Multi-view Subspace Clustering with Joint Tensor Representation and Indicator Matrix Learning

Jing Wang[1], Xiaoqian Zhang[1,2,3], Zhigui Liu[1(✉)], Zhuang Yue[1], and Zhengliang Huang[1]

[1] School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China
liuzhigui@swust.edu.cn, yuez@mails.swust.edu.cn
[2] SCII Innovation Center for Convergence and Innovation Industry Technology, Mianyang 621010, China
[3] Tianfu Institute of Research and Innovation, Southwest University of Science and Technology, Chengdu 610213, China

Abstract. Multi-view subspace clustering (MVSC), as an extension of single-view subspace clustering, can exploit more information and has achieved excellent performance. In particular, the MVSC methods with sparse and low-rank basing have become a research priority as they can improve the clustering effect in an effective way. However, the following problems still exist: 1) focusing only on the connections between two views, ignoring the relationship of higher-order views; 2) performing representation matrix learning and indicator matrix learning separately, unable to get the clustering result in one step and obtain the global optimal solution. To tackle these issues, a novel sparsity and low-rank based MVSC algorithm is designed. It jointly conducts tensor representation learning and indicator matrix learning. More specifically, the Tensor Nuclear Norm (TNN) is utilized to exploit the relationships among higher-order views; besides, by incorporating the subsequent spectral clustering, the indicator matrix learning is conducted during the optimization framework. An iterative algorithm, the alternating direction method of multipliers (ADMM) is derived for the solving of the proposed algorithm. Experiments over five baseline datasets prove the competitiveness and excellence of the presented method with comparisons to other eight state-of-the-art algorithms.

Keywords: Multi-view subspace clustering · Tensor nuclear norm · Spectral clustering · Higher-order correlations

Supported by National Natural Science Foundation of China (Grant No. 62102331, 62176125, 61772272), Natural Science Foundation of Sichuan Province (Grant No. 2022NSFSC0839), Doctoral Research Foundation of Southwest University of Science and Technology (Grant No. 22zx7110).

L. Fang et al. (Eds.): CICAI 2022, LNAI 13605, pp. 450–461, 2022.
https://doi.org/10.1007/978-3-031-20500-2_37

1 Introduction

As a typical data analysis methodology, clustering is already widely used in different fields, like artificial intelligence [1,2], biology [3], marketing [4] and so on. In general, clustering algorithms are categorized into five classes: partition-based clustering [5], hierarchical clustering [6], fuzzy clustering [7], density-based clustering [8], and model-based clustering [9,10], in which, Subspace Clustering(SC), for its validity in processing high-dimensional data, has received much public interest as a model-based clustering method. Out of different SC methods, SSC [11] and LRR [12] have made significant contributions to the growth of subspace clustering. Specifically, SSC exploits the sparse representation of data points, promoting data points represented by a linear combination of other points from the same subspace; LRR can effectively recover the subspace representation of corrupted data. Furthermore, many other variants [13,14] of SSC and LRR have also achieved superior performance. However, all these methods are intended for single-feature data and cannot handle multi-feature data.

To address above issuses, Multi-view subspace clustering methods are presented. The general MVSC methods are crudely classified as three steps. The first step is specific subspace representation learning, as with single-view subspace clustering algorithms, this step produces a coefficient matrix in each view. Two ways—self-expressiveness property [15–18] and nonnegative matrix decomposition [19] are generally adopted. The second step is multi-view correlation exploitation. After acquiring the specific representation matrices, inevitably, the shared matrix is required by merging them together. And to improve the quality of the unified representation matrix, some strategies like concatenating [20], center [21,22] and pairwise-based [17,23] regularization, and tensor singular value decomposition (t-svd) [24,25] are proposed. The last step is Spectral clustering. By inputting the shared representation matrix into the spectral clustering can obtain the final clustering results.

Although satisfactory progress has been achieved with the aforementioned methods, still there is space for enhancement. For instance, 1) Low-rank and sparsity constraints have been proved advantageous, yet few methods consider them simultaneously, and of those that do, only the relationship between two views is explored, ignoring higher-order associations. 2) Most MVSC algorithms treat the shared representation matrix as an optimization objective, which means the indicator matrix obtained from spectral clustering is not involved in the optimization process, resulting in the inability to obtain a globally optimal solution. To tackle these two problems, a new MVSC algorithm with joint tensor representation and indicator matrix learning (MVSCTI) is proposed. To illustrate our contributions more clearly, we list them as follows.

