Supplementary Materials: Automatic and Aligned Anchor Learning Strategy for Multi-View Clustering

Anonymous Authors

1 ADDITIONAL REMARK FOR OPTIMIZATION

In order to optimize the alignment matrix $\mathbf{P}$, the problem of aligning two anchor sets is formulated as follows:

$$\begin{aligned} &\min_{\mathbf{P}} \|\mathbf{Z}_b - \mathbf{Z}_i\mathbf{P}\|_F^2 + \lambda\|\mathbf{G}_b - \mathbf{P}^\top\mathbf{G}_i\mathbf{P}\|_F^2 \\ &\Leftrightarrow \max_{\mathbf{P}} \mathrm{Tr}(\mathbf{Z}_b^\top\mathbf{Z}_i\mathbf{P} + \lambda\mathbf{G}_b\mathbf{P}^\top\mathbf{G}_i\mathbf{P}), \\ &\text{s.t. } \mathbf{P}\mathbf{1} = \mathbf{1}, \mathbf{P}^\top\mathbf{1} = \mathbf{1}, \mathbf{P} \in \{0,1\}^{m_i \times m_b}. \end{aligned} \tag{1}$$

The above problem is a quadratic assignment problem (QAP) and generally is proved to be NP-hard[1]. The feasible region constraint often relaxes into its convex hull, the Birkhoff polytope with double stochastic region. Then the optimization problem can be transformed as

$$\begin{aligned} &\max_{\mathbf{P}} \mathrm{Tr}(\mathbf{Z}_b^\top\mathbf{Z}_i\mathbf{P} + \lambda\mathbf{G}_b\mathbf{P}^\top\mathbf{G}_i\mathbf{P}), \\ &\text{s.t. } \mathbf{P}\mathbf{1} = \mathbf{1}, \mathbf{P}^\top\mathbf{1} = \mathbf{1}, \mathbf{P} \in [0,1]^{m_i \times m_b}. \end{aligned} \tag{2}$$

To solve Eq. (2), we use the Projection Fixed-Point Algorithm [2] to update $\mathbf{P}$ as follows:

$$\begin{aligned} \mathbf{P}^{(t+1)} &= (1-\alpha)\mathbf{P}^{(t)} + \alpha\Gamma\left(\nabla f\left(\mathbf{P}^{(t)}\right)\right) \\ &= (1-\alpha)\mathbf{P}^{(t)} + \alpha\Gamma\left(\mathbf{K}^\top + 2\lambda\mathbf{G}_i\mathbf{P}^{(t)}\mathbf{G}_b^\top\right), \alpha \in [0,1], \end{aligned} \tag{3}$$

where α denotes the step size parameter, t denotes the number of iterations, $\mathbf{\Gamma}$ denotes the double stochastic projection operator and $\mathbf{K} = \mathbf{Z}_b^\top\mathbf{Z}_i$. We set $\alpha = 0.5$ in this paper.

2 EXPERIMENTS

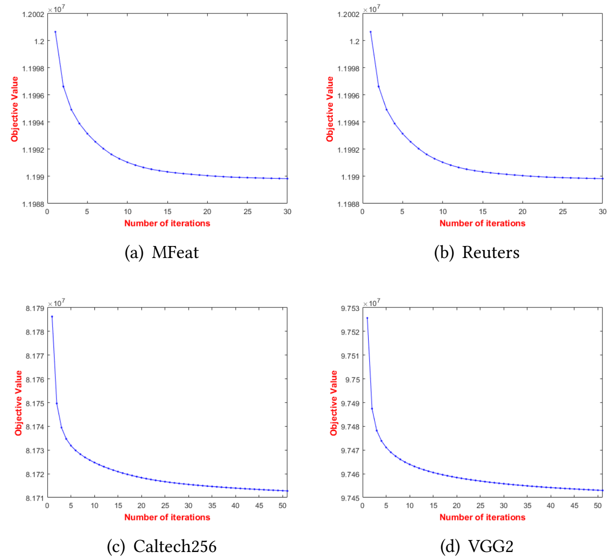
2.1 Convergence

We demonstrate the convergence of the proposed 3AMVC on four datasets (MFeat, Reuters, Caltech256, VGG2) in Figure 1. As seen, our 3AMVC algorithm can gradually converge to a stable state as the number of iterations increases.

