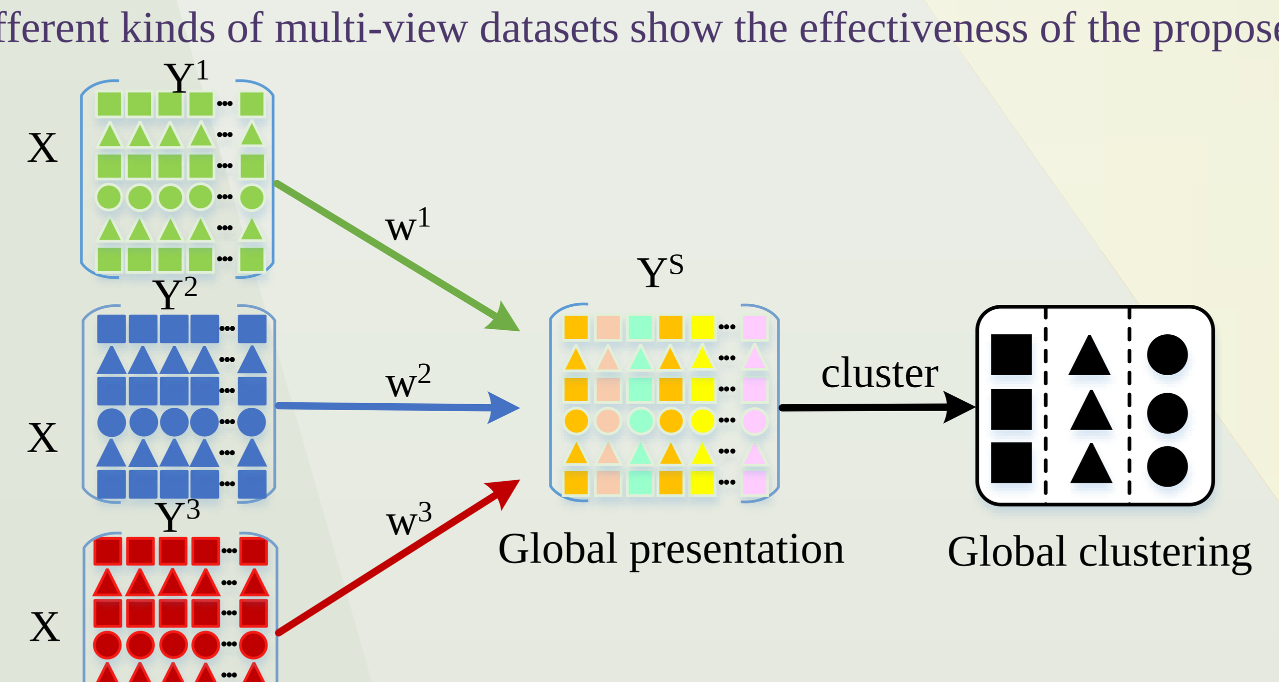
A Parameter-free Multi-view Information Bottleneck Clustering Method by Cross-view Weighting

Shizhe Hu, Ruilin Geng, Zhaoxu Cheng, Chaoyang Zhang, Guoliang Zou, Zhengzheng Lou, Yangdong Ye*

School of Computer and Artificial Intelligence Zhengzhou University, Zhengzhou China