- Unlike available MVSC methods based on low-rank sparse representations that only consider the pair-wise connections of views, t-svd based TNN learning is conducted to dig the higher-order connections of multiple views.
- MVSCTI jointly pursues tensor representation and indicator matrix, which can get the indicator matrix in one step and obtain the global optimal solution.

– The effectiveness of the proposed algorithm is demonstrated by comparing it with eight advanced algorithms on five datasets.

The rest of this paper is structured as follows. The second section shows the work most relevant to this paper; The third section describes the construction of MVSCTI; For the fourth section, the superiority of MVSCTI is verified by the experimental comparison of several comparative algorithms on different datasets; The fifth section concludes the work of this paper.

2 Related Works

This section lists the related works that is most relevant to this paper.

2.1 Spectral Clustering

Spectral clustering is used extensively for the ability to deal with complex structural data and does not require any assumptions about the shape of the data. It suggests using the adjacency matrix's eigenvectors to determine the classification. The general procedure is first to generate an affinity matrix and input it into the following model to obtain the indicator matrix.

$$\min_{\mathbf{F}} Tr(\mathbf{F}^T\mathbf{L}\mathbf{F}) \quad s.t.\ \ \mathbf{F}^T\mathbf{F} = \mathbf{I} \tag{1}$$

where $\mathbf{L} = \mathbf{D} - \mathbf{W}$, $\mathbf{F}$ is the indicator matrix. Each row of $\mathbf{F}$ is considered as a point, and these points are divided into groups to which they belong, employing existing algorithms such as k-means.

2.2 Multi-view Subspace Clustering

To take best benefit of multiple views information for clustering, many MVSC methods were proposed, and their general form could be elaborated as Eq. (2).

$$\min_{\mathbf{Z}^{(v)}} \sum_{v=1}^{V} \Psi\left(\mathbf{Z}^{(v)}\right) \quad s.t. \quad \mathbf{X}^{(v)} = \mathbf{X}^{(v)}\mathbf{Z}^{(v)},\ \ diag\left(\mathbf{Z}^{(v)}\right) = 0 \quad (v = 1, 2, 3...V) \tag{2}$$

where $\mathbf{X}^{(v)}$ represents the feature matrix of the v-th view, $\Psi\left(\mathbf{Z}^{(v)}\right)$ represents the regularization term to induce the desired performance of the coefficient matrix $\mathbf{Z}^{(v)}$.

3 Model Proposal and Optimization

In this section, a MVSC method with joint tensor representation and indicator matrix learning (MVSCTI) is designed. The flowchart of MVSCTI is shown in Fig. 1.

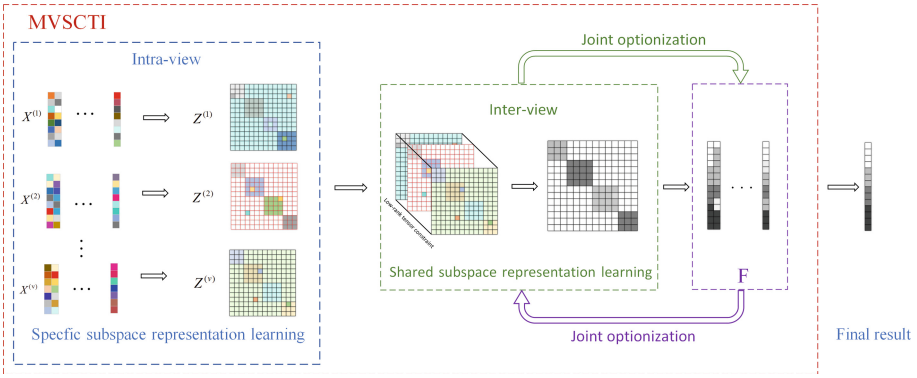

Fig. 1. Input multiple original data matrices $\mathbf{X}^{(v)}$. Then, specific subspace learning is performed using self-expressiveness property and l_1 norm constraints to obtain multiple coefficient matrices $Z^{(v)}$. Further, stack these matrices as tensor and apply low-rank constraints from lateral; after that, integrate them to a shared subspace representation matrix. Next, input it into the spectral clustering algorithm for the indicator matrix F, which is also the final optimization objective of our model. Eventually, the k-means algorithm is adopted for achieving the final results.