2.2 Visualization of Anchor Selection

We record the number of anchors selected and the baseline view of alignment on all datasets using our HBNC under the best parameter settings when the best clustering performance is achieved. As shown in Table 1, our method does not always select the view with the largest number of anchors as the baseline view, because the quality of the anchor graph is not positively correlated with the number of anchors. Figure 2 can also prove that determining the baseline view by evaluating the quality of the anchor graph has a beneficial effect on the clustering performance.

In addition, we show the anchor selection results of the K-means and HBNC algorithms in three views on the ForestTypes dataset in Figure 3. From the experimental results of the ForestTypes dataset, as with the MFeat dataset, the baseline view we picked generally has the strongest representation ability of anchors and can better represent the internal structure in the original sample space.



(a) MFeat (b) Reuters (c) Caltech256 (d) VGG2

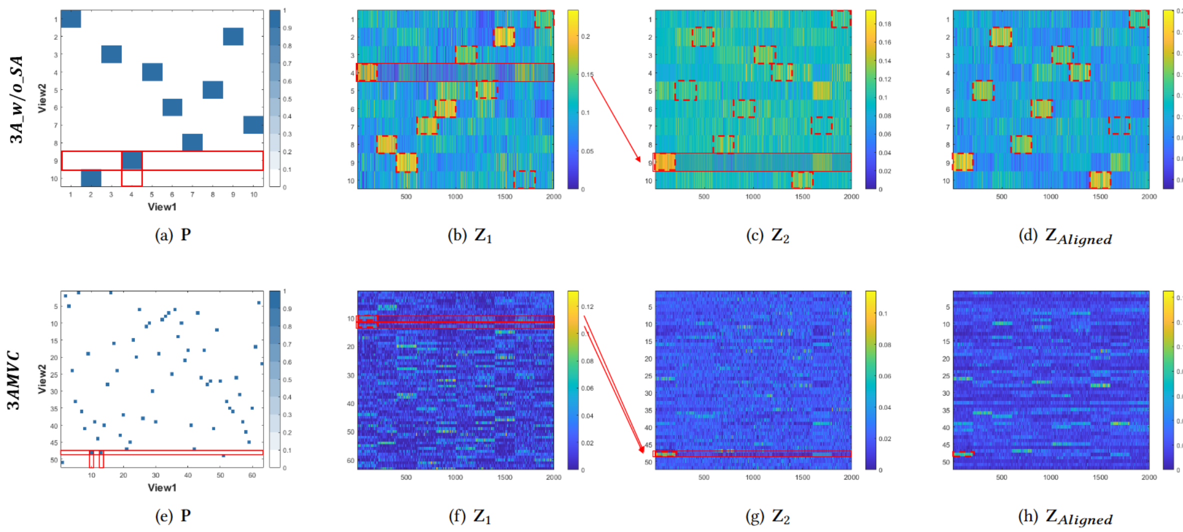
Figure 1: The variation of the objective value with the number of iterations on four datasets.

Table 1: The number of anchors selected by HBNC on all datasets.

Dataset	Baseline	View	Number of anchors
ForestTypes		1	10
		2	16
	✓	**3**	**20**
MFeat	✓	**1**	**54**
		2	64
Reuters		1	62
		2	48
		3	45
		4	13
	✓	**5**	**53**
Caltech256		1	48
		2	67
	✓	**3**	**62**
		4	45
VGG2		1	33
		2	74
	✓	**3**	**66**
		4	38

2.3 Visualization of Ablation Studies

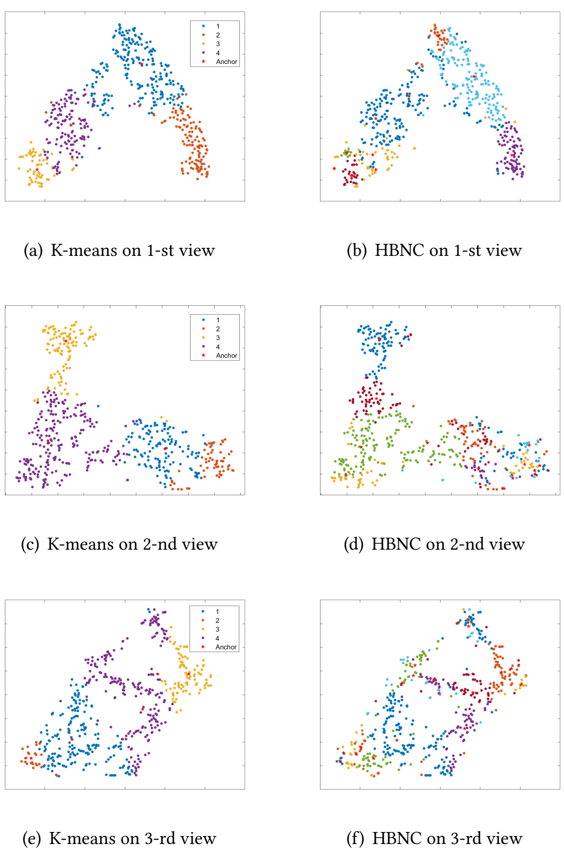
We visualize the ablation experiment on the MFeat dataset. Since the baseline view of the MFeat dataset is also the first view, in order to illustrate the effectiveness of our beaseline view alignment method, we exchange two views of the MFeat dataset and reconstruct the



