2023 IEEE 39th International Conference on Data Engineering (ICDE) | 979-8-3503-2227-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICDE55515.2023.00307

Consensus One-step Multi-view Subspace Clustering (Extended abstract)

Pei Zhang[1], Xinwang Liu[1] *, Jian Xiong [2], Sihang Zhou [3], Wentao Zhao [1], En Zhu [1], Zhiping Cai [1]

[1] *School of Computer, National University of Defense Technology, Changsha, China*

[2] *School of Business Administration, Southwestern University of Finance and Economics, Chengdu, China*

[3] *School of Intelligence Science and Technology, National University of Defense Technology, Changsha, China*

{zhangpei, xinwangliu, wtzhao, enzhu, zpcai}@nudt.edu.cn, xiongjian2017@swufe.edu.cn, sihangjoe@gmail.com

***Abstract*—Multi-view clustering has attracted increasing attention in data mining communities. Despite superior clustering performance, we observe that existing multi-view subspace clustering methods *directly fuse multi-view information in the similarity level by merging noisy affinity matrices*; and *isolate the processes of affinity learning, multiple information fusion and clustering*. Both factors may cause insufficient utilization of multi-view information, leading to unsatisfying clustering performance. This paper proposes a novel consensus one-step multi-view subspace clustering (COMVSC) method to address these issues. Instead of directly fusing affinity matrices, COMVSC optimally integrates discriminative partition-level information, which is helpful in eliminating noise among data. Moreover, the affinity matrices, consensus representation and final clustering labels are learned simultaneously in a unified framework. Extensive experiment results on benchmark datasets demonstrate the superiority of our method over other state-of-the-art approaches.**

***Index Terms*—Multi-view Clustering, Subspace Clustering, Data Fusion.**

I. Introduction

Multi-view clustering (MVC) aims to combine multiple feature information and search for consistent clustering results across views. MVC methods can be summarized into four categories: co-training, multi-kernel clustering, multi-view matrix factorization clustering and multi-view subspace clustering (MVSC). Existing MVSC methods can be improved from the following two points: i) The majority of existing MVSC methods learn a shared affinity matrix or graph and then apply spectral clustering to obtain the final clustering result. However, directly learning a common graph from the original data may affect the clustering structure since the original space often consists of noise and redundancy. ii) Previous approaches are usually conducted in a two-step fashion, which isolates the representation learning from the clustering task. To address these issues, we propose a novel Consensus One-step Multi-view Subspace Clustering (COMVSC) method to fuse the multiple partition-level information and integrate the representation learning and clustering into a unified framework.

II. Proposed Approach

Given n data point $\mathbf{X} \in \mathbb{R}^{d\times n}$, the self-representation method is utilized to express each data point with a linear combination of the data themselves. It can be formulated as $\mathbf{X} = \mathbf{X}\mathbf{Z} + \mathbf{E}$, where $\mathbf{Z}$ is the subspace representation matrix with each column being the representation of the corresponding data point. $\mathbf{E}$ is the noise matrix. When datasets have multiple features $\{\mathbf{X}^v\}_{v=1}^m \in \mathbb{R}^{d_v\times n}$, the framework of existing MVSC methods can be summarized as follows,

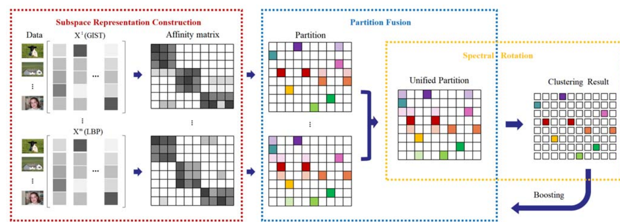


Fig. 1. Framework of the proposed COMVSC algorithm.

$$\begin{aligned} \min_{\mathbf{Z}^v,\mathbf{Z}^*} \ & \mathbf{L}(\mathbf{X}^v, \mathbf{X}^v\mathbf{Z}^v) + \lambda\, \boldsymbol{\Omega}(\mathbf{Z}^v) + \mathbf{Cons}(\mathbf{Z}^v, \mathbf{Z}^*) \\ \text{s.t. } & 0 \le Z_{i,j}^v \le 1, (\mathbf{Z}^v)^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathrm{diag}(\mathbf{Z}^v) = 0\ , \end{aligned} \tag{1}$$

where $\mathbf{Z}^v \in \mathbb{R}^{n\times n}$ is regarded as the subspace representation matrix of v-th view while $\mathbf{Z}^* \in \mathbb{R}^{n\times n}$ is the consensus subspace representation across multiple views. $\mathbf{Cons}(\cdot)$ refers to strategies for reaching a consensus from several view-specific subspace representations. Considering the limitations of similarity-level fusion and the isolation of the two-step strategy, we proposed the following COMVSC method in Eq. (2). In this framework, as shown in Figure 1, we jointly conduct the representation learning, partition fusion and clustering.

