Graph embedded incomplete multi-view clustering method with proximity relation estimation

Runze Chen$^{1,2}$, Jie Wen$^{1,2}$, Xiaoyue Chen$^3$, Yong Xu$^{1,2}$

$^1$Bio-Computing Research Centre, Harbin Institute of Technology, Shenzhen, Shenzhen, People's Republic of China
$^2$Shenzhen Key Laboratory of Visual Object Detection and Recognition, Shenzhen, People's Republic of China
$^3$School of Electric Power, South China University of Technology, Guangzhou, People's Republic of China

E-mail: yongxu2017@gmail.com

Abstract: More and more importance has been attached to multi-view clustering due to its outstanding performance against single-view clustering. Existing studies are on the basis of the assumption that every sample exists in all views. However, in reality, this is often not the case and the real data are always with incomplete views, which lead to the failure of the conventional methods. To address the issue, the authors present a novel method, referred to as incomplete multi-view spectral clustering with proximity relation estimation (IMSC_PRE), which exploits the proximity relations of samples in different views to complete their affinity graphs. Different from existing methods, IMSC_PRE fully excavates the manifold structures of all views to estimate the proximity relations of missing samples in each view, which presents a brand-new strategy to exploit the proximity information lying in the whole samples with good interpretation. Experimental results on four datasets demonstrate the advantages of the proposed method in comparison with the state-of-the-art methods.

1 Introduction

Nowadays, a great many objects to be recognised or detected are comprised of multiples modalities due to the prosperity of diversiform-data acquiring technique. For example, evidence used for real estate valuation crawled by web spiders may contain environmental images, press comments, economic indicators etc. A certain story is always described in different languages by several newspaper offices, and each kind of language can be regarded as a single view. All above cases illustrate that almost everything can be displayed from distinct perspectives. Clustering data with multiple modalities is proved to be more effective than clustering with single modality by lots of experiments these years, one palpable reason of which is that features from distinct channels always hold complementary and discrepancy information, which can enhance the performance [1, 2].

To this day, extensive studies on multi-view clustering have been carried out [1, 3, 4]. Kumar et al. [5] introduced a co-regularised constraint to utilise traditional spectral clustering method on multi-view data. By minimising the disagreement of kernel matrices constructed from low-dimensional representation of data in diverse views, the method learns a common compact representation of original data. By constraining the $l_1$ norm on the basis matrix learnt from each view, Liu et al. [6] provided a joint non-negative matrix factorisation (NMF) learning framework rendering meaningful probabilistic interpretation to the common representation matrix. All above techniques hold a common assumption that all modalities contain the whole samples; however, it is unreachable in lots of situations. For instance, a story may not be carried by all the newspaper offices. In other words, there is no newspaper agency covering all the news, which essentially provides an incomplete view. Such situations result in the failure of the existing methods. The incompleteness of views leads to the massive loss of information, and it is difficult to learn a consensus representation due to the different missing rates of views, making the incomplete multi-view clustering become a challenging problem.

To deal with the incomplete challenge, based on NMF, Li et al. [7] presented partial multi-view clustering (PVC) to partition data with two views, in which the representation of paired samples (i.e. the samples which embody all views) are forced to be the same. Remedying the objective function of PVC by appending a graph Laplacian term, method proposed by Zhao et al. [8] can learn a unified affinity graph over all views for incomplete multi-view clustering. Nevertheless, above two methods can only handle datasets with two views, and they both hold a hypothesis that there are some samples embodying all views, which restrict the fields of application [9]. Hu and Chen [10] first defined an indicator matrix for every individual view and proposed a weighted matrix factorisation-based incomplete multi-view learning model, which is able to handle the clustering cases with more than two views. In [11], missing instances were filled with the average features of the corresponding view, and then a diagonal indicator matrix for each view was set to endow available instances and filled instances with different weights to minimise the negative effect of the missing samples. To some extent, the methods mentioned above can meet the challenge of incomplete multi-view clustering problem; however, they are all based on NMF, resulting in their incapability to datasets with non-linear intrinsic structure in the incomplete multi-view clustering task.

In this paper, a novel graph embedded incomplete multi-view clustering method, dubbed as incomplete multi-view spectral clustering with proximity relation estimation (IMSC_PRE), is brought forward to address the issue. First, IMSC_PRE constructs an affinity graph with the $k$-nearest criterion ($k$-nearest graph) from the available instances of each view. Second, we make estimations of the proximity relations related to missing instances by averaging the proximity relations from the other views. Third, a co-regularisation term is introduced to get a more discriminative and compact representation shared by all views. On account of the fact that the information carried by each individual view may have significant difference, distinct weights are assigned to the learning models of distinct views.