ABSTRACT

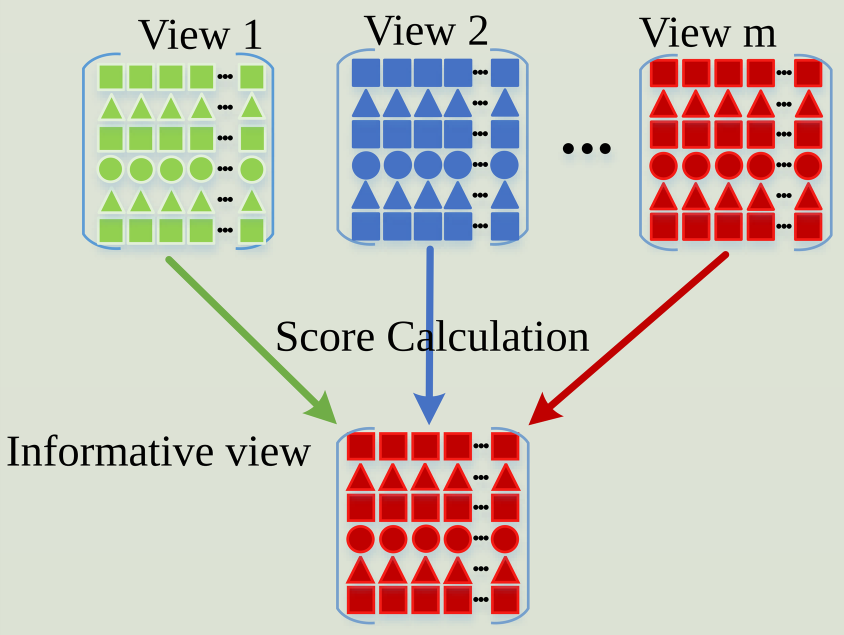
With the fast-growing multi-modal/media data in the Big Data era, multi-view clustering (MVC) has attracted lots of attentions lately. Most MVCs focus on integrating and utilizing the complementary information among views by linear sum of the learned view weights and have shown great success in some fields. However, they fail to quantify how complementary the information across views actually utilized for benefiting final clustering. Additionally, most of them contain at least one parameter for regularization without prior knowledge, which puts pressure on the parameter tuning and thus makes them impractical. In this paper, we propose a novel parameter-free multi-view information bottleneck (PMIB) clustering method to automatically identify and exploit useful complementary information among views, thus reducing the negative impact from the harmful views. Specifically, we first discover the informative view by measuring the relevant information preserved by the original data and the compact clusters with mutual information. Then, a new cross-view weight learning scheme is designed to learn how complementary between the informative view and remaining views. Finally, the quantitative correlations among views are fully exploited to improve the clustering performance without needing any additional parameters or prior knowledge. Experimental results on different kinds of multi-view datasets show the effectiveness of the proposed method.



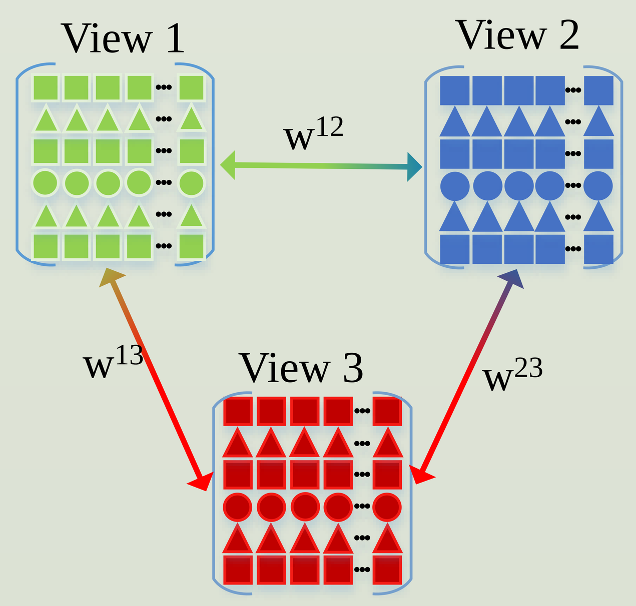
Traditional study

METHODS

Informative View Discovery. We discover the informative view by measuring the relevant information preserved by the original data and the compact clusters with the concept of MI. On one hand, for the i-th view, we calculate and measure the relevant information between the data variable X and its feature variables Y^i via MI. The higher the MI value obtains, the more informative or relevant the i-view is. On the other hand, we perform the IB method on each view to obtain the data assignments, by which we compute the MI between it and the feature variable Y^i. The higher the MI value obtains, the bigger contribution the current view makes, which thus makes it more informative. Based on the above, we obtain the following score calculation for informative view discovery:



Informative View Discovery



Cross-View Weighting

$$\text{Score}=\frac{I(X;Y^i)+I(T^i;Y^i)}{\sum_{i=1}^{m} I(X;Y^i)+I(T^i;Y^i)}$$

$$\lambda^{ij}=\frac{I(V_{row}^{ij};V_{column}^{ij})+I(V_{row}^{ji};V_{column}^{ji})}{I(X;Y^i)+I(X;Y^j)},\ w^{ij}=w^{ji}=\frac{\lambda^{ij}}{\sum \lambda^{ij}}$$

Cross-view Weights Learning. Most existing MVCs utilize the view-weighted scheme to learn a weight for every view by measuring the importance of views in order to employ the complementary information. However, they only focus on each individual view and then combine views together by linear weighted sum, failing to quantify how complementary the information across views is actually utilized for benefiting the data clustering. Unlike existing view weighted schemes only considering the importance of individual view, cross-view weight takes into account the importance of pairwise views by quantifying how complementary across views.