$$\begin{aligned} \min_{\substack{\mathbf{F}^*,\mathbf{R},\mathbf{Y},\\ \mathbf{F}^v,\mathbf{Z}^v}} & \underbrace{\sum_{v=1}^{m} \|\mathbf{X}^v - \mathbf{X}^v\mathbf{Z}^v\|_F^2 + \lambda\|\mathbf{Z}^v\|_F^2 + \mathrm{Tr}((\mathbf{F}^v)^{\mathrm{T}}\mathbf{L}^v\mathbf{F}^v)}_{Subspace\ Representation\ Construction} \\ & + \underbrace{\sum_{v=1}^{m} \|\mathbf{F}^v - \mathbf{F}^*\|_F^2}_{Partition\ Fusion} + \underbrace{\sum_{i=1}^{n}\sum_{c=1}^{k} (Y_{i,c})^{\gamma} \|\mathbf{t}_c - \mathbf{F}^*_{i,:}\mathbf{R}\|_2^2}_{Spectral\ Rotation} \\ \text{s.t. } & 0 \le Z_{i,j}^v \le 1, (\mathbf{Z}^v)^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \mathrm{diag}(\mathbf{Z}^v) = 0, \mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}_k, \\ & (\mathbf{F}^v)^{\mathrm{T}}\mathbf{F}^v = \mathbf{I}_k, (\mathbf{F}^*)^{\mathrm{T}}\mathbf{F}^* = \mathbf{I}_k, Y_{i,c} \ge 0,\ \mathbf{Y}_{i,:}\mathbf{1}_k = 1\ , \end{aligned} \tag{2}$$

where $\mathbf{F}^v \in \mathbb{R}^{n\times k}$ is the clustering indicator matrix of v-th view. Rotation matrix $\mathbf{R} \in \mathbb{R}^{k\times k}$ is introduced to jointly optimize the representation and the clustering results. $Y_{i,c}$ signifies the probability of the i-th sample belonging to the c-th cluster. $\mathbf{t}_c$ terms the one-hot vector that the c-th element equals

2375-026X/23/$31.00 ©2023 IEEE
DOI 10.1109/ICDE55515.2023.00307

to 1 and others are 0, where $c \in \{1, 2, \cdots, k\}$. Based on the assumption that the clustering structure of individuals should be similar, we propose to fuse the partition information $\mathbf{F}^v$ of different views into a consensus one $\mathbf{F}^*$. Furthermore, spectral rotation $\mathbf{R}$ is introduced to the consensus indicator matrix to obtain clustering labels directly, avoiding the additional post-process step in previous methods. In this way, the affinity matrices, consensus partition, and final clustering labels matrix are learned simultaneously in a unified framework. The three steps negotiate with each other to better serve the clustering tasks, leading to promising clustering performance. Then, we propose an iterative algorithm to optimize each variable with others fixed. Entire process is summarized in Algorithm 1.

Algorithm 1 Algorithm of COMVSC

Input: $\{\mathbf{X}^v\}_{v=1}^m$, clusters k, hyper-parameters λ and γ.
Output: Probability clustering labels $\mathbf{Y}$.

while not converged **do**
 $\mathbf{Z}^v_{i,:} = \max\left(\hat{\mathbf{Z}}^v_{i,:} + \alpha_i \mathbf{1}, 0\right)$, α_i is Lagrange multiplier.
 $\max_{\mathbf{F}^*} \operatorname{Tr}((\mathbf{F}^*)^{\mathrm{T}}\mathbf{N})$, where $\mathbf{N} = 2\sum_{v=1}^m \mathbf{F}^v + \mathbf{G}\mathbf{R}^{\mathrm{T}}$.
 Update $\mathbf{F}^v$ by algorithm [1].
 $\max_{\mathbf{R}} \operatorname{Tr}(\mathbf{R}^{\mathrm{T}}\mathbf{H})$, where $\mathbf{H} = (\mathbf{F}^*)^{\mathrm{T}}\mathbf{G}$.
 $Y_{i,c} = \frac{(P_{i,c})^{\frac{1}{1-\gamma}}}{\sum_{c=1}^k (P_{i,c})^{\frac{1}{1-\gamma}}}$.
end while
return clustering labels $\mathbf{Y}$.

