering accuracy. *b*) when the number of neighbors is greater than three but less than nine, the performance is relatively good and stable. On the one hand, a smaller

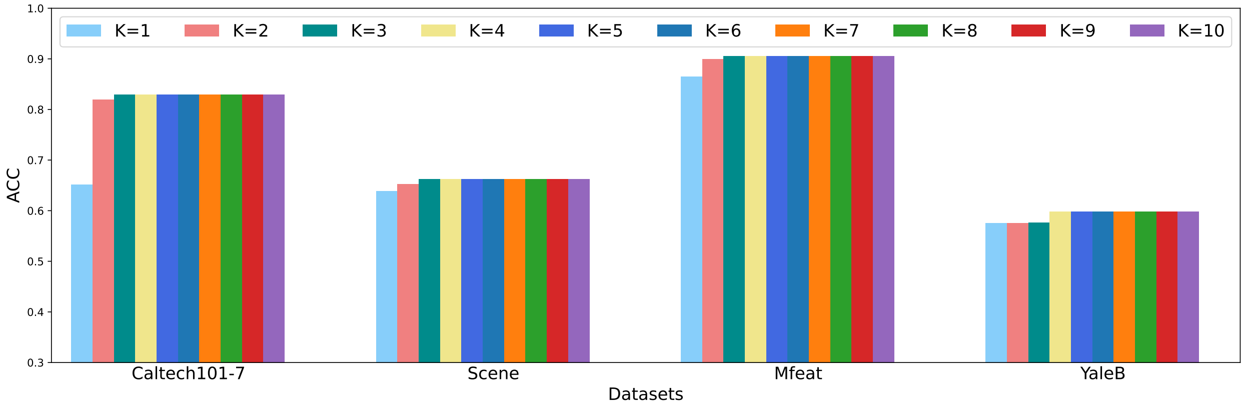
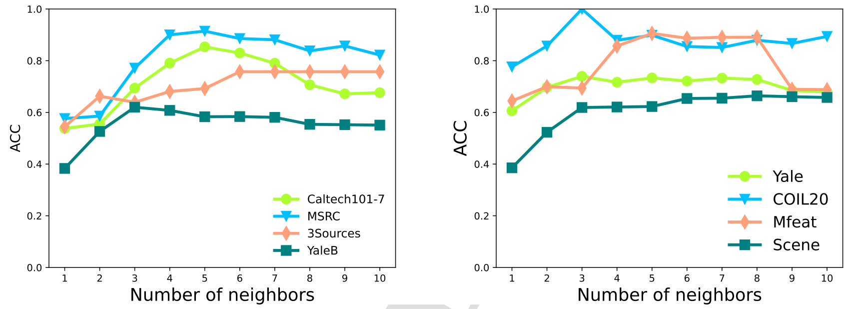
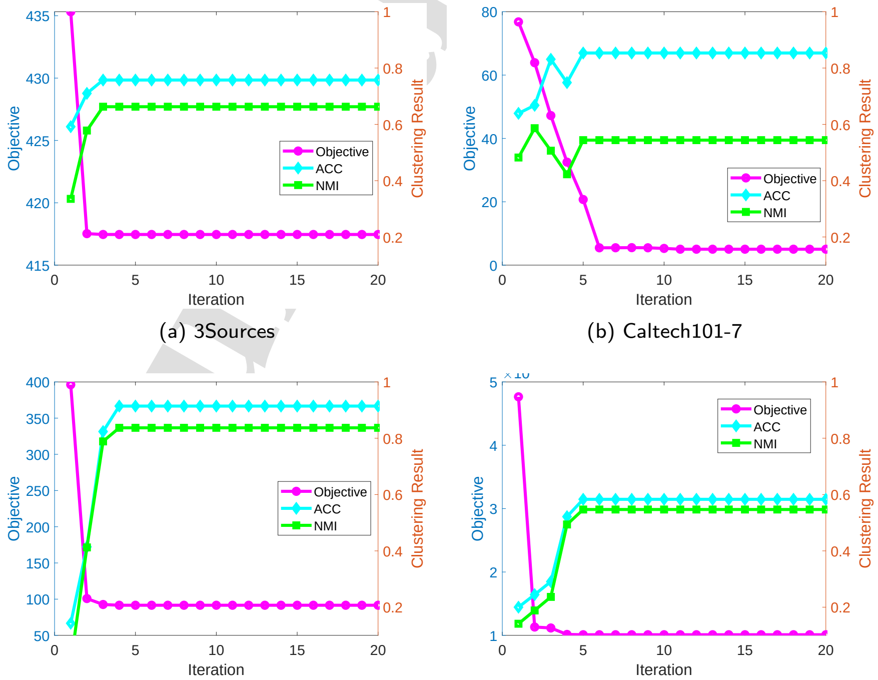