Algorithm 1 The Proposed PMIB method

Input: m joint distributions $\{p(X,Y^i)\}_i^{m=1}$, cluster number c.
Output: Final data assignments $p(t|x)$.
Initialization: $T \leftarrow$ Random partition of X into c clusters;
Discover the informative view among m views via informative view strategy
Compute the cross-view weights between the informative view and remaining views via cross-view weighting scheme.
repeat
  for all $x \in$ X do
    Draw: Draw x from the original cluster t_{old};
    Merger: Assign x into the clusters of m view, and choose a new cluster t_{new} with the minimal merger cost by optimization operation. Then, x is merged into the new chosen cluster;
  end for
until Data samples of clusters in views remain unchanged.

Objective function: $F_{\max}[p(t|x)] = S^a \cdot I(T;Y^a) + \sum_{i=1, i\neq a}^{m} w^{ai} \cdot I(T;Y^i)$

The advantages of PMIB:

- Usage of quantitative complementary information
- Cross-view weighting
- No parameters are involved

EXPERIMENTS

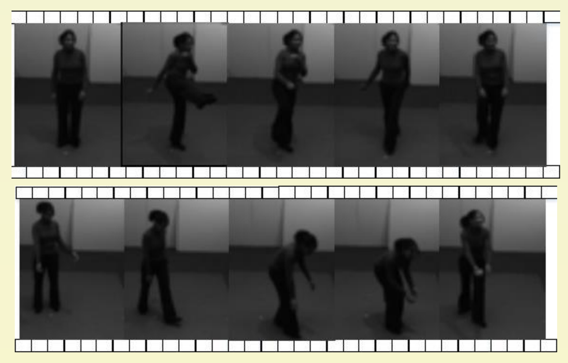
Methods	Parameters
MVIB	β: {1,1.5,2,4,6,8}
Co(reg)	λ:[0.01,0.05] (step=0.01)
RMKMC	γ:[0.1,2] (step=0.2)
MfIB	$\{\lambda_i\}^m{}_i$=1:[0.1,0.9] (step=0.1)
RMSC	λ:{0.005,0.01,0.05,0.1,0.5}
DEKM	γ:[1,10] (step=1)
MLAN	λ:[1,30] (step=λ/2 or $\lambda * 2$)
GMC	$\lambda \geq 1$: (step=λ/2 or $\lambda * 2$)
ONMSC	α:[2^0, 2^3, . . . , 2^{15}]
SMVSC	*anchors* = {c, $2c$, $3c$}
Ours	Parameter-free

Parameter settings of the compared multi-view clustering methods and ours

(a) COIL100

Dataset	Type	View	Samples	Clusters	Dimensionality
20NGs	Text	3	500	5	2000
COIL20	Image	3	1440	20	1000
COIL100	Image	2	7200	100	800
IXMAS	Video	4	330	11	1000
Large Dataset	Type	View	Samples	Clusters	Dimensionality
GHIM	Image	2	10000	20	500
ALOI	Image	3	11025	100	(77, 64, 64)
CIFAR	Image	2	15000	10	1000