III. Experiments

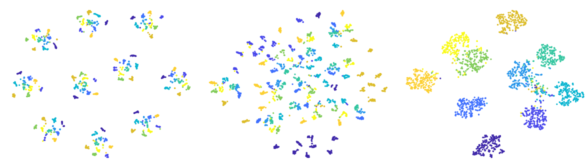
In this section, we extensively evaluate the clustering property of the proposed method on seven widely used multi-view benchmark datasets. The performance of COMVSC is compared with a single-view clustering algorithm and six state-of-the-art multi-view methods. Experimental results in Table I have well demonstrated the effectiveness of our proposed COMVSC in comparison with other compared methods.

TABLE I
The accuracy of compared methods.

Datasets	BBCSport	YALE	Cornell	MSRCv1	Wiki	webkb	HW
FeaCon	0.3448	0.5091	0.3564	0.4541	0.2338	0.5821	0.5830
MLRSSC [2]	0.4086	0.2539	0.3864	0.4852	0.3033	0.9049	0.3702
LMSC [3]	0.4741	0.7394	0.3487	0.7476	0.4141	0.9163	0.6135
RMKMC [4]	0.3190	0.6121	0.4308	0.7095	0.5743	0.7812	0.6710
mPAC [5]	0.6121	0.7636	0.5692	0.8143	0.5382	0.8211	0.8890
FMR [6]	0.3793	0.6303	0.4000	0.8524	0.5556	0.5423	0.6660
GMC [7]	0.5603	0.6788	0.3692	0.8952	0.3939	0.7869	0.8820
LMVSC [8]	0.4828	0.4849	0.4410	0.8810	0.5137	0.8820	0.8540
Ours	**0.6983**	**0.9152**	**0.5949**	**0.9571**	**0.6061**	**0.9343**	**0.9450**

We summarize the superiority of the proposed approach in two aspects: 1) COMVSC employs a joint fusion to optimize self-representation, partition matrices, and clustering labels. The three steps can negotiate with each other to best serve the clustering task, leading to improved performance. 2) COMVSC combines high-level partition indicators with more information and less noise and redundancy. These two factors contribute to improvement in clustering performance.

Besides, we use t-Distributed Stochastic Neighbor Embedding (t-SNE) to visualize the learned representation during the iterations in Fig. 2. As the algorithm is iterated, the clustering structure becomes clearer, which visually demonstrates the feasibility and validity of the proposed method.



(a) 1 iter (b) 5 iter (c) 50 iter

Fig. 2. The cluster structure evolution on Handwritten dataset.

To further illustrate the effectiveness of the one-step strategy, we remove the last spectral rotation term in Eq. (2) and feed the consensus $\mathbf{F}^*$ into k-means to get the clustering results (w/o rotation). Compared with the two-step strategy, our algorithm outperforms it on all datasets in terms of all metrics. The notable results in Table II demonstrate the effectiveness and importance of the one-step strategy.

TABLE II
Comparison between one-step and two-step w.r.t accuracy.

Dataset	BBCSport	yaleA	Wikipedia	Handwritten
w/o rotation	0.5603	0.8000	0.5498	0.9200
COMVSC	**0.6983**	**0.9273**	**0.6061**	**0.9450**

IV. Conclusion

In this paper, we propose a unified framework that jointly optimizes graphs, partitions and final clustering labels, which avoids the sub-optimal solution of existing two-step approaches. Our COMVSC incorporates multiple information at the partition level, which not only preserves the view-specific local clustering structure but also guarantees consistency among multiple views. Extensive experiments demonstrate the effectiveness and superiority of our proposed method.

References

[1] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, no. December, pp. 977–986, 2014.

[2] M. Brbić and I. Kopriva, "Multi-view low-rank sparse subspace clustering," *Pattern Recognition*, vol. 73, pp. 247–258, 2018.

[3] C. Zhang, Q. Hu, H. Fu, P. Zhu, and X. Cao, "Latent multi-view subspace clustering," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[4] H. H. Xiao Cai, Feiping Nie, "Multi-view k-means clustering on big data," *IJCAI*, vol. 31, no. 5, pp. 724–730, 2013.

[5] Z. Kang, Z. Guo, S. Huang, S. Wang, W. Chen, Y. Su, and Z. Xu, "Multiple partitions aligned clustering," *IJCAI*, pp. 2701–2707, 2019.

[6] R. Li, C. Zhang, Q. Hu, P. Zhu, and Z. Wang, "Flexible multi-view representation learning for subspace clustering," *IJCAI*, vol. 2019-Augus, pp. 2916–2922, 2019.

[7] H. Wang, Y. Yang, and B. Liu, "GMC: Graph-based multi-view clustering," *IEEE Transactions on Knowledge and Data Engineering*, 2019.

[8] Z. Kang, W. Zhou, Z. Zhao, J. Shao, M. Han, and Z. Xu, "Large-scale multi-view subspace clustering in linear time," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 04, 2020, pp. 4412–4419.