Overall, the proposed method holds three contributions as follows: (i) our method is the first work to address the incomplete challenging problem from the angle of learning multiple completed affinity graphs from multiple views with arbitrary incomplete conditions. (ii) Compared to the existing methods, our method can well-excavate the manifold structure lying in the multiple views, which is beneficial to group the data with non-linear structure. (iii) In our method, suitable weights of distinct views are adaptively learnt, which can help the modal acquire a compact and
discriminative consensus representation. Experimental results conducted on several datasets show that our method significantly outperforms the state-of-the-art methods.

2 Related work

In this section, we briefly introduce the single-view spectral clustering [12]. For data \( X \in \mathbb{R}^{d \times n} \) with \( n \) samples and \( d \) dimensions, we first construct a symmetric similarity graph \( W \in \mathbb{R}^{n \times n} \), whose element \( W_{ij} \) represents the similarity weights of samples \( x_i \) and \( x_j \). The learning model is formulated as follows:

\[
\min_U \text{Tr}(U^T L U) \quad \text{s.t.} \quad U^T U = I. \tag{1}
\]

In the above formula, \( U \in \mathbb{R}^{n \times \ell} \) is the low-dimensional representation of origin data \( X \), \( \ell \) is the number of clusters. Under the radio-cut measure, \( L \in \mathbb{R}^{n \times n} \) is the normalised Laplacian matrix of similarity graph \( W \) calculated by \( L = D - W \), where \( D \) is a diagonal matrix with its \( i \)th diagonal element calculated by \( D_{ii} = \sum_j W_{ij} \). While under the normalised-cut measure, \( L \in \mathbb{R}^{n \times n} \) is the normalised Laplacian matrix of similarity graph \( W \) calculated by \( L = I - D^{-1/2}WD^{-1/2} \), in which \( D \) is the same as before.

It is easy to optimise problem (1), with the columns of solution \( U \) equaling to the eigenvector set corresponding to \( c \) smallest eigenvalues. Then, traditional clustering method such as \( k \)-means is conducted on the representation \( U \) to obtain the final clustering result.

3 Proposed method

3.1 Objective function

3.1.1 Construct complete affinity graphs of distinct views: Let \( X_v^{(i)} \in \mathbb{R}^{d \times n_v} \) be the set of instances in the \( v \)th view, where \( n_v \) and \( d_v \) denote the number of available instances and dimension of the \( v \)th view, respectively. To excavate the local manifold structure of each view, we first construct a \( k \)-nearest neighbour graph for each view, noted as \( W_v^{(i)} \in \mathbb{R}^{n_v \times n_v} \), whose elements are acquired by the formula as follows [13]:

\[
W_v^{(i,j)} = \begin{cases} 
1, & \text{if } x_v^{(i)} \in \Psi(x_v^{(j)}) \cap x_v^{(j)} \in \Psi(x_v^{(i)}), \\
0, & \text{otherwise} 
\end{cases} \tag{2}
\]

where \( \Psi(x_v^{(i)}) \) denotes the \( k \)-nearest neighbour set of the instance \( x_v^{(i)} \) based on the Euclidean metric.

Owing to the incompleteness of the problem, the size of individual affinity graph \( W_v^{(i)} \) is less than the total number of samples, i.e. \( n_v < n_v \), where \( n_v \) denotes the total number of samples. Here, we present a novel strategy to complete the affinity graph by fully excavating the intrinsic structures of all views. To enable the affinity graph \( W_v^{(i)} \) to reveal the proximity relations across all samples, we complete the missing relations concerning absent instances according to the proximity relations of all views.

First, we can exploit the following formula to extend the size of each incomplete graph to the same size \( n \):

\[
\tilde{W}_v^{(i)} = G_v^{(i)} \gamma_v W_v^{(i)} G_v^{(i)}, \tag{3}
\]

where \( G_v^{(i)} \in \mathbb{R}^{n_v \times n_v} \) is the index matrix of the \( v \)th view, defined as follows:

\[
G_v^{(i)} = \begin{cases} 
1, & \text{if } x_v^{(i)} \text{ is the original instance } \tilde{x}_v^{(i)}, \\
0, & \text{otherwise} 
\end{cases} \tag{4}
\]

where \( \tilde{X}_v = [\tilde{x}_1, ..., \tilde{x}_n] \in \mathbb{R}^{d_v \times n} \) denotes the set of original instances including the available instances and missing instances of the \( v \)th view.