Details of the involved multi-view datasets



(b) IXMAS

Some datasets examples

Method	20NGs		COIL20		COIL100		IXMAS	
	Acc	NMI	Acc	NMI	Acc	NMI	Acc	NMI
KM	23.68±0.39	6.41±0.64	52.05±2.25	64.59±1.22	30.17±1.94	62.31±1.91	14.33±2.40	11.56±3.59
IB	85.48±10.57	76.07±7.68	68.43±3.89	78.44±2.61	40.76±1.43	68.36±0.78	43.63±4.54	41.41±3.70
AVKM	21.12±0.53	1.49±0.45	44.56±6.84	59.30±7.40	30.53±3.16	51.10±2.93	12.03±2.15	4.80±3.28
AVIB	91.80±9.88	92.72±6.71	77.52±5.99	91.42±2.55	37.45±2.29	67.68±2.45	26.82±8.98	48.06±8.24
MVIB (DASFAA'07)	94.22±1.37	83.21±3.18	62.19±10.50	73.54±6.81	46.35±1.68	70.21±0.67	59.03±3.92	57.23±3.06
Co(reg) (NeurIPS'11)	20.35±0.83	3.87±0.73	64.75±1.28	84.13±0.37	48.68±0.47	71.19±0.14	17.26±1.20	31.06±1.14
RMKMC (IJCAI'13)	48.20±3.43	36.86±4.12	53.06±3.82	72.02±2.99	45.01±1.61	68.79±0.68	34.27±4.34	36.20±3.49
MfIB (IJCAI'13)	93.76±2.89	85.11±4.54	78.99±3.88	88.54±1.27	56.52±0.08	80.81±0.46	64.85±6.42	67.48±3.21
RMSC (AAAI'14)	37.26±0.91	15.70±0.84	65.45±4.26	79.13±1.89	46.32±0.28	69.33±0.45	52.61±2.76	68.79±1.61
DEKM (CVPR'16)	31.60±0.00	13.48±0.00	40.69±0.00	59.52±0.00	42.76±0.00	69.57±0.00	17.27±0.00	9.75±0.00
MLAN (TIP'18)	96.40±0.11	89.18±0.17	**86.91±0.40**	**94.49±0.00**	45.05±0.41	59.55±0.53	68.15±2.57	72.23±1.71
GMC (TKDE'20)	98.20±0.00	93.92±0.00	73.33±0.00	92.11±0.00	38.86±0.00	67.55±0.00	67.58±0.00	**78.51±0.00**
ONMSC (AAAI'20)	97.20±0.00	91.48±0.00	13.06±0.00	11.88±0.00	6.47±0.00	18.47±0.00	65.76±0.00	**77.67±0.00**
SMVSC (ACM MM'21)	63.70±5.53	50.69±5.18	66.15±3.93	82.42±1.16	42.13±0.91	67.42±0.73	36.06±2.15	46.61±0.00
PMIB (Parameter-free)	**99.18±0.24**	**97.28±0.66**	**85.83±3.98**	**93.60±1.72**	**62.01±1.87**	**84.59±0.88**	**73.67±9.51**	**76.85±6.15**

Clustering Results (+/-Standard Deviation) on the Four Multi-view Datasets

CONCLUSION

In this paper, a novel parameter-free multi-view information bottleneck (PMIB) method is proposed to improve the multi-view clustering performance. Unlike existing MVCs utilizing complementary information by linear sum of view weights, PMIB can fully exploit the quantitative complementary information by cross-view weights. Instead of learning view-based weights, a novel cross-view weight learning scheme is designed to accurately quantify the complementary information between views. Additionally, our PMIB is totally parameter-free without involving any parameters or prior knowledge, thus being more practical. Experimental results on different kinds of multi-view datasets show the effectiveness of the proposed method. In the future, we plan to extend our method to address more complicated incomplete multi-view clustering for broad applications.

REFERENCES

[1] Anton Bardera, Jaume Rigau, Imma Boada, Miquel Feixas, and Mateu Sbert. 2009. Image Segmentation Using Information Bottleneck Method. IEEE Trans. Image Process. 18, 7 (2009), 1601–1612.

[2] Herbert Bay, Tinne Tuytelaars, and Luc J. Van Gool. Graz, Austria, May 2006. SURF: Speeded Up Robust Features. In Proc. Eur. Conf. Comput. Vis. 404–417.

[3] Steffen Bickel and Tobias Scheffer. 2004. Multi-View Clustering. In Proceedings of the 4th IEEE International Conference on Data Mining (ICDM). 19–26.

[4] Shaohua Fan, Xiao Wang, Chuan Shi, Emiao Lu, Ken Lin, and Bai Wang. 2020. One2Multi Graph Autoencoder for Multi-view Graph Clustering. In WWW '20: The Web Conference. 3070–3076

ACKNOWLEDGEMENTS

This work is supported by National Natural Science Foundation of China（62176239, 61906172, 62002330）.

CONTACT

E-mail: ieshizhehu@zzu.edu.cn; healinrlgeng@outlook.com

Personal homepage: http://www5.zzu.edu.cn/mlis/info/1011/1051.htm