Sample-level Multi-view Graph Clustering

Yuze Tan, Yixi Liu, Shudong Huang,* Wentao Feng, Jiancheng Lv
Sichuan University
{tanyuze,liuyixi}@stu.scu.edu.cn,{huangsd,Wtfeng2021,lvjiancheng}@scu.edu.cn

Abstract

Multi-view clustering has hitherto been studied due to their effectiveness in dealing with heterogeneous data. Despite the empirical success made by recent works, there still exists several severe challenges. Particularly, previous multi-view clustering algorithms seldom consider the topological structure in data, which is essential for clustering data on manifold. Moreover, existing methods cannot fully explore the consistency of local structures between different views as they uncover the clustering structure in a intra-view way instead of a inter-view manner. In this paper, we propose to exploit the implied data manifold by learning the topological structure of data. Besides, considering that the consistency of multiple views is manifested in the generally similar local structure while the inconsistent structures are the minority, we further explore the intersections of multiple views in the sample level such that the cross-view consistency can be better maintained. We model the above concerns in a unified framework and design an efficient algorithm to solve the corresponding optimization problem. Experimental results on various multi-view datasets certificate the effectiveness of the proposed method and verify its superiority over other SOTA approaches.

1. Introduction

In many real scenarios, data are usually collected from diverse sources in various domains or described by multiple feature sets [1, 2]. A case in point is the image dataset, which can be represented by different visual descriptors, such as LBP, GIST, CENTRIST, HOG, SIFT and Color Moment [3]. Moreover, a document can be written in different languages, and a website can be described by its linkage, page content, etc. These data are generally known as multi-view data [1, 4]. Each view contains partial and complementary information, any of which suffices for learning, and they all together agree to a consensus latent representation [5–7]. Multi-view clustering aims to accurately par-

tition data into distinct clusters according to their compatible and complementary information embedded in heterogeneous features [1, 8].

In general, multi-view clustering methods can be divided into four categories according to the mechanisms and principles on which these methods are based. [1] initiated the trend of *co-training* algorithm by first carrying out a co-training strategy and making the clustering results on all view tend to each other as well as satisfying the broadest agreement across all views. Owing to the efficiency of exploiting similarities among multiple views, *multi-kernel learning* is widely utilized to boost the performance of multi-view clustering methods [9–11]. *Multi-task* multi-view clustering [12,13] inherits the property of multi-task clustering. Each view of the data is treated with at least one related task. Moreover, inter-task knowledge is transferred from one to another so that the relationship between multi-view and multi-task is fully exploited to improve the clustering outcomes. Another kind of *graph-based* multi-view clustering methods usually attempt to explore an optimal consensus graph across all views, and utilize graph-cut algorithm on the optimal graph to obtain final clustering result [3, 5, 14].

A great number of multi-view clustering methods have been proposed and illustrated remarkable empirical successes up to now [15–17]. However, there are still severe drawbacks urgently need to be overcome. For one thing, existing multi-view clustering algorithms seldom consider the topological structure in data, which is essential for clustering data on manifold. Considering that the data sampled from real world typically lie in the nonlinear manifold, data points geometrically far from one to another may keep high consistency when they are linked by a series of consecutive neighbors. For another, they suffer from a coarse-grained problem that ignores the view correlation between samples. Note that redundancy or partial structure mistakes in certain views may lead to a sub-optimal cluster structure. Besides, traditional view-wise fusion strategy would result in a superimposition of redundancies, and hence acquire a more imprecise common cluster structure.

Regarding the above issues, we propose to exploit the

---
*Corresponding author.

implied data manifold by learning the topological structure of data. Besides, considering that the consistency of multiple views is manifested in the generally similar local structure while the inconsistent structures are the minority, we further explore the intersections of multiple views in the sample level such that the cross-view consistency can be better maintained. By leveraging the subtasks of topological relevance learning and the sample-level graph fusion in our collaborative model, each subtask is alternately boosted towards an optimal solution. Experimental results on various multi-view datasets certificate the effectiveness of the proposed method.

2. Related Work

Manifold learning has attracted widespread attention in recent decades [18, 19]. A plethora of works are mainly spent on exploring methods for discovering nonlinear low-dimensional manifolds between individuals where the topological relevance can be high [20]. The ability of those algorithms is to learn the manifold structure embedded in the high-dimensional ambient space.

Existing studies show that interactions between the instances place heavy reliance on the topological relationship instead of the metric range [21]. That is, a higher cohesion of the aggregation can be maintained by a topological interaction. Therefore, the work in [22] learns the topological relevance between different samples to uncover the hidden structures of the crowded and measure collectiveness. From the perspective of topological space, the relevance between two individuals is high if they are connected by countable neighbors, even though they show low similarity in the spatial scene. It is based on a simple yet intuitive assumption that the topological connectivities between individuals could be propagated from near to far. In light of this, we replace the most often-used Euclidean distance with a powerful manifold topological relevance to capture the intrinsic structures.

Given a precomputed affinity matrix $\mathbf{G} = [g_{jk}] \in \mathbb{R}^{n \times n}$, where n is the number of data samples. Based on the assumption that if two data points keep high consistency, their topological relevance to any other point is assumed to be similar, [22] introduced a structure-based clustering strategy to explore the topological relevance between different samples. The objective function can be defined as

$$\min_{\mathbf{Z}} \frac{1}{2} \sum_{i,j,k=1}^{n} g_{jk} \left(z_{ij} - z_{ik} \right)^2 + \lambda \left\| \mathbf{Z} - \mathbf{I} \right\|_F^2, \quad (1)$$

where $\mathbf{Z}$ represents the topological relationship graph and the element z_{ij} is the topological relevance of i-th point to j-th point. The first term in Eq. (1), as a smoothness constraint, is incorporated to meet the aforementioned assumption, which ensures that for data points j and k, their
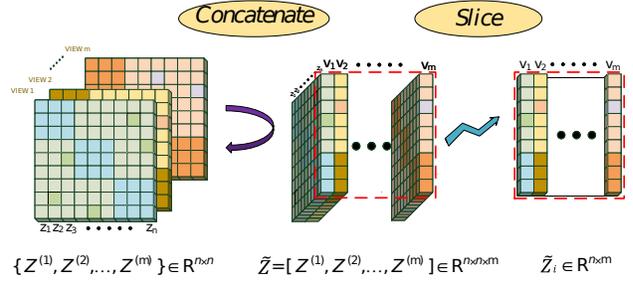

Figure 1. By concatenating all views to a tensor, we obtain $\widetilde{\mathbf{Z}}$. Rotating the obtained tensor $\widetilde{\mathbf{Z}}$ and cutting it vertically, we have the i-th frontal slice $\widetilde{\mathbf{Z}}_i$.

topological relevance to point i is assumed to be similar if they have high consistency. The second term in Eq. (1), as a fitting constraint, is used to rule out the trivial solution. We also utilize a weighting coefficient α to balance the two terms. According to the above form, we detect the consistency by propagating topological relationship, and the relevance between two distant points will be high if they are connected by countable neighbors.