3.1 The Proposed Model

In consideration of the efficiency of sparse representation, a MVSC framework based on sparse representation is first proposed, as shown in Eq. (3).

$$\begin{aligned} &\min_{\mathbf{A}^{(v)}} \underbrace{\sum_{\mathrm{v}=1}^{n_v} \frac{1}{2}\left\|\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{A}^{(v)}\right\|_F+\alpha\left\|\mathbf{A}^{(v)}\right\|_{\text{sparsity}}}_{\text{intra-views}}+\underbrace{\sum_{v\neq w, v\geqslant 1, w\leqslant n_v} corre(\mathbf{A}^{(v)},\mathbf{A}^{(w)})}_{\text{inter-views}} \\ &s.t.\ \ diag\left(\mathbf{A}^{(v)}\right)=0 \end{aligned} \tag{3}$$

where sparsity represents sparsity constraint and *corre* is abbreviations for correlation, which represents inter-view correlation. The formulation is divided into two terms, the first term—intra-views performs specific subspace learning within each view and the other term mines the inter-view connections. For sparsity constraints, there are many different choices, the most efficient of which is the schatten-p norm, which represents the number of non-zero elements, but it is non-convex, so we choose its convex approximate l_1 norm here, and use the TNN to dig the higher-order connections of multiple views, and the formula is shown in Eq. (4).

$$\min_{\mathbf{A}^{(v)}} \underbrace{\sum_{\mathrm{v}=1}^{n_v} \frac{1}{2}\left\|\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{A}^{(v)}\right\|_F+\alpha_2\left\|\mathbf{A}^{(v)}\right\|_1}_{\text{intra-views}}+\underbrace{\sum_{v\neq w, v\geqslant 1, w\leqslant n_v} \alpha_1\|\mathcal{A}\|_{\circledast}}_{\text{inter - views}} \quad (4)$$
$$s.t.\ \ diag\left(\mathbf{A}^{(v)}\right)=0$$

where $\|\mathcal{A}\|_{\circledast}$ is t-svd based TNN [24]. The above formula takes $\mathbf{A}^{(v)}$ as the final optimization objective, then, the shared similarity matrix $\mathbf{A}$ is obtained using $\sum_{v=1}^{n_v} \frac{\mathbf{A}^{(v)T}+\mathbf{A}^{(v)}}{2n_v}$, after which it is input into the spectral clustering algorithm for the indicator matrix, which is a two-step method, and the final results are not incorporated into our optimization model, implying that what is obtained in this way is not a globally optimal solution. So we further integrate the spectral clustering $Tr(\mathbf{F}^T\mathbf{L}\mathbf{F})$ into our model, and the final model is as shown in Eq. (5).

$$\min_{\mathbf{A}^{(v)},\mathbf{F}} \sum_{v=1}^{n_v} \overbrace{(\underbrace{\frac{1}{2}\left\|\mathbf{X}^{(v)}-\mathbf{X}^{(v)}\mathbf{A}^{(v)}\right\|_F}_{\text{Self-expressiveness}}+\underbrace{\alpha_2\left\|\mathbf{A}^{(v)}\right\|_1}_{\text{Sparsity constrain}})}^{\text{Specific subspace representation learning}}+\overbrace{(\underbrace{\alpha_1\|\mathcal{A}\|_{\circledast}}_{\text{Low-rank tensor constrain}}+\underbrace{\lambda \operatorname{Tr}\left(\mathbf{F}^T\mathbf{L}\mathbf{F}\right)}_{\text{Spectral clustering}})}^{\text{Shared subspace representation learning}}$$
$$s.t. \quad \operatorname{diag}\left(\mathbf{A}^{(v)}\right)=\mathbf{0}, \quad \mathbf{F}^T\mathbf{F}=\mathbf{I} \quad (5)$$

where $\mathbf{L}=\mathbf{D}-\mathbf{W}$, $\mathcal{A}=\Phi\left(\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(n_v)}\right)$, $\mathbf{W}=\sum_{v=1}^{n_v}\frac{\mathbf{A}^{(v)T}+\mathbf{A}^{(v)}}{2n_v}$, $\mathbf{F}^T\mathbf{F}=\mathbf{I}$ is set to prevent a trivial solution, $\mathbf{D}=diag\left(\sum_j[\mathbf{W}]_{1j},\sum_j[\mathbf{W}]_{2j},\cdots,\sum_j[\mathbf{W}]_{nj}\right)$. This formulation is a model end-to-end, with which the global optimal solution can be obtained. The proposed model can be optimized with an efficient algorithm Alternating Direction Method of Multipliers (ADMM), and the specific optimization process and the complexity analysis is provided in the Appendix.

4 Experiment and Analysis

This section verifies the validity and superiority of MVSCTI through comparative experiments. Furthermore, the factors of excellence of the model are analyzed by conducting ablation study, parameter sensitivity and convergence analysis experiments.