(a) P (b) Z_1 (c) Z_2 (d) $Z_{Aligned}$

(e) P (f) Z_1 (g) Z_2 (h) $Z_{Aligned}$

Figure 2: The visualization of the anchor graphs on MFeat_test dataset.

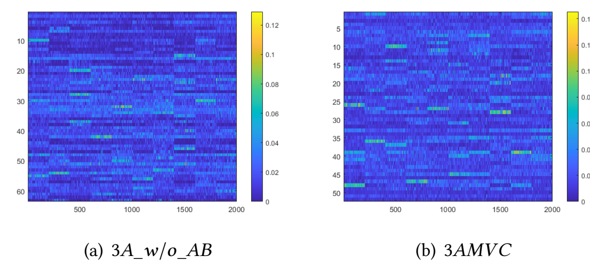


(a) K-means on 1-st view (b) HBNC on 1-st view

(c) K-means on 2-nd view (d) HBNC on 2-nd view

(e) K-means on 3-rd view (f) HBNC on 3-rd view

Figure 3: The visualization of anchor selection on Forest-Types dataset.



(a) *3A_w/o_AB* (b) *3AMVC*

Figure 4: The comparison of visualization of the fused anchor graphs with *3A_w/o_AB* and *3AMVC* on MFeat_test dataset.

Table 2: Experiment results of ablation studies

Datasets	Methods	ACC	NMI	Fscore
MFeat_test	*3A_neither*	0.8107	0.7226	0.6970
	3A_w/o_SA	0.8107	0.7226	0.6970
	3A_w/o_AB	0.8520	0.8178	0.7806
	3AMVC	**0.8775**	**0.8298**	**0.8044**

MFeat_test dataset, aiming to provide a clearer and more illustrative depiction of the impact of our strategies. As mentioned earlier, methods that do not use either strategy are recorded as *3A_neither*. *3A_w/o_SA* means that the anchors are not automatically selected, and we use the kmeans method instead. *3A_w/o_AB* stands for not aligning based on the baseline view, we align based on the first view uniformly. *3AMVC* means the complete algorithm.

The results of ablation experiments on the MFeat_test dataset are shown in Table 2. The effect of the baseline view is not reflected on *3A_w/o_SA*. We analyze that the k anchors may not adequately capture the original spatial structure inherent in the data, which results in an ineffective alignment outcome.

We visualize the anchor graph of $3A\_w/o\_SA$ and $3AMVC$ in Figure 2. In both methods, we take the representative anchor of the first cluster as an example. $3A\_w/o\_SA$ realizes the one-to-one correspondence of the anchor 4 in view 1 according to the anchor 9 in the baseline view 2, while $3AMVC$ method matches the anchor 48 of the base view 2 with the two representative anchors 10 and 13 in view 1. Consequently, compared with $3A\_w/o\_SA$, our anchor selection strategy is more flexible, and the alignment matrix $\mathbf{P}$ can handle one-to-many or many-to-one matching relationships. In addition, from the fused anchor graph, no matter how the anchor selection result is, the alignment operation can obtain a clearer clustering structure.

In addition, we compare the visualization of the fusion anchor graph $\mathbf{Z}_{Aligned}$ of $3A\_w/o\_AB$ and $3AMVC$. Observing Figure 4, we find that the clustering structure of the fused anchor graph can be clearer by aligning the views according to the view with the best anchor graph quality, which is conducive to improving the performance of clustering.

REFERENCES

[1] Eugene L Lawler. 1963. The quadratic assignment problem. *Management science* 9, 4 (1963), 586–599.
[2] Yao Lu, Kaizhu Huang, and Cheng-Lin Liu. 2016. A fast projected fixed-point algorithm for large graph matching. *Pattern Recognition* 60 (2016), 971–982.