However, the model in Eq. (1) is applied for solving the singe-view problem, and it is not suitable for multi-view clustering tasks. In this paper, we propose to explore the manifold topological structure for multi-view data. Given $\mathbf{G}^{(v)}(v = 1, \ldots, m) \in \mathbb{R}^{n \times n}$ as the input similarity graphs of multi-view data with m views. Based on Eq. (1), the topological relationship for each view can be measured by

$$\min_{\mathbf{Z}^{(v)}} \frac{1}{2} \sum_{v=1}^{m} \sum_{i,j,k=1}^{n} g_{jk}^{(v)} \left(z_{ij}^{(v)} - z_{ik}^{(v)} \right)^2 + \lambda \left\| \mathbf{Z}^{(v)} - \mathbf{I} \right\|_F^2$$

$$\text{s.t.} \quad \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0, \quad (2)$$

where all $z_{ij}^{(v)}$ is constrained to be non-negative and the sum of $\mathbf{z}_i^{(v)}$ is one. Considering that a data point connected with many similar neighbors would largely affect the objective value. We utilize a normalized form of Eq. (2), which is formulated as

$$\min_{\mathbf{Z}^{(v)}} \frac{1}{2} \sum_{v=1}^{m} \sum_{i,j,k=1}^{n} g_{jk}^{(v)} \left(\frac{z_{ij}^{(v)}}{\sqrt{d_{jj}^{(v)}}} - \frac{z_{ik}^{(v)}}{\sqrt{d_{kk}^{(v)}}} \right)^2 + \lambda \left\| \mathbf{Z}^{(v)} - \mathbf{I} \right\|_F^2$$

$$\text{s.t.} \quad \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0, \quad (3)$$

where $\mathbf{D}^{(v)}$ is the degree matrix of $\mathbf{G}^{(v)}$.

3. The proposed Methodology

As aforementioned, existing multi-view clustering methods generally suffer from a coarse-grained problem that ig-

nores the view correlation between samples. To tackle this inappropriate fusion strategy, we design a novel model to explore the intersections of multiple views in the sample level. It is able to capture more intrinsic cross-view consistency and is robust to the inconsistency introduced by certain views. Inspired by [23, 24], here we proposed a transformation on a set of multi-view matrices. As shown in Figure 1, first we start with concatenating all views to a tensor $\tilde{\mathbf{Z}} \in \mathbb{R}^{n \times n \times m}$. Then we get the i-th frontal slice $\tilde{\mathbf{Z}}_i \in \mathbb{R}^{n \times m}$ corresponding to the i-th sample to explore sample-level manifold information. By conducting such transformation, it is feasible for us to manipulate fuse all views on the sample level as follow:

$$
\min_{\mathbf{S}} \sum_{i=1}^{n} \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Z}}_i \right\|_2^2 + \beta \|\mathbf{S}\|_F^2
$$
$$
\text{s.t.} \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0, \tag{4}
$$
$$
\mathbf{S} \geq 0, \mathbf{W} \geq 0, \mathbf{s}_i^T \mathbf{1} = 1, \mathbf{W} \mathbf{1} = \mathbf{1},
$$

where $\mathbf{w}_i^T$ corresponds to the weight of different views with respect to the i-th sample. Each column of $\mathbf{S}$ integrates the consistent information of one instance across all views. β is the balanced parameter for the penalty term, which is utilized to prevent the trivial solution when a specific view dominants the others. The constraints that we apply to the optimal fusion graph $\mathbf{S}$ and weight matrix $\mathbf{W}$ assist to narrow down the solution space and strongly boost the efficiency of our optimization process. Based on the elementary yet intuitive assumption that the topological relevance between instances tends to be propagated from near to far. That is, the individuals have low spatial similarity may achieve high topological relevance if they are linked by the consecutive neighbors. Thus, we explore the similarity between instances on a more appropriate topological manifold structure rather than on the Euclidean structure. With the help of topological manifold exploration in the multi-view clustering domain, by combining Eq. (3) and Eq. (4), we can easily derive:

$$
\min_{\mathbf{z}^{(v)}, \mathbf{W}, \mathbf{S}} \sum_{v=1}^{m} \underbrace{\frac{1}{2} \sum_{i,j,k=1}^{n} g_{jk}^{(v)} \left(\frac{z_{ij}^{(v)}}{\sqrt{d_{jj}^{(v)}}} - \frac{z_{ik}^{(v)}}{\sqrt{d_{kk}^{(v)}}} \right)^2 + \lambda \left\| \mathbf{Z}^{(v)} - \mathbf{I} \right\|_F^2}_{\text{topological relevance learning}}
$$
$$
\underbrace{+ \alpha \sum_{i=1}^{n} \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Z}}_i \right\|_2^2 + \beta \|\mathbf{S}\|_F^2}_{\text{sample-level graph fusion}}
$$
$$
\text{s.t.} \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0,
$$
$$
\mathbf{S} \geq 0, \mathbf{w}_i \geq 0, \mathbf{S} \mathbf{1} = \mathbf{1}, \mathbf{w}_i^T \mathbf{1} = \mathbf{1}. \tag{5}
$$

Moreover, considering the graph for sample-level graph fusion might be a translation from the topological manifold,

we introduce an intermediate variable $\mathbf{Q}$ to link the topological relevance learning and the sample-level graph fusion. As a result, the joint optimization framework of sample-level multi-view clustering with topological relevance considered can be formulated as:

$$
\min_{\mathbf{z}^{(v)}, \mathbf{W}, \mathbf{S}} \sum_{v=1}^{m} \underbrace{\frac{1}{2} \sum_{i,j,k=1}^{n} g_{jk}^{(v)} \left(\frac{z_{ij}^{(v)}}{\sqrt{d_{jj}^{(v)}}} - \frac{z_{ik}^{(v)}}{\sqrt{d_{kk}^{(v)}}} \right)^2 + \lambda \left\| \mathbf{Z}^{(v)} - \mathbf{I} \right\|_F^2}_{\text{topological relevance learning}}
$$
$$
\underbrace{+ \gamma \sum_{v=1}^{m} \|\mathbf{Z}^{(v)} - \mathbf{Q}^{(v)}\|_F^2 + \alpha \sum_{i=1}^{n} \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \right\|_2^2 + \beta \|\mathbf{S}\|_F^2}_{\text{sample-level graph fusion}}
$$
$$
\text{s.t.} \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0,
$$
$$
\mathbf{S} \geq 0, \mathbf{w}_i \geq 0, \mathbf{S} \mathbf{1} = \mathbf{1}, \mathbf{w}_i^T \mathbf{1} = \mathbf{1}, \tag{6}
$$

where γ and α are trade-off parameters. It is noteworthy that our model in Eq. (6) enjoys several properties that are distinct from existing approaches:

- Instead of performing the inter-view graph fusion, we propose to conducting a cross-view graph fusion process in a sample-wise way. Considering that the consistency of multiple views is manifested in the generally similar local structure while the inconsistent structures are the minority, we further explore the intersections of multiple views in the sample level such that the cross-view consistency can be better maintained.

- Note that the topological structure is essential for clustering data on manifold. Based on a simple yet intuitive assumption that the topological connectivities between individuals could be propagated from near to far, we propose to exploit the implied data manifold with the topological relevance considered.

- We model the subtasks of topological relevance learning and the sample-level graph fusion in a unified framework, each subtask is alternately boosted towards an optimal solution. An efficient algorithm to solve the corresponding optimization problem is also introduced. Experimental results on several benchmark datasets certificate the effectiveness of our method.