4.1 Experiment Setting

Datasets. Experiments are conducted on five datasets covering different types and domains of data, such as articles, images, and biology. More details are as follows.

- **BBCsports** [26]: This dataset contains 554 articles owned by the BBC, containing five topics and two views.
- **NUS** [21]: This dataset is from NUS-WIDE. It includes 12 categories of images and 6 views, of which we have selected 2400 images.
- **Prokaryotic** [27]: This dataset contains 551 prokaryotic species with four classes and three views (textual, the proteome composition, and the gene repertoire).
- **WebKB** [22]: This dataset includes 203 web-pages with 4 categories and three views.
- **Reuters** [28]: The archive contains 1200 documents over the six labels. It comprises five views on the same documents.

Table 1 enumerates the general information of the datasets.

Table 1. Details of the five datasets.

Datasets	Views	Dimensionality	Instances	Classes
BBCsports	2	3183/3203	554	5
NUS	6	64/144/73/128/255/500	2400	12
Prokaryotic	3	438/3/393	551	4
WebKB	3	1703/230/230	203	4
Reuters	5	2000/2000/2000/2000/2000	1200	6

Baselines and Metrics. We compare the algorithm in this paper with eight advanced algorithms, which includes GBS-KO [22], SMVSC [21], LTMSC [29], Co-reg [30], DIMSC [31], FPMVS-CAG [32], CoMSC [28] and MLRSSC [27]. Three commonly used evaluation metrics are adopted for our experiment, i.e., Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI). All comparison algorithms codes were downloaded from the URL provided by the authors. We choose the parameters provided in the original article that correspond to the optimal performance. To present a stable experimental result, we execute 20 times for every method, take the average result and its standard deviation is calculated. All experiments were run on the identical device with an Intel(R) Core(TM) i5-7400 3.00 GHz CPU.

4.2 Experimental Results and Analysis

Based on the results of the comparison experiments displayed in Table 2 as well as in Fig. 2, where - indicate that the algorithm is not applicable to the corresponding dataset. We present the following observations:

- Compared to all contrast algorithms, our method is optimal all the time, e.g., TIMVSC outperforms the second best algorithm MLRSSC by 14.4%, 18.4%, and 21.8% in ACC, NMI, and ARI, respectively, on the prokaryotic dataset.

In addition, on the Reuters dataset, TIMVSC outperforms the second best algorithm MLRSSC exceeds 27.6%, 38.4%, and 40.7% in ACC, NMI, and ARI, respectively. These show the superiority of TIMVSC.
- Figure 2 shows a two-dimensional visualization of the embedding representation F using the T-SNE algorithm. It can be observed that MVSCTI exhibits a clearer structure than most other algorithms.
- Compared with MLRSSC, our algorithm always outperforms it, while the most significant difference between them is the way of low-rank constraint, MLRSSC uses matrix svd decomposition, while MVSCTI employs t-svd decomposition, which illustrates that for multi-view clustering, adopting t-svd decomposition for low-rank constraint is able to effectively exploit the higher-order associations between views.

4.3 Model Discussion

We validate our model by performing ablation study, parameter sensitivity, run time and convergence experiments.

Ablation Study. One of the innovations of the proposed model is the joint low-rank tensor representation learning and indicator matrix learning. To further verify the advantages of this joint learning strategy, the following two experiments were designed, in which the first one eliminates the spectral clustering term (indicator matrix learning), and the second one eliminates the low-rank tensor constraint term. The experimental results are shown in Fig. 3, where No SC is eliminating spectral clustering term and No tensor is eliminating tensor learning term. Both experiments' results present a decrease compared to MVSCTI and the decrease is more significant after eliminating the tensor learning term, which illustrates the effectiveness of our joint learning strategies, and incorporating spectral clustering into the model not only reduces the clustering steps, but also improves the clustering effectiveness.

Parameter Sensitivity. For our proposed model, there are three parameters to be traded off: α_1, α_2, and λ, corresponding to the tensor term, the sparsity term, and the spectral clustering term, respectively. Figure 5 shows the parameter sensitivity experiments on NUS, where we did a logarithmic treatment of the parameter coordinate label values. One can see that the experiments can get a good results when selecting parameters in a certain range. Here we provide a parameter selection range. For all datasets, α_2 and λ can be chosen from a range of [0.01, 0.1, 1, 10] except the BBCsports dataset, whose α_2 can be chosen from a range of [0.01, 1000, 10000] and λ can be chosen from a range of [0.1, 100, 1000]; for Prokaryotic, α_1 can be chosen from a range of [0.01, 1, 10, 1000], and for the rest of datasets, α_1 can be chosen from a range of [1000, 10000, 100000].