Figure 4: Sensitivity analysis w.r.t. the number of the order K.



(a) (b)

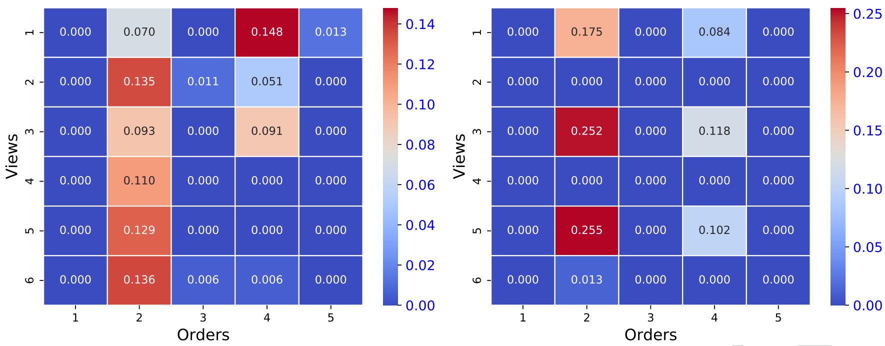
Figure 5: Impact w.r.t the number of neighbors.



(a) 3Sources (b) Caltech101-7

(c) MSRC (d) YaleB

Figure 6: Convergence analysis.



(a) Caltech101-7 (b) Mfeat

Figure 7: Visualization of learned weights.

number of neighbors implies that the first-order graph lacks essential must-link information, hindering the ability of the model to explore structural information. On the other hand, a larger number of neighbors may introduce some cannot-link information to the first-order graph, which could mislead the learning of the consensus graph.

4.2.4. Impact of The Number of The Order K

In the MoMvGC model, the order K is a significant parameter that should not be overlooked. To investigate the impact of the order K on the performance of the MoMvGC model, we present a histogram of the optimal ACC w.r.t. the order K. As shown in Figure 4, ACC obtained is low when K is small. This can be attributed to the fact that graphs from different views lack necessary must-link information, making that the learned consensus graph is of poor quality. With the increase of K, the model can mine more rich structural information that is not available in the low-order graphs but in the high-order graphs, which promotes the clustering performance. When K is relatively large, a significant amount of noise is introduced into the high-order graphs, and consequently, the adaptive selection mechanism discards these high-order graphs. Therefore, we suggest that, in general, setting K to 5 is sufficient to fully exploit useful high-order information.

4.2.5. Visualization of Learned Weights

The core idea of MoMvGC lies in the simultaneous selection and learning of weights for multi-order graphs using the constrained $l_{1,2}$-norm. To illustrate this, we visualize the weight distributions learned for the Caltech101-7 and Mfeat datasets under optimal performance, as shown in Figure 7. For the Caltech101-7 dataset, the learned non-zero weights corresponding to the optimal clustering accuracy concentrate in the first, third, and fifth views, which means that other views with noise are not very suitable for consensus graph learning. All views of the Mfeat dataset contribute to the optimal clustering performance. Moreover, it is worth noting that only a few of the multi-order graphs from the same view are selected, which highlights the role of row sparsity of the constrained $l_{1,2}$-norm.