4. Optimization

In this section, we design an iterative updating algorithm to solve the optimization problem in Eq. (6). Since it is not jointly convex in all variable s, we propose to optimize the objective function with respect to one variable while fixing other variables. And the procedure repeats until convergence.

Algorithm 1: Algorithm to solve Eq. (10)

Input: a nonzero matrix $\mathbf{A}$ and a nonzero vector $\mathbf{b}$.
Set $1 < \rho < 2$, initialize $\eta > 0$, $\mathbf{h}$.
Output: $\mathbf{Z}^{(v)}$.
1: **repeat**
2:   Update $\mathbf{p}$ according to (13).
3:   Update $\mathbf{z}_i^{(v)}$ according to (15).
4:   Update $\eta \leftarrow \rho\eta$.
5:   Update $\mathbf{h} \leftarrow \mathbf{h} + \eta \left(\mathbf{z}_i^{(v)} - \mathbf{p} \right)$.
6: **until** converge

Algorithm 2: Algorithm to solve Eq. (6)

Input: A set of similarity graph $\mathbf{G}^{(v)}$. ($v = 1,2,...,m$)
Parameter α, β, λ and γ.
Output: Clustering result.
1: **repeat**
2:   Update $\mathbf{Z}^{(v)}$ according to Algorithm 1.
3:   Update $\mathbf{S}$ by solving Eq. (18).
4:   Update $\mathbf{W}$ according to Eq. (21).
5:   Update $\widetilde{\mathbf{Q}}_i$ according to Eq. (24).
6: **until** converge
7: Conduct the standard spectral clustering on the optimal graph $\mathbf{S}$ to obtain the final clustering result.

Update $\mathbf{Z}^{(v)}$

For each $\mathbf{Z}^{(v)}$, we need to solve

$$
\min_{\mathbf{Z}^{(v)}} \frac{1}{2} \sum_{v=1}^{m} \sum_{i,j,k=1}^{n} g_{jk}^{(v)} \left(\frac{z_{ij}^{(v)}}{\sqrt{d_{jj}^{(v)}}} - \frac{z_{ik}^{(v)}}{\sqrt{d_{kk}^{(v)}}} \right)^2
$$
$$
+ \lambda \sum_{v=1}^{m} \left\| \mathbf{Z}^{(v)} - \mathbf{I} \right\|_F^2 + \gamma \sum_{v=1}^{m} \| \mathbf{Q}^{(v)} - \mathbf{Z}^{(v)} \|_F^2 \quad (7)
$$
$$
\text{s.t. } \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0.
$$

Since Eq. (7) is independent for different v, for each v we have

$$
\min_{\mathbf{Z}^{(v)}} \frac{1}{2} \sum_{i=1}^{n} \left\{ \sum_{j,k=1}^{n} g_{jk}^{(v)} \left(\frac{z_{ij}^{(v)}}{\sqrt{d_{jj}^{(v)}}} - \frac{z_{ik}^{(v)}}{\sqrt{d_{kk}^{(v)}}} \right)^2 \right.
$$
$$
\left. + \lambda \sum_{j=1}^{n} \left\| z_{ij}^{(v)} - e_{ij} \right\|_F^2 + \gamma \sum_{j=1}^{n} \| q_{ij}^{(v)} - z_{ij}^{(v)} \|_F^2 \right\} \quad (8)
$$
$$
\text{s.t. } \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0.
$$

For each i, Eq. (8) can be further rewritten in a vector form as

$$
\min_{\mathbf{z}_i^{(v)}} \sum_{i=1}^{n} \left(\mathbf{z}_i^{(v)} \right)^T \left(\mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{G}^{(v)} \mathbf{D}^{-\frac{1}{2}} \right) \mathbf{z}_i^{(v)}
$$
$$
+ \lambda \left\| \mathbf{z}_i^{(v)} - \mathbf{e}_i \right\|_F^2 + \gamma \| \mathbf{q}_i^{(v)} - \mathbf{z}_i^{(v)} \|_F^2 \quad (9)
$$
$$
\text{s.t. } \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1} = 1, z_{ij}^{(v)} \geq 0.
$$

Denote $\mathbf{A}^{(v)} = (1 + \lambda + \gamma) \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{G}^{(v)} \mathbf{D}^{-\frac{1}{2}}$ and $\mathbf{b} = 2\lambda\mathbf{e}_i + 2\gamma\mathbf{q}_i^{(v)}$, Eq. (9) can be stated as:

$$
\min_{\left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1}=1, z_{ij}^{(v)} \geq 0} \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{A}^{(v)} \mathbf{z}_i^{(v)} - \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{b} \quad (10)
$$

It is obvious that Eq. (10) is a quadratic convex optimization problem which can be solved with the classical augmented Lagrangian multiplier (ALM) method [25]. In detail, Eq. (10) can be cracked by tackling its counterpart:

$$
\min_{\left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1}=1, z_{ij}^{(v)} \geq 0, \mathbf{p}=\mathbf{z}_i^{(v)}} \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{A}^{(v)} \mathbf{p} - \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{b} \quad (11)
$$

whose augmented Lagrangian function can be defined as:

$$
\min_{\left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1}=1, z_{ij}^{(v)} \geq 0} \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{A}^{(v)} \mathbf{p} - \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{b}
$$
$$
+ \frac{\eta}{2} \| \mathbf{z}_i^{(v)} - \mathbf{p} + \frac{1}{\eta} \mathbf{h} \|_2^2 \quad (12)
$$

where the second term in Eq. (12) is a penalty function term which guarantees that $\mathbf{p} = \mathbf{z}_i^{(v)}$, η and $\mathbf{h}$ are the corresponding penalty coefficient and parameter, respectively.
Note that $\mathbf{p}$ and $\mathbf{z}_i^{(v)}$ can be iteratively optimized:
1) Update $\mathbf{p}$ with fixed $\mathbf{z}_i^{(v)}$ The Lagrange function of Eq. (12) w.r.t. $\mathbf{p}$ is

$$
\mathcal{L}_p = \left(\mathbf{z}_i^{(v)} \right)^T \mathbf{A}^{(v)} \mathbf{p} + \frac{\eta}{2} \| \mathbf{z}_i^{(v)} - \mathbf{p} + \frac{1}{\eta} \mathbf{h} \|_2^2 \quad (13)
$$

Taking the derivative of $\mathcal{L}_p$ w.r.t $\mathbf{p}$ and setting the derivative to zero. Thus, we have:

$$
\mathbf{p} = \mathbf{z}_i^{(v)} - \frac{1}{\eta} \left(\left(\mathbf{A}^{(v)} \right)^T \mathbf{z}_i^{(v)} + \mathbf{h} \right) \quad (14)
$$

2) Update $\mathbf{z}_i^{(v)}$ with fixed $\mathbf{p}$. The Lagrange function of Eq. (13) w.r.t. $\mathbf{z}_i^{(v)}$ can be written as

$$
\min_{\left(\mathbf{z}_i^{(v)} \right)^T \mathbf{1}=1, z_{ij}^{(v)} \geq 0} \left\| \mathbf{z}_i^{(v)} - \mathbf{p} + \frac{1}{\eta} \mathbf{h} + \frac{\mathbf{A}^{(v)} \mathbf{p} - \mathbf{b}}{\eta} \right\|_2^2,
$$
$$
(15)
$$