Table 2. The experimental metrics of nine algorithms on five datasets.

Method	Metrics	BBCsports	NUS	Prokaryotic	WebKBs	Reuters
Co-reg	ACC	0.356(0.003)	0.008(0.008)	0.537(0.005)	0.613(0.016)	0.018(0.018)
	ARI	0.004(0.001)	0.141(0.003)	0.045(0.005)	0.295(0.011)	0.866(0.129)
	NMI	0.021(0.004)	0.127(0.002)	0.111(0.004)	0.258(0.005)	0.168(0.001)
LT-MSC	ACC	0.460(0.046)	0.241(0.010)	0.419(0.005)	0.538(0.003)	0.418(0.031)
	ARI	0.166(0.042)	0.012(0.000)	0.0281(0.000)	0.180(0.003)	0.151(0.018)
	NMI	0.221(0.027)	0.124(0.007)	0.130(0.006)	0.165(0.001)	0.212(0.011)
DIMSC	ACC	0.795(0.003)	0.127(0.003)	0.362(0.002)	–	–
	ARI	0.563(0.005)	0.006(0.000)	0.027(0.002)	–	–
	NMI	0.583(0.004)	0.023(0.001)	0.031(0.000)	–	–
MLRSSC	ACC	0.840(0.013)	0.294(0.008)	0.654(0.007)	0.698(0.007)	0.475(0.024)
	ARI	0.770(0.015)	0.091(0.003)	0.339(0.009)	0.476(0.013)	0.204(0.015)
	NMI	0.762(0.009)	0.157(0.003)	0.319(0.003)	0.451(0.116)	0.283(0.010)
S-MVSC	ACC	0.789(0.111)	0.295(0.009)	0.411(0.103)	0.700(0.255)	0.313(0.036)
	ARI	0.698(0.132)	0.097(0.005)	0.061(0.011)	0.501(0.021)	0.067(0.034)
	NMI	0.712(0.093)	0.164(0.004)	0.155(0.003)	0.463(0.011)	0.159(0.023)
GBS-KO	ACC	0.807(0.000)	0.165(0.000)	0.510(0.000)	0.744(0.000)	0.199(0.000)
	ARI	0.722(0.000)	0.012(0.000)	0.102(0.000)	0.368(0.000)	0.013(0.000)
	NMI	0.760(0.000)	0.122(0.000)	0.217(0.000)	0.378(0.000)	0.132(0.000)
FPMVS-CAG	ACC	0.423(0.000)	0.258(0.001)	0.523(0.000)	0.576(0.000)	0.443(0.000)
	ARI	0.132(0.000)	0.012(0.000)	0.135(0.000)	0.326(0.000)	0.169(0.000)
	NMI	0.151(0.000)	0.124(0.007)	0.154(0.000)	0.327(0.000)	0.212(0.000)
CoMSC	ACC	0.850(0.067)	0.206(0.006)	0.579(0.000)	0.735(0.030)	0.541(0.026)
	ARI	0.683(0.053)	0.041(0.003)	0.017(0.000)	0.517(0.040)	0.517(0.040)
	NMI	0.681(0.038)	0.086(0.002)	0.05(0.000)	0.492(0.022)	0.353(0.017)
MVSCTI	**ACC**	**0.939(0.047)**	**0.345(0.013)**	**0.798(0.010)**	**0.865(0.052)**	**0.751(0.006)**
	ARI	**0.890(0.042)**	**0.270(0.007)**	**0.558(0.009)**	**0.730(0.057)**	**0.611(0.012)**
	NMI	**0.890(0.025)**	**0.330(0.028)**	**0.504(0.008)**	**0.695(0.032)**	**0.667(0.010)**

Table 3. Run time of all MVC comparative algorithms on all datasets (in seconds).

Method	Co-reg	LT-MSC	DIMSC	MLRSSC	S-MVSC	GBS-KO	FPMVS-CA	CoMSC	MVSCTI
BBCsports	8.31	43.76	10.87	3.21	41.35	13.01	12.60	1.65	2.24
NUS	155.83	1939.20	1840.83	913.65	2.31	33.76	174.49	11.11	507.68
Prokaryotic	3.50	26.96	8.86	4.23	0.29	–	5.09	3.75	3.27
WebKBs	1.63	5.67	–	0.80	0.18	0.27	3.05	0.51	0.70
Reuter	25.37	493.25	–	96.55	359.40	98.23	32.99	7.76	67.47