4.2.6. Convergence Analysis

For the sake of demonstrating the effectiveness of the MoMvGC model, we plot the trend of the objective and the corresponding metrics for the four datasets with the increase of the number of iterations. As depicted in Figure 6, as the number of iterations increases, the objective value steadily decreases while the ACC and NMI gradually improve. This observation suggests that the model continuously learns better consensus graphs through each optimization step. When the objective function converges, ACC and NMI also reach their maximum.

5. Conclusion and Discussion

In this paper, we introduce high-order graphs into multi-view clustering, and propose Multi-view and Multi-order Graph Clustering via Constrained $l_{1,2}$-norm (MoMvGC), which alleviates the influence of sparse graphs on clustering performance. Specifically, by introducing constrained $l_{1,2}$-norm, the model can simultaneously select and learn the weights of multi-order graphs, avoiding the information loss or information redundancy caused by the two-step strategy. Moreover, the MoMvGC model is capable of jointly selecting multi-order graphs and views, reducing the impact of noisy views on consensus graph learning. Extensive experiments conducted on nine datasets provide ample evidence of the feasibility and effectiveness of our model.

On the other hand, utilizing the $n \times n$ fully connected graph as input introduces a relatively high computational complexity due to eigenvalue decomposition, which is $\mathcal{O}(n^2)$. This is not conducive to clustering task of large-scale multi-view datasets. Therefore, the design of a structure graph learning method based on multi-view and multi-order anchor graphs is our future work.

CRediT Authorship Contribution Statement

Haonan Xin: Conceptualization, Methodology, Validation, Software, Formal analysis, Investigation, Writing – original draft. Zhezheng Hao: Methodology, Validation, Formal analysis, Writing – original draft. Zhensheng Sun:

Methodology, Validation, Formal analysis, Investigation, Resources, Writing – review & editing. Rong Wang: Methodology, Validation, Formal analysis, Investigation, Resources, Project administration, Supervision, Funding acquisition. Zongcheng Miao: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – review & editing, Project administration, Supervision, Funding acquisition. Feiping Nie: Conceptualization, Validation, Project administration, Funding acquisition, Supervision.

References

[1] Ben-Hur, A., Horn, D., Siegelmann, H.T., Vapnik, V., 2001. Support vector clustering. J. Mach. Learn. Res. 2, 125–137.

[2] El Hajjar, S., Dornaika, F., Abdallah, F., 2022. Multi-view spectral clustering via constrained nonnegative embedding. Inf. Fusion 78, 209–217.

[3] Fan, K., 1949. On a theorem of weyl concerning eigenvalues of linear transformations i, in: Proc. PNAS, pp. 652–655.

[4] Fang, S.G., Huang, D., Cai, X.S., Wang, C.D., He, C., Tang, Y., 2023a. Efficient multi-view clustering via unified and discrete bipartite graph learning. IEEE Trans. Neural Netw. Learn. Syst. 34, 4412–4419.

[5] Fang, U., Li, M., Li, J., Gao, L., Jia, T., Zhang, Y., 2023b. A comprehensive survey on multi-view clustering. IEEE Trans. Knowl. Data Eng. 35, 12350–12368.

[6] He, Z., Wan, S., Zappatore, M., Lu, H., 2023. A similarity matrix low-rank approximation and inconsistency separation fusion approach for multi-view clustering. IEEE Trans. Artif. Intell. 5, 868–880.

[7] Hu, Z., Nie, F., Wang, R., Li, X., 2020. Multi-view spectral clustering via integrating nonnegative embedding and spectral embedding. Inf. Fusion 55, 251–259.

[8] Huang, D., Wang, C.D., Lai, J.H., 2023. Fast multi-view clustering via ensembles: Towards scalability, superiority, and simplicity. IEEE Trans. Knowl. Data Eng. 35, 11388–11402.

[9] Huang, S., Tsang, I., Xu, Z., Lv, J., Liu, Q., 2022a. Multi-view clustering on topological manifold, in: Proc. AAAI, pp. 6944–6951.

[10] Huang, S., Tsang, I.W., Xu, Z., Lv, J., 2021. Measuring diversity in graph learning: A unified framework for structured multi-view clustering. IEEE Trans. Knowl. Data Eng. 34, 5869–5883.