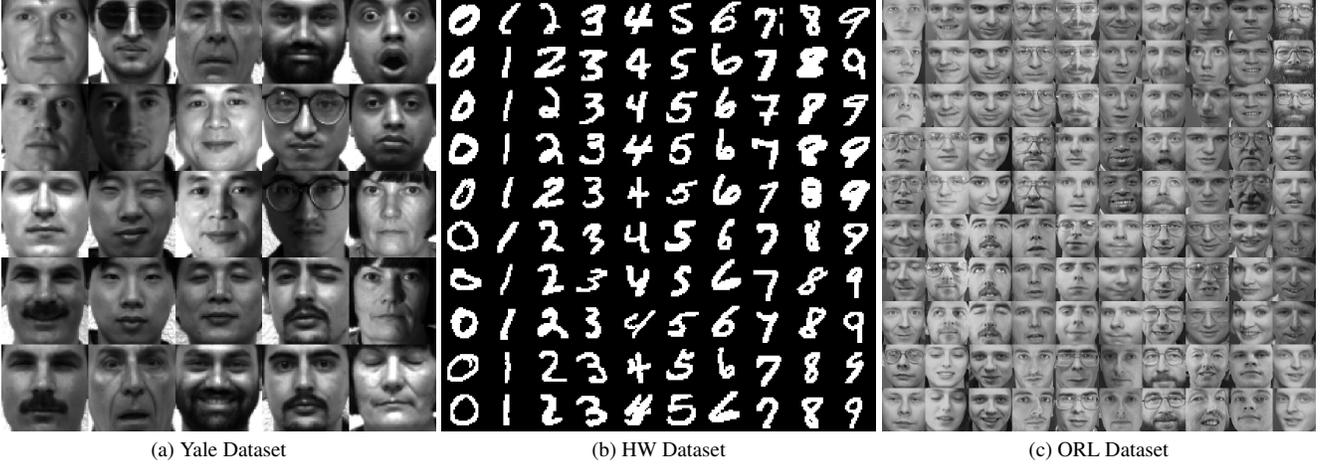(a) Yale Dataset   (b) HW Dataset   (c) ORL Dataset

Figure 2. Sample images of datasets Yale, HW and ORL.

which has a closed-form solution and can be readily achieved by the off-the-shelf optimization algorithm [26].

According to the ALM principles [25], η can be magnified increasingly in each iteration, and $\mathbf{h}$ is updated by $\mathbf{h} \leftarrow \mathbf{h} + \eta\left(\mathbf{z}_i^{(v)} - \mathbf{h}\right)$. The detailed algorithm to solve Eq. (10) is summarized in Algorithm 1.

Update S

Drop all unrelated terms of Eq. (6) w.r.t. $\mathbf{S}$, thus we have

$$\min_{\mathbf{S}} \alpha \sum_{i=1}^{n} \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \right\|_2^2 + \beta \|\mathbf{S}\|_F^2 \tag{16}$$
$$\text{s.t. } \mathbf{S} \geq 0, \mathbf{S1} = \mathbf{1}.$$

We can reformulate Eq. (16) in a vector form as

$$\min_{\mathbf{s}_i} (\alpha + \beta) \, \mathbf{s}_i \mathbf{s}_i^T - 2\alpha \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \mathbf{s}_i \tag{17}$$
$$\text{s.t. } \mathbf{s}_i \mathbf{1} = 1, \mathbf{s}_i \geq 0.$$

Based on Eq. (17), we get the following compact formulation

$$\min_{\mathbf{s}_i \mathbf{1} = 1, \mathbf{s}_i \geq 0} \left\| \mathbf{s}_i - \frac{\alpha}{\alpha + \beta} \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \right\|_2^2, \tag{18}$$

which can be effectively solved by the optimization algorithm proposed in [26].

Update W

Optimizing Eq. (6) w.r.t. $\mathbf{W}$ is equivalent to solving

$$\min_{\mathbf{w}_i} \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \right\|_2^2 \tag{19}$$
$$\text{s.t. } \mathbf{w}_i \geq 0, \mathbf{w}_i^T \mathbf{1} = \mathbf{1}.$$

Eq. (19) can be reorganized as

$$\min_{\mathbf{w}_i} \left\| \mathbf{w}_i^T \mathbf{B}_i \right\|_2^2 \tag{20}$$
$$\text{s.t. } \mathbf{w}_i \geq 0, \mathbf{w}_i^T \mathbf{1} = \mathbf{1}.$$

where $\mathbf{B}_i = \left(\mathbf{1}\mathbf{s}_i^T - \widetilde{\mathbf{Q}}_i \right) \in \mathbb{R}^{m \times n}$.

Taking the derivative of Eq. (20) w.r.t $\mathbf{w}_i$, it yields

$$\mathbf{w}_i = \frac{\left(\mathbf{B}_i \mathbf{B}_i^T \right)^{-1} \mathbf{1}}{\mathbf{1}^T \left(\mathbf{B}_i \mathbf{B}_i^T \right)^{-1} \mathbf{1}}. \tag{21}$$

Update $\widetilde{\mathbf{Q}}_i$

Fixing all the other irrelevant variables in Eq. (6), we have:

$$\min_{\widetilde{\mathbf{Q}}_i} \gamma \sum_{v=1}^{m} \|\mathbf{Z}^{(v)} - \mathbf{Q}^{(v)}\|_F^2 + \alpha \sum_{i=1}^{n} \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \right\|_2^2 \tag{22}$$

Obviously, Eq. (22) is sample-wise independent. Therefore, it can be written as:

$$\min_{\widetilde{\mathbf{Q}}_i} \gamma \|\widetilde{\mathbf{Z}}^{(v)} - \widetilde{\mathbf{Q}}^{(v)}\|_F^2 + \alpha \left\| \mathbf{s}_i^T - \mathbf{w}_i^T \widetilde{\mathbf{Q}}_i \right\|_2^2 \tag{23}$$

Naturally, by taking the derivative of Eq. (23), it yields:

$$\widetilde{\mathbf{Q}}_i = \left(\gamma \mathbf{I} + \alpha \mathbf{w}_i \mathbf{w}_i^T \right)^{-1} \left(\gamma \widetilde{\mathbf{Z}}_i + \alpha \mathbf{w}_i \mathbf{s}_i^T \right) \tag{24}$$

The detailed algorithm to solve the objective in Eq. (6) is summarized in Algorithm 2.

5. Experiment

This section is comprised of four sub-sections: experimental setting, result analysis, robustness analysis and convergence analysis.