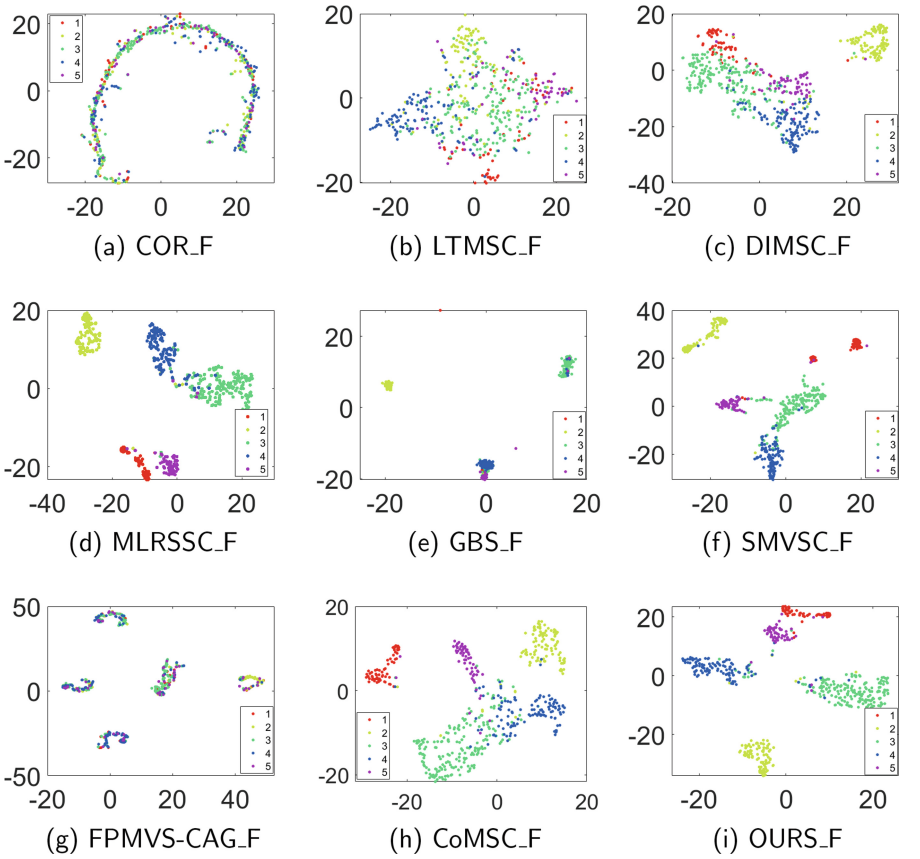

Fig. 2. The visualization of the embedding representation F of different MVC methods on BBCsports.

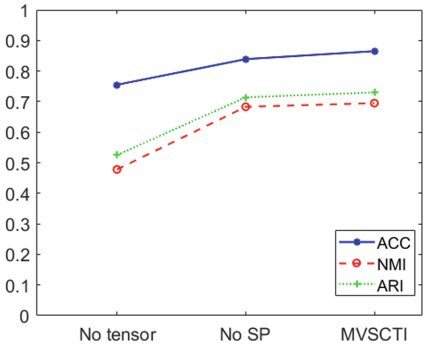

Fig. 3. Ablation study on WebKBs.

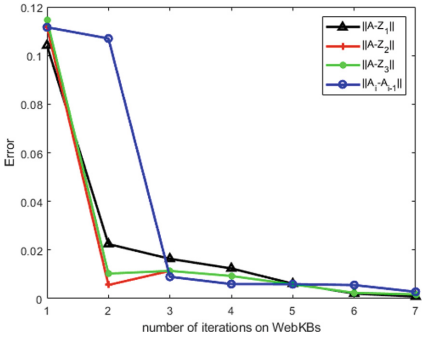

Fig. 4. The convergence curve on WebKBs.