[11] Huang, S., Wu, H., Ren, Y., Tsang, I., Xu, Z., Feng, W., Lv, J., 2022b. Multi-view subspace clustering on topological manifold, in: Proc. NeurIPS, pp. 25883–25894.

[12] Jouirou, A., Baâzaoui, A., Barhoumi, W., 2019. Multi-view information fusion in mammograms: A comprehensive overview. Inf. Fusion 52, 308–321.

[13] Li, A., Feng, C., Cheng, Y., Zhang, Y., Yang, H., 2024. Incomplete multiview subspace clustering based on multiple kernel low-redundant representation learning. Inf. Fusion 103, 102086.

[14] Li, L., Zhang, J., Wang, S., Liu, X., Li, K., Li, K., 2023. Multi-view bipartite graph clustering with coupled noisy feature filter. IEEE Trans. Knowl. Data Eng. 35, 12842–12854.

[15] Li, Y., Yang, M., Zhang, Z., 2018. A survey of multi-view representation learning. IEEE Trans. Knowl. Data Eng. 31, 1863–1883.

[16] Liu, J., Liu, X., Yang, Y., Guo, X., Kloft, M., He, L., 2021. Multiview subspace clustering via co-training robust data representation. IEEE Trans. Neural Netw. Learn. Syst. 33, 5177–5189.

[17] Lu, X., Feng, S., 2023. Structure diversity-induced anchor graph fusion for multi-view clustering, in: Proc. TKDD, pp. 1–18.

[18] Ming, D., Ding, C., 2019. Robust flexible feature selection via exclusive l21 regularization, in: Proc. IJCAI, pp. 3158–3164.

[19] Mohar, B., Alavi, Y., Chartrand, G., Oellermann, O., 1991. The laplacian spectrum of graphs. Graph comb. 2, 12.

[20] Nie, F., Cai, G., Li, X., 2017a. Multi-view clustering and semi-supervised classification with adaptive neighbours, in: Pro. AAAI.

[21] Nie, F., Li, J., Li, X., et al., 2016. Parameter-free auto-weighted multiple graph learning: a framework for multiview clustering and semi-supervised classification., in: Proc. IJCAI, pp. 1881–1887.

[22] Nie, F., Li, J., Li, X., et al., 2017b. Self-weighted multiview clustering with multiple graphs., in: Proc. IJCAI, pp. 2564–2570.

[23] Nie, F., Wang, X., Huang, H., 2014. Clustering and projected clustering with adaptive neighbors, in: Proc. SIGKDD, pp. 977–986.

[24] Nie, F., Wu, D., Wang, R., Li, X., 2020. Truncated robust principle component analysis with a general optimization framework. IEEE Trans. Pattern Anal. Mach. Intell. 44, 1081–1097.

[25] Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22, 888–905.

[26] Sun, M., Zhang, P., Wang, S., Zhou, S., Tu, W., Liu, X., Zhu, E., Wang, C., 2021. Scalable multi-view subspace clustering with unified anchors, in: Proc. MM, pp. 3528–3536.

[27] Sun, S., Dong, W., Liu, Q., 2020. Multi-view representation learning with deep gaussian processes. IEEE Trans. Pattern Anal. Mach. Intell. 43, 4453–4468.

[28] Tang, C., Liu, X., Zhu, X., Zhu, E., Luo, Z., Wang, L., Gao, W., 2020a. Cgd: Multi-view clustering via cross-view graph diffusion, in: Proc. AAAI, pp. 5924–5931.

[29] Tang, C., Liu, X., Zhu, X., Zhu, E., Luo, Z., Wang, L., Gao, W., 2020b. Cgd: Multi-view clustering via cross-view graph diffusion, in: Proc. AAAI, pp. 5924–5931.

[30] Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q., 2015. Line: Large-scale information network embedding, in: Proc. WWW, pp. 1067–1077.

[31] Wang, D., Li, T., Huang, W., Luo, Z., Deng, P., Zhang, P., Ma, M., 2023a. A multi-view clustering algorithm based on deep semi-nmf. Inf. Fusion 99, 101884.