5.1. Experiment Setting

We compare our proposed method with several multi-view clustering methods: Multi-view Spectral Clustering with Co-training strategy(Co-train) [27]. Multi-

Table 1. The clustering results on **100leaves** dataset (%)

Method	ACC	NMI	Purity	F-score
Co-train	33.21 ± 0.85	57.59 ± 0.73	36.10 ± 0.81	15.88 ± 0.79
Co-reg	32.16 ± 0.93	58.97 ± 0.82	36.00 ± 0.82	17.33 ± 0.99
DiMSC	48.52 ± 1.57	71.24 ± 0.55	51.14 ± 1.41	33.66 ± 1.17
AMGL	77.66 ± 1.58	90.35 ± 0.97	82.29 ± 1.16	59.54 ± 7.04
MVGL	54.12 ± 0.00	63.96 ± 0.00	57.44 ± 0.00	8.58 ± 0.00
WMSC	77.26 ± 0.99	89.81 ± 0.47	79.62 ± 0.68	71.11 ± 1.31
AWP	66.13 ± 0.00	85.52 ± 0.00	68.19 ± 0.00	60.75 ± 0.00
MCGC	90.25 ± 0.00	93.90 ± 0.00	91.13 ± 0.00	77.79 ± 0.00
LRMSC	75.25 ± 1.56	87.64 ± 0.54	77.60 ± 1.31	66.62 ± 1.46
MVCTM	68.63 ± 0.00	76.46 ± 0.00	72.25 ± 0.00	16.52 ± 0.00
SMVSC	37.94 ± 0.27	65.04 ± 0.29	39.42 ± 0.23	23.22 ± 0.39
FPMVS	35.14 ± 0.20	63.06 ± 0.43	36.96 ± 0.26	22.12 ± 0.38
CSMSC	76.06 ± 1.15	88.83 ± 0.43	78.96 ± 0.81	68.91 ± 1.25
Ours	**90.86 ± 0.01**	**96.38 ± 0.00**	**92.46 ± 0.01**	**88.69 ± 0.01**

Table 4. The clustering results on **ORL** dataset (%)

Method	ACC	NMI	Purity	F-score
Co-train	61.90 ± 3.19	79.23 ± 1.58	65.60 ± 2.23	50.88 ± 3.26
Co-reg	61.65 ± 2.81	79.38 ± 1.69	66.22 ± 2.14	51.39 ± 3.51
DiMSC	80.03 ± 1.78	90.63 ± 0.88	82.93 ± 1.17	74.61 ± 1.91
AMGL	73.15 ± 3.22	85.59 ± 1.09	78.27 ± 2.01	58.89 ± 3.62
MVGL	49.00 ± 0.00	66.30 ± 0.00	58.00 ± 0.00	18.24 ± 0.00
WMSC	76.60 ± 3.70	88.26 ± 1.68	79.75 ± 2.85	69.84 ± 4.38
AWP	74.50 ± 0.00	86.02 ± 0.00	76.25 ± 0.00	65.18 ± 0.00
MCGC	77.00 ± 0.00	87.22 ± 0.00	82.75 ± 0.00	56.25 ± 0.00
LRMSC	79.62 ± 3.09	90.49 ± 1.50	83.42 ± 2.46	73.99 ± 3.96
MVCTM	59.50 ± 0.00	72.58 ± 0.00	66.75 ± 0.00	23.98 ± 0.00
SMVSC	58.28 ± 0.51	76.07 ± 0.14	61.75 ± 0.30	43.96 ± 0.33
FPMVS	55.45 ± 0.59	73.72 ± 0.48	59.27 ± 0.56	41.02 ± 0.69
CSMSC	77.57 ± 1.88	89.20 ± 0.95	80.77 ± 1.78	71.69 ± 2.58
Ours	**84.35 ± 0.02**	**91.78 ± 0.01**	**86.45 ± 0.01**	**78.58 ± 0.02**

Table 2. The clustering results on **HW** dataset (%)

Method	ACC	NMI	Purity	F-score
Co-train	81.28 ± 5.37	72.04 ± 2.98	82.25 ± 4.11	70.93 ± 4.19
Co-reg	54.96 ± 4.35	47.46 ± 1.97	58.64 ± 2.87	42.62 ± 2.40
DiMSC	46.26 ± 0.45	32.97 ± 0.33	46.79 ± 0.47	27.92 ± 0.27
AMGL	93.76 ± 8.57	95.02 ± 4.23	94.89 ± 6.78	93.63 ± 7.16
MVGL	85.35 ± 0.00	85.19 ± 0.00	85.75 ± 0.00	82.48 ± 0.00
WMSC	74.87 ± 0.14	66.79 ± 0.19	74.95 ± 0.16	63.65 ± 0.22
AWP	69.55 ± 0.00	59.49 ± 0.00	71.45 ± 0.00	58.35 ± 0.00
MCGC	53.95 ± 0.00	61.78 ± 0.00	54.10 ± 0.00	57.25 ± 0.00
LRMSC	79.39 ± 2.95	76.49 ± 1.33	83.25 ± 1.47	72.83 ± 1.89
MVCTM	98.75 ± 0.00	96.91 ± 0.00	98.75 ± 0.00	97.52 ± 0.00
SMVSC	78.59 ± 0.46	69.03 ± 0.57	78.90 ± 0.38	67.65 ± 0.76
FPMVS	78.40 ± 0.18	70.88 ± 0.89	78.77 ± 0.41	69.66 ± 0.85
CSMSC	86.87 ± 0.02	80.04 ± 0.09	86.87 ± 0.02	78.05 ± 0.04
Ours	**99.25 ± 0.00**	**98.15 ± 0.00**	**99.25 ± 0.00**	**98.51 ± 0.00**

Table 5. The clustering results on **Cornell** dataset (%)

Method	ACC	NMI	Purity	F-score
Co-train	41.38 ± 1.67	22.36 ± 1.43	54.26 ± 1.90	34.86 ± 0.90
Co-reg	36.26 ± 2.00	13.68 ± 1.45	47.28 ± 1.77	31.98 ± 1.41
DiMSC	40.15 ± 1.76	17.95 ± 0.82	52.77 ± 0.67	34.58 ± 0.57
AMGL	37.38 ± 4.44	4.26 ± 0.37	44.82 ± 0.34	37.24 ± 3.51
MVGL	44.10 ± 0.00	9.42 ± 0.00	45.64 ± 0.00	38.10 ± 0.00
WMSC	46.10 ± 0.15	21.69 ± 0.17	52.26 ± 0.15	37.24 ± 0.12
AWP	38.97 ± 0.00	13.63 ± 0.00	49.74 ± 0.00	33.72 ± 0.00
MCGC	36.41 ± 0.00	22.60 ± 0.00	53.85 ± 0.00	33.79 ± 0.00
LRMSC	34.77 ± 1.62	8.51 ± 1.48	43.95 ± 0.76	28.99 ± 0.72
MVCTM	40.00 ± 0.00	15.20 ± 0.00	51.79 ± 0.00	36.89 ± 0.00
SMVSC	48.67 ± 2.21	20.61 ± 1.36	53.90 ± 0.87	**42.14 ± 2.28**
FPMVS	35.90 ± 0.00	17.67 ± 0.00	53.33 ± 0.00	33.20 ± 0.00
CSMSC	47.33 ± 0.33	23.86 ± 0.36	55.49 ± 0.21	37.87 ± 0.37
Ours	**51.28 ± 0.00**	**34.81 ± 0.00**	**61.03 ± 0.00**	42.03 ± 0.00

Table 3. The clustering results on **MSRC** dataset (%)