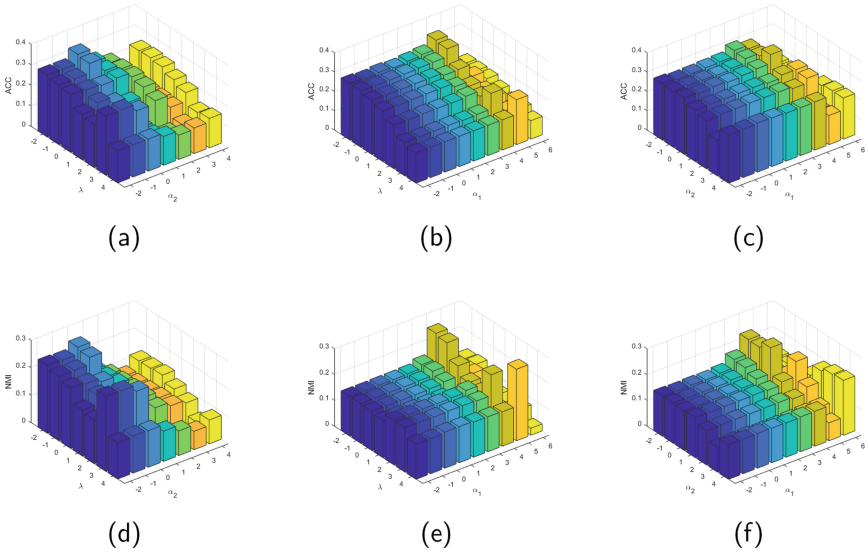

Fig. 5. Influence of parameter variation (The 1st column is about α_2 & λ with α_1 fixed as 10000; The 2nd column is about α_1 & λ with α_2 fixed as 1; The 3rd column is about α_1 & α_2 with λ fixed as 0.01) for ACC and NMI on the NUS dataset.

Run Time and Convergence. We calculated the average run time of all algorithms, and the experimental results are shown in Table 3. MVSCTI has the least run time on the BBCsports dataset in comparison with other algorithms and is in the middle on the other datasets. Furthermore, for most algorithms, it can be observed that the larger the size of the dataset, the longer the run time, except for S-MVSC, whose run time on the BBCsports dataset with a size of 554 is 41.35 s, which is the second-longest among all algorithms. Still, its run time is only 2.31 s on the NUS dataset with a size of 2400, which is much smaller than the other algorithms. This might be because it is more sensitive to data dimensionality than data size. For example, the size of NUS is twice as large as Reuter, but its dimensionality is much smaller than Reuter, which leads to its run time on Reuter (359.40 s) to be much higher than that of NUS (2.31 s). We also conduct a convergence experiment on WebKBs. As shown in Fig. 4, it can reach convergence within 7 times.

5 Conclusions

In this paper, an effective method—MVSCTI is proposed. It exploits the correlations of multiple views by low-rank tensor learning, and the global optimal solution can be gained by combining spectral clustering within the MVSC framework. In the end, we compare MVSCTI with ten advanced benchmark algorithms by experimenting on seven public datasets, and the results show the superiority

of our method. Further, we validate the proposed model by conducting several general experiments. For future works, We believe the following issues remain to be explored: 1) Inter-view consistency and complementary must be exploited more effectively, which is essential for improving clustering effect; 2) Reducing the computational complexity while improving clustering performance, which is critical for large-scale data.

References

1. Cui, Z., Jing, X., Zhao, P., Zhang, W., Chen, J.: A new subspace clustering strategy for ai-based data analysis in iot system. IEEE Internet Things J. (2021)
2. He, R., et al.: A kernel-power-density-based algorithm for channel multipath components clustering. IEEE Trans. Wireless Commun. **16**(11), 7138–7151 (2017)
3. Chowdhury, H.A., Bhattacharyya, D.K., Kalita, J.K.: Uicpc: centrality-based clustering for scrna-seq data analysis without user input. Comput. Biol. Med. **137**, 104820 (2021)
4. Cheng, Y., Cheng, M., Pang, T., Liu, S.: Using clustering analysis and association rule technology in cross-marketing. Complexity **2021**, 1–11 (2021)
5. Mittal, M., Sharma, R.K., Singh, V.P., Kumar, R.: Adaptive threshold based clustering: a deterministic partitioning approach. Int. J. Inf. Syst. Modeling Des. (IJISMD) **10**(1), 42–59 (2019)
6. Murtagh, F., Contreras, P.: Algorithms for hierarchical clustering: an overview. Wiley Interdisciplinary Rev. Data Mining Knowl. Discovery **2**(1), 86–97 (2012). https://doi.org/10.1002/widm.53
7. Bagherinia, A., Minaei-Bidgoli, B., Hosseinzadeh, M., Parvin, H.: Reliability-based fuzzy clustering ensemble. Fuzzy Sets Syst. **413**, 1–28 (2021)
8. Campello, R.J., Kröger, P., Sander, J., Zimek, A.: Density-based clustering. Wiley Interdisc. Rev. Data Mining Knowl. Dis. **10**(2), e1343 (2020)
9. Bouveyron, C., Brunet-Saumard, C.: Model-based clustering of high-dimensional data: a review. Comput. Stat. Data Anal. **71**, 52–78 (2014) https://doi.org/10.1016/j.csda.2012.12.008,https://www.sciencedirect.com/science/article/pii/S0167947312004422
10. Xue, X., Zhang, X., Feng, X., Sun, H., Chen, W., Liu, Z.: Robust subspace clustering based on non-convex low-rank approximation and adaptive kernel. Inf. Sci. **513**, 190–205 (2020)
11. Elhamifar, E., Vidal, R.: Sparse subspace clustering: algorithm, theory, and applications. IEEE Trans. Pattern Anal. Mach. Intell. **35**(11), 2765–2781 (2013)
12. Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., Ma, Y.: Robust recovery of subspace structures by low-rank representation. IEEE Trans. Pattern Anal. Mach. Intell. **35**(1), 171–184 (2012)
13. Kumar, S., Dai, Y., Li, H.: Multi-body non-rigid structure-from-motion. In: 2016 Fourth International Conference on 3D Vision (3DV), pp. 148–156. IEEE (2016)
14. Tierney, S., Gao, J., Guo, Y.: Subspace clustering for sequential data. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1019–1026 (2014)
15. Kang, Z., et al.: Partition level multiview subspace clustering. Neural Netw. **122**, 279–288 (2020)
16. Yang, Z., Xu, Q., Zhang, W., Cao, X., Huang, Q.: Split multiplicative multi-view subspace clustering. IEEE Trans. Image Process. **28**(10), 5147–5160 (2019)