[32] Wang, H., Yang, Y., Liu, B., 2019. Gmc: Graph-based multi-view clustering. IEEE Trans. Knowl. Data Eng. 32, 1116–1129.

[33] Wang, P., Wu, D., Wang, R., Nie, F., 2023b. Multi-view graph clustering via efficient global-local spectral embedding fusion, in: Proc. MM, pp. 3268–3276.

[34] Wang, Q., Dou, Y., Liu, X., Xia, F., Lv, Q., Yang, K., 2018. Local kernel alignment based multi-view clustering using extreme learning machine. Neural Comput. 275, 1099–1111.

[35] Wang, R., Chen, H., Lu, Y., Zhang, Q., Nie, F., Li, X., 2023c. Discrete and balanced spectral clustering with scalability. IEEE Trans. Pattern Anal. Mach. Intell. 45, 14321–14335.

[36] Wang, S., Liu, X., Zhu, X., Zhang, P., Zhang, Y., Gao, F., Zhu, E., 2021. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. IEEE Trans. Image Process. 31, 556–568.

[37] Wu, D., Chang, W., Lu, J., Nie, F., Wang, R., Li, X., 2022a. Adaptive-order proximity learning for graph-based clustering. Pattern Recogn. 126, 108550.

[38] Wu, D., Lu, J., Nie, F., Wang, R., Yuan, Y., 2022b. Emgc2f: Efficient multi-view graph clustering with comprehensive fusion, in: Proc. IJCAI, pp. 3566–3572.

[39] Wu, D., Nie, F., Dong, X., Wang, R., Li, X., 2021. Parameter-free consensus embedding learning for multiview graph-based clustering. IEEE Trans. Neural Netw. Learn. Syst. 33, 7944–7950.

[40] Xu, J., Ren, Y., Shi, X., Shen, H.T., Zhu, X., 2023. Untie: Clustering analysis with disentanglement in multi-view information fusion. Inf. Fusion 100, 101937.

[41] Yang, Y., Lan, C., Li, X., Luo, B., Huan, J., 2014. Automatic social circle detection using multi-view clustering, in: Proc. CIKM, pp. 1019–1028.

[42] Yu, S., Wang, S., Dong, Z., Tu, W., Liu, S., Lv, Z., Li, P., Wang, M., Zhu, E., 2024. A non-parametric graph clustering framework for multi-view data, in: Proc. AAAI, pp. 16558–16567.

[43] Zhan, K., Zhang, C., Guan, J., Wang, J., 2017. Graph learning for multiview clustering. IEEE Trans. Cybern. 48, 2887–2895.

[44] Zhao, H., Ding, Z., Fu, Y., 2017a. Multi-view clustering via deep matrix factorization, in: Proc. AAAI, pp. 2921–2927.

[45] Zhao, J., Xie, X., Xu, X., Sun, S., 2017b. Multi-view learning overview: Recent progress and new challenges. Inf. Fusion 38, 43–54.

[46] Zhao, M., Yang, W., Nie, F., 2023. Auto-weighted orthogonal and nonnegative graph reconstruction for multi-view clustering. Inf. Sci. 632, 324–339.

Highlights

- We introduce the idea of high-order graphs to multi-view clustering.
- We innovatively propose constrained $l_{1,2}$-norm for weight learning.
- Our model enables simultaneous selection of views and multi-order graphs.
- We design a comprehensive optimization framework for MoMvGC model.
- Our model achieves decent clustering performance.

Credit Author Statement

Haonan Xin: Conceptualization, Methodology, Validation, Software, Formal analysis, Investigation, Writing – original draft. Zhezheng Hao: Methodology, Validation, Formal analysis, Writing – original draft. Zhensheng Sun: Methodology, Validation, Formal analysis, Investigation, Resources, Writing – review \& editing. Rong Wang: Methodology, Validation, Formal analysis, Investigation, Resources, Project administration, Supervision, Funding acquisition. Zongcheng Miao: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – review \& editing, Project administration, Supervision, Funding acquisition. Feiping Nie: Conceptualization, Validation, Project administration, Funding acquisition, Supervision.

Declaration of interest

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.