Method	ACC	NMI	Purity	F-score
Co-train	64.95 ± 6.24	54.27 ± 4.18	66.52 ± 5.00	52.59 ± 4.97
Co-reg	63.95 ± 4.50	57.89 ± 2.99	67.24 ± 4.47	55.31 ± 3.98
DiMSC	71.90 ± 0.95	59.62 ± 0.96	71.90 ± 0.95	58.02 ± 1.08
AMGL	76.71 ± 4.33	70.98 ± 2.20	79.00 ± 2.80	67.13 ± 2.49
MVGL	70.48 ± 0.00	58.18 ± 0.00	70.48 ± 0.00	54.56 ± 0.00
WMSC	69.05 ± 0.00	59.55 ± 0.17	71.43 ± 0.00	57.57 ± 0.08
AWP	63.33 ± 0.00	54.88 ± 0.00	63.33 ± 0.00	53.76 ± 0.00
MCGC	80.48 ± 0.00	70.18 ± 0.00	80.95 ± 0.00	72.46 ± 0.00
LRMSC	71.81 ± 2.89	62.42 ± 1.58	74.14 ± 1.78	60.08 ± 1.57
MVCTM	85.71 ± 0.00	76.15 ± 0.00	85.71 ± 0.00	74.90 ± 0.00
SMVSC	81.43 ± 0.00	70.18 ± 0.00	81.43 ± 0.00	69.36 ± 0.00
FPMVS	78.57 ± 0.00	66.84 ± 0.00	78.57 ± 0.00	68.36 ± 0.00
CSMSC	80.43 ± 0.14	71.36 ± 0.19	80.43 ± 0.14	70.06 ± 0.21
Ours	**87.14 ± 0.00**	**78.21 ± 0.00**	**87.14 ± 0.00**	**77.87 ± 0.00**

Table 6. The clustering results on **Yale** dataset(%)

Method	ACC	NMI	Purity	F-score
Co-train	53.94 ± 4.21	57.90 ± 3.14	55.52 ± 4.11	38.68 ± 3.72
Co-reg	56.36 ± 3.49	60.30 ± 3.05	57.64 ± 3.25	41.43 ± 3.79
DiMSC	64.36 ± 3.40	66.57 ± 2.30	64.79 ± 3.22	49.41 ± 2.90
AMGL	59.94 ± 3.64	62.95 ± 2.78	61.33 ± 2.93	41.87 ± 4.18
MVGL	44.85 ± 0.00	49.19 ± 0.00	46.67 ± 0.00	26.05 ± 0.00
WMSC	61.88 ± 3.64	65.40 ± 2.89	62.12 ± 3.84	48.08 ± 3.53
AWP	61.82 ± 0.00	65.77 ± 0.00	61.82 ± 0.00	49.71 ± 0.00
MCGC	67.27 ± 0.00	68.92 ± 0.00	67.27 ± 0.00	48.33 ± 0.00
LRMSC	68.36 ± 1.11	70.36 ± 1.11	68.55 ± 1.10	52.27 ± 1.67
MVCTM	62.42 ± 0.00	66.14 ± 0.00	62.42 ± 0.00	47.40 ± 0.00
SMVSC	60.36 ± 0.95	62.23 ± 1.37	60.73 ± 0.89	43.05 ± 1.51
FPMVS	44.24 ± 0.00	49.76 ± 0.00	46.67 ± 0.00	30.45 ± 0.00
CSMSC	64.06 ± 2.91	68.22 ± 1.49	64.12 ± 2.80	51.84 ± 1.76
Ours	**71.03 ± 0.01**	**74.51 ± 0.01**	**71.21 ± 0.00**	**59.72 ± 0.01**

view Spectral Clustering with Co-regularized strategy(Co-reg) [28]. Diversity-induced Multi-view Subspace Clustering(DiMSC) [14]. Auto-weighted Multiple Graph Learning(AMGL) [29]. Consistent and Specific Multi-View Subspace Clustering(CSMSC) [6]. Graph Learning for Multi-view Clustering(MVGL) [30] . Weighted Multi-view Spec-

tral Clustering(WMSC) [31]. multi-view clustering via Adaptively Weighted Procrustes(AWP) [32]. Multi-view Consensus Graph Clustering(MCGC) [33]. Generalized Latent Multi-View Subspace Clustering(LRMSC) [34]. Multi-view Clustering on Topological Manifold(MVCTM) Fast Parameter-free Multi-view Subspace Clustering with
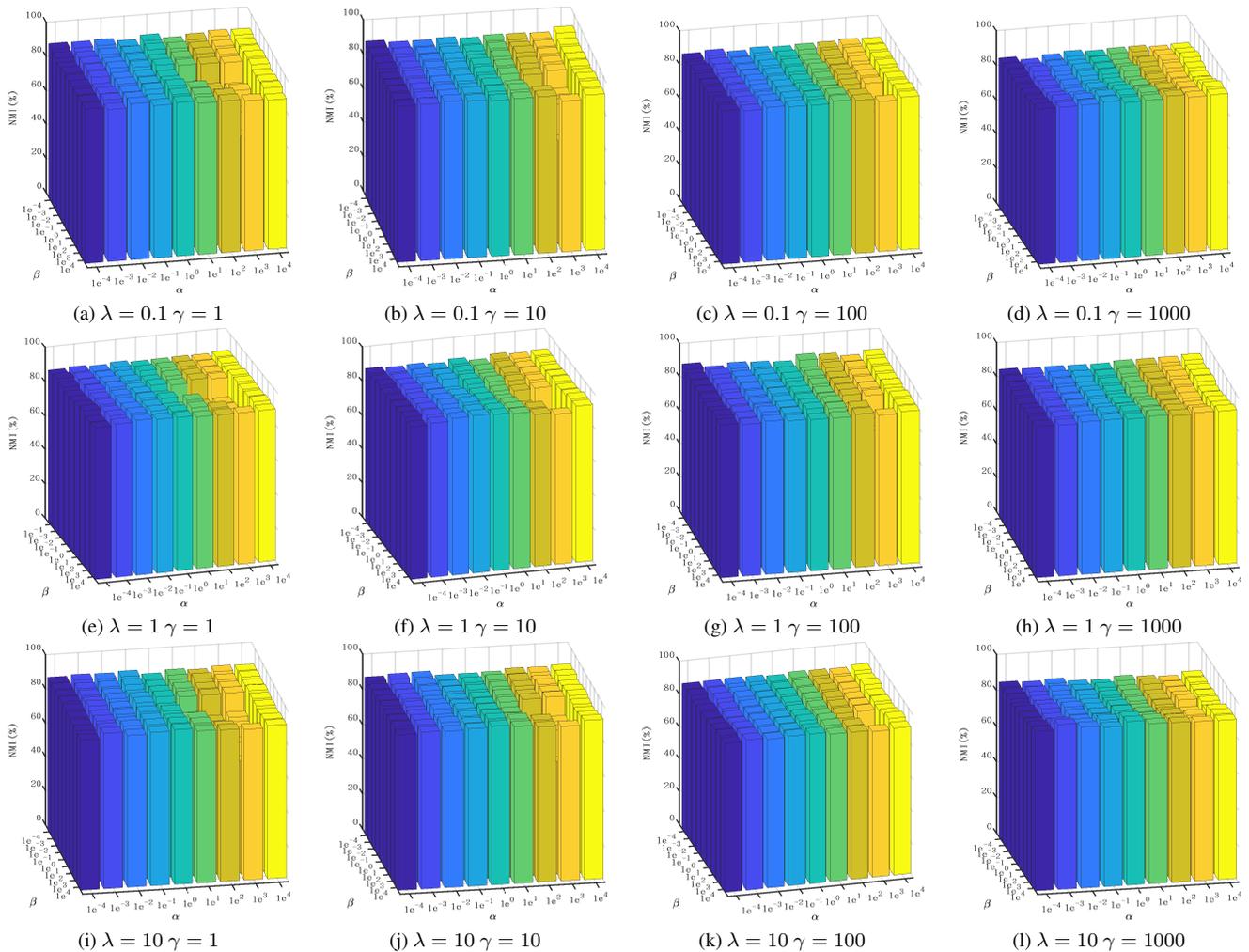
(a) $\lambda = 0.1\ \gamma = 1$ (b) $\lambda = 0.1\ \gamma = 10$ (c) $\lambda = 0.1\ \gamma = 100$ (d) $\lambda = 0.1\ \gamma = 1000$