17. Yu, H., Zhang, T., Lian, Y., Cai, Y.: Co-regularized multi-view subspace clustering. In: Asian Conference on Machine Learning, pp. 17–32. PMLR (2018)
18. Zhang, X., Sun, H., Liu, Z., Ren, Z., Cui, Q., Li, Y.: Robust low-rank kernel multi-view subspace clustering based on the schatten p-norm and correntropy. Inf. Sci. **477**, 430–447 (2019)
19. Tolić, D., Antulov-Fantulin, N., Kopriva, I.: A nonlinear orthogonal non-negative matrix factorization approach to subspace clustering. Pattern Recogn. **82**, 40–55 (2018)
20. Zheng, Q., Zhu, J., Li, Z., Pang, S., Wang, J., Li, Y.: Feature concatenation multi-view subspace clustering. Neurocomputing **379**, 89–102 (2020)
21. Hu, Z., Nie, F., Chang, W., Hao, S., Wang, R., Li, X.: Multi-view spectral clustering via sparse graph learning. Neurocomputing **384**, 1–10 (2020)
22. Wang, H., Yang, Y., Liu, B., Fujita, H.: A study of graph-based system for multi-view clustering. Knowl.-Based Syst. **163**, 1009–1019 (2019)
23. Tang, C., Zhu, X., Liu, X., Li, M., Wang, P., Zhang, C., Wang, L.: Learning a joint affinity graph for multiview subspace clustering. IEEE Trans. Multimed. **21**(7), 1724–1736 (2018)
24. Xie, Y., Tao, D., Zhang, W., Liu, Y., Zhang, L., Qu, Y.: On unifying multi-view self-representations for clustering by tensor multi-rank minimization. Int. J. Comput. Vision **126**(11), 1157–1179 (2018)
25. Wu, J., Lin, Z., Zha, H.: Essential tensor learning for multi-view spectral clustering. IEEE Trans. Image Process. **28**(12), 5910–5922 (2019)
26. Huang, S., Xu, Z., Lv, J.: Adaptive local structure learning for document co-clustering. Knowl.-Based Syst. **148**, 74–84 (2018)
27. Brbić, M., Kopriva, I.: Multi-view low-rank sparse subspace clustering. Pattern Recogn. **73**, 247–258 (2018)
28. Liu, J., Liu, X., Yang, Y., Guo, X., Kloft, M., He, L.: Multiview subspace clustering via co-training robust data representation. IEEE Trans. Neural Networks Learn. Syst. (2021)
29. Zhang, C., Fu, H., Liu, S., Liu, G., Cao, X.: Low-rank tensor constrained multi-view subspace clustering. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1582–1590 (2015)
30. Kumar, A., Rai, P., Daume, H.: Co-regularized multi-view spectral clustering. Adv. Neural. Inf. Process. Syst. **24**, 1413–1421 (2011)
31. Cao, X., Zhang, C., Fu, H., Liu, S., Zhang, H.: Diversity-induced multi-view subspace clustering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586–594 (2015)
32. Wang, S., et al.: Fast parameter-free multi-view subspace clustering with consensus anchor guidance. IEEE Trans. Image Process. **31**, 556–568 (2022). https://doi.org/10.1109/TIP.2021.3131941