(e) $\lambda = 1\ \gamma = 1$ (f) $\lambda = 1\ \gamma = 10$ (g) $\lambda = 1\ \gamma = 100$ (h) $\lambda = 1\ \gamma = 1000$

(i) $\lambda = 10\ \gamma = 1$ (j) $\lambda = 10\ \gamma = 10$ (k) $\lambda = 10\ \gamma = 100$ (l) $\lambda = 10\ \gamma = 1000$

Figure 3. Parameter analysis on ORL dataset

Consensus Anchor Guidance(FPMVS) [35]. Multi-view Subspace Clustering with Unified Anchors(SMVSC) [36] Consistent and Specific Multi-View Subspace Clustering(CSMSC) [6]. Rather than adopting view-wise weighing strategy like the above mentioned methods, we delicately conduct graph fusion in a sample-wise manner. Thus, the superiority of our proposed approach could be demonstrated in the experiment section.

As for benchmark datasets, we adopt **100leaves**, **HW**, **MRSC**, **ORL**, **Cornell** and **Yale**. Specifically, the leaves data set (**100leaves**) consists of 1,600 leaves from each of 100 plant species. For each sample, margin, shape and texture are given as three distinct views. Handwritten numerals data (**HW**) is composed of 2,000 data points for digits 0 to 9 from UCI machine learning repository and two public features are available. **MSRC** data set contains 210 images and can be separated into 7 classes including cow, airplane, face, bicycle, tree, building and car. There are five visual

features from each sample. **ORL** consists of 400 face images in 40 different themes in total. For each subject, the images are described in three features: facial expressions, facial details and the lighting. **Cornell**, which is collected by Cornell University, contains 195 web pages and a web page is made of 2 views: content features and cites features. Dataset **Yale** contains 165 gray-scale images under varying poses and illuminations, where every image is of size 64×64.

5.2. Result Analysis

With the aim of evaluating our proposed approach adequately with above mentioned SOTA multi-view clustering methods, we adopt four widely used criteria including Normalized Mutual Information(NMI), Accuracy(ACC), Purity and F-Score. The clustering results are reported in Tables (1)- (6). We can arrive at a conclusion that our method is very effective and competitive as the proposed method
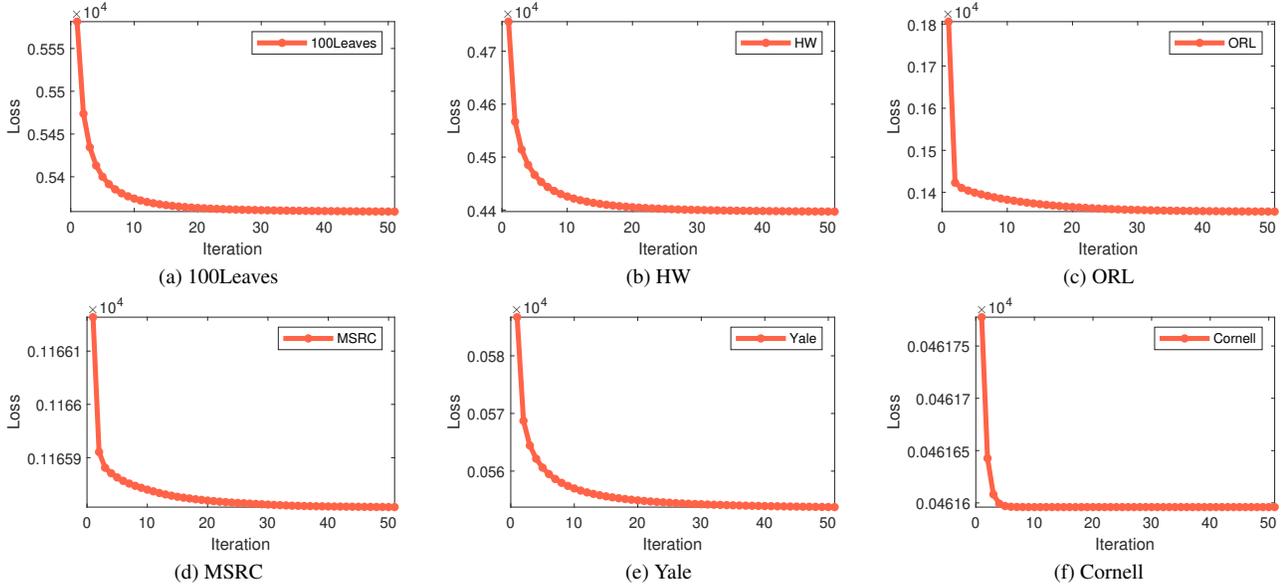
Figure 4. Convergence curves on all datasets

outperforms other competitors in the majority of cases. In detail, our method consistently obtains the best results in terms of NMI, ACC and Purity on all datasets. While for F-score, the proposed method surpasses the compared methods except in one case on dataset **Cornell**. The superiority of our method explicitly demonstrates the effectiveness of the sample-wise learning strategy and corroborates our theoretical findings.

5.3. Parameter Discussion

With the aim of studying what impact different parameter settings will have on the clustering results, we vary α, β, λ and γ in the ranges $\left[1e^{-4}, \ldots, 1e^{4}\right]$, $\left[1e^{-4}, \ldots, 1e^{4}\right]$, $\left[1e^{-1}, \ldots, 1e^{2}\right]$ and $\left[1, \ldots, 1e^{3}\right]$, respectively. Taking the ORL dataset as an example, we can see the clustering performance is quite stable with respect to different parameter settings, as shown in Figure (3). Considering that there is little difference among the performances of diverse parameter combinations, we can come to a conclusion that our approach is robust to the hyper parameters.

5.4. Convergence Analysis

Owing to the fact that the optimization of our proposed method is essentially a non-convex problem that is solved by an iterative algorithm, it is critical to validate the convergence of our model. Therefore, this section empirically showcases the convergence property and how fast our algorithm can converge. The convergence curves of our model on six datasets shown in Fig. (4) demonstrate the effectiveness of our optimization approach. Note that despite the non-convexity of Eq. (6), our model can still discover an optimal solution for each variable and achieve a local minima within a few iterations, which verifies the effectiveness of the proposed optimization algorithm.

6. Conclusion

In this paper, we propose to exploit the implied data manifold by learning the topological structure of data. Besides, considering that the consistency of multiple views is manifested in the generally similar local structure while the inconsistent structures are the minority, we further explore the intersections of multiple views in the sample level such that the cross-view consistency can be better maintained. By leveraging the subtasks of topological relevance learning and the sample-level graph fusion in our collaborative model, each subtask is alternately boosted towards an optimal solution. We also design an efficient algorithm to solve the corresponding optimization problem. Meanwhile, our method generates outstanding results on several authoritative benchmark datasets and is proven to outperform current state-of-the-art methods. In the future, we are interested in extending the proposed model to other machine learning framework such as kernel learning and deep learning.

References

[1] Steffen Bickel and Tobias Scheffer. Multi-view clustering. In *Proceedings of the 4th IEEE International Conference on Data Mining*, pages 19–26, 2004. 1

[2] Shudong Huang, Zenglin Xu, Ivor W Tsang, and Zhao Kang. Auto-weighted multi-view co-clustering with bipartite graphs. *Information Sciences*, 512:18–30, 2020. 1

[3] Hongchang Gao, Feiping Nie, Xuelong Li, and Heng Huang. Multi-view subspace clustering. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4238–4246, 2015. 1

[4] Shudong Huang, Ivor W Tsang, Zenglin Xu, and Jiancheng Lv. Cgdd: Multiview graph clustering via cross-graph diversity detection. *IEEE Transactions on Neural Networks and Learning Systems*, 2022. 1

[5] Changqing Zhang, Qinghua Hu, Huazhu Fu, Pengfei Zhu, and Xiaochun Cao. Latent multi-view subspace clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4279–4287, 2017. 1

[6] Shirui Luo, Changqing Zhang, Wei Zhang, and Xiaochun Cao. Consistent and specific multi-view subspace clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3730–3737, 2018. 1, 6, 7

[7] Zhiyong Yang, Qianqian Xu, Weigang Zhang, Xiaochun Cao, and Qingming Huang. Split multiplicative multi-view subspace clustering. *IEEE Transactions on Image Processing*, 28(10):5147–5160, 2019. 1

[8] Hong Tao, Chenping Hou, Yuhua Qian, Jubo Zhu, and Dongyun Yi. Latent complete row space recovery for multi-view subspace clustering. *IEEE Transactions on Image Processing*, 29:8083–8096, 2020. 1

[9] Lynn Houthuys, Rocco Langone, and Johan A.K. Suykens. Multi-view kernel spectral clustering. *Information Fusion*, 44:46–56, 2018. 1

[10] Shudong Huang, Zhao Kang, Ivor W Tsang, and Zenglin Xu. Auto-weighted multi-view clustering via kernelized graph learning. *Pattern Recognition*, 88:174–184, 2019. 1

[11] Yongyong Chen, Xiaolin Xiao, and Yicong Zhou. Jointly learning kernel representation tensor and affinity matrix for multi-view clustering. *IEEE Transactions on Multimedia*, 22(8):1985–1997, 2020. 1

[12] Xiaotong Zhang, Xianchao Zhang, Han Liu, and Xinyue Liu. Multi-task multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 28(12):3324–3338, 2016. 1

[13] Yiling Zhang, Yan Yang, Tianrui Li, and Hamido Fujita. A multitask multiview clustering algorithm in heterogeneous situations based on lle and le. *Knowledge-Based Systems*, 163:776–786, 2019. 1

[14] Xiaochun Cao, Changqing Zhang, Huazhu Fu, Si Liu, and Hua Zhang. Diversity-induced multi-view subspace clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–594, 2015. 1, 6

[15] Yan Yang and Hao Wang. Multi-view clustering: A survey. *Big Data Mining and Analytics*, 1(2):83–107, 2018. 1

[16] Rui Zhang, Feiping Nie, Xuelong Li, and Xian Wei. Feature selection with multi-view data: A survey. *Information Fusion*, 50:158–167, 2019. 1

[17] Shudong Huang, Ivor W Tsang, Zenglin Xu, and Jiancheng Lv. Measuring diversity in graph learning: a unified framework for structured multi-view clustering. *IEEE Transactions on Knowledge and Data Engineering*, 34(12):5869–5883, 2021. 1

[18] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, 2003. 2

[19] Zhenyue Zhang, Jing Wang, and Hongyuan Zha. Adaptive manifold learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(2):253–265, 2011. 2

[20] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000. 2

[21] Michele Ballerini, Nicola Cabibbo, Raphael Candelier, Andrea Cavagna, Evaristo Cisbani, Irene Giardina, Vivien Lecomte, Alberto Orlandi, Giorgio Parisi, Andrea Procaccini, et al. Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4):1232–1237, 2008. 2

[22] Qi Wang, Mulin Chen, and Xuelong Li. Quantifying and detecting collective motion by manifold learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 4292–4298, 2017. 2

[23] Quanxue Gao, Wei Xia, Zhizhen Wan, Deyan Xie, and Pu Zhang. Tensor-svd based graph learning for multi-view subspace clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3930–3937, 2020. 3

[24] Yidi Wang, Xiaobing Pei, and Haoxi Zhan. Fine-grained graph learning for multi-view subspace clustering. *arXiv preprint arXiv:2201.04604*, 2022. 3

[25] Dimitri P Bertsekas. Nonlinear programming. *Journal of the Operational Research Society*, 48(3):334–334, 1997. 4, 5

[26] Jin Huang, Feiping Nie, and Heng Huang. A new simplex sparse learning model to measure data similarity for clustering. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, page 3569–3575, 2015. 5

[27] Abhishek Kumar and Hal Daumé. A co-training approach for multi-view spectral clustering. In *Proceedings of the 28th International Conference on Machine Learning*, pages 393–400, 2011. 5

[28] Abhishek Kumar, Piyush Rai, and Hal Daume. Co-regularized multi-view spectral clustering. *Advances in Neural Information Processing Systems*, 24:1–9, 2011. 6

[29] Feiping Nie, Jing Li, and Xuelong Li. Parameter-free auto-weighted multiple graph learning: A framework for multi-view clustering and semi-supervised classification. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 1881–1887, 2016. 6

[30] Kun Zhan, Changqing Zhang, Junpeng Guan, and Junsheng Wang. Graph learning for multiview clustering. *IEEE Transactions on Cybernetics*, 48(10):2887–2895, 2017. 6

[31] Linlin Zong, Xianchao Zhang, Xinyue Liu, and Hong Yu. Weighted multi-view spectral clustering based on spectral perturbation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 4621–4628, 2018. 6

[32] Feiping Nie, Lai Tian, and Xuelong Li. Multiview clustering via adaptively weighted procrustes. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2022–2030, 2018. 6

[33] Kun Zhan, Feiping Nie, Jing Wang, and Yi Yang. Multiview consensus graph clustering. *IEEE Transactions on Image Processing*, 28(3):1261–1270, 2019. 6

[34] Changqing Zhang, Huazhu Fu, Qinghua Hu, Xiaochun Cao, Yuan Xie, Dacheng Tao, and Dong Xu. Generalized latent multi-view subspace clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(1):86–99, 2018. 6

[35] Siwei Wang, Xinwang Liu, Xinzhong Zhu, Pei Zhang, Yi Zhang, Feng Gao, and En Zhu. Fast parameter-free multi-view subspace clustering with consensus anchor guidance. *IEEE Transactions on Image Processing*, 31:556–568, 2021. 7

[36] Mengjing Sun, Pei Zhang, Siwei Wang, Sihang Zhou, Wenxuan Tu, Xinwang Liu, En Zhu, and Changjian Wang. Scalable multi-view subspace clustering with unified anchors. In *Proceedings of the 29th ACM International Conference on Multimedia*, pages 3528–3536, 2021. 7