In (3), all elements related to the absent instances are filled in with 0 in the extended affinity graph \( \tilde{W}_v^{(i)} \in \mathbb{R}^{n \times n} \) of each view. However, when the missing ratio is large, these graphs do not contain sufficient information to guide the clustering, resulting in bad performance. To address the issue, based on the idea that proximity relations between instances are approximately consistent across multiple views, we propose a simple and effective approach by exploiting the proximity relations of all views to further complete the graph. Specifically, for each missing instance \( x_v^{(i)} \) of the \( v \)th view (that is to say, the \( v \)th view of sample \( x_i \) is absent), we exploit the following formula to estimate its proximity relation:

\[
\tilde{W}_v^{(i)} = \left( \sum_{k=1}^n H_{v,k} \tilde{W}_k^{(i)} \right) / \left( \sum_{k=1}^n H_{v,k} \right), \tag{5}
\]

where \( \tilde{W}_k^{(i)} \) denotes the \( i \)th row of the affinity graph \( W_k^{(i)} \) of the \( k \)th view, i.e. the proximity relations related to the \( i \)th instance in the \( k \)th view \( x_k^{(i)} \). \( H_v \in \mathbb{R}^{n \times n} \) is a matrix pre-defined to denote the occurrence of the instances in the views, the detail is formulated as follows:

\[
H_{v,j} = \begin{cases} 
1, & \text{if the } i \text{th sample occurs in the } j \text{th view} \\
0, & \text{otherwise} 
\end{cases}. \tag{6}
\]

Formula (5) enables us to estimate the proximity relations in the missing view and excavate complementary information across all views simultaneously. Then, we symmetrise the graph acquired above by exploiting the following formula:

\[
\tilde{W}^{(i)} = \left( \tilde{W}^{(i)} + W^{(i)\top} \right) / 2. \tag{7}
\]

Till now, the affinity graph for each view is completed, by fully excavating the local manifold structures of multiple views. We denote the completed graph as \( \tilde{W}^{(i)} \).

3.1.2 Learn a consensus representation across all views: After we obtain the completed graphs of all views, we can learn the compact representation of each view by exploiting the spectral clustering as follows:

\[
\min_{U_v^{(i)}} \sum_v \text{Tr}(U_v^{(i)\top} L_v^{(i)} U_v^{(i)}) \quad \text{s.t.} \quad U_v^{(i)\top} U_v^{(i)} = I. \tag{8}
\]

where \( U_v^{(i)} \in \mathbb{R}^{n \times \ell_v} \) is the low-dimensional representation of the \( v \)th view, \( \ell_v \) is the number of clusters, \( L_v^{(i)} \) is the Laplacian graph for the \( v \)th view, which is calculated by \( L_v^{(i)} = D_v^{(i)} - W_v^{(i)} \), and \( D_v^{(i)} \) is a diagonal matrix, whose \( i \)th diagonal element is calculated by \( D_v^{(i)}_{ii} = \sum_j W_v^{(i,j)} \).

By virtue of the fact that each view is just a representation form of the same sample set, samples across all views hold a consensus representation \( U_v^{(i)} \) naturally [1, 14]. To learn the consensus representation \( U_v^{(i)} \) for clustering, we introduce a co-regularisation term as follows:

\[
\min_{U_v^{(i)}, U_v^{(i)}} \sum_v \left( \text{Tr}(U_v^{(i)\top} L_v^{(i)} U_v^{(i)}) + \lambda \Gamma(U_v^{(i)}, U_v^{(i)}) \right) \quad \text{s.t.} \quad U_v^{(i)\top} U_v^{(i)} = I, \quad U_v^{(i)} = U_v^{(i)} \tag{9}
\]

where function \( \Gamma(U_v^{(i)}, U_v^{(i)}) \) is the disagreement term to judge the disagreement of the consensus representation \( U_v^{(i)} \) and representation \( U_v^{(i)} \), \( \lambda \) is the trade-off parameter between the above disagreement term and spectral clustering objects. Specifically, introducing the co-regularisation term enables the method to capture the complementary information across different views, which is beneficial to improve the clustering performance. \( \Gamma(U_v^{(i)}, U_v^{(i)}) \) is defined as follows [5]:

\[
\Gamma(U_v^{(i)}, U_v^{(i)}) = \text{Tr}(U_v^{(i)\top} L_v^{(i)} U_v^{(i)}) + \lambda \text{Tr}(U_v^{(i)\top} Y_v^{(i)} U_v^{(i)})
\]

with \( Y_v^{(i)} \) defined as follows:

\[
Y_v^{(i)} = \begin{cases} 
0, & \text{if } U_v^{(i)}_{ii} = 0 \\
I, & \text{otherwise}
\end{cases}
\]

J. Eng., 2020, Vol. 2020 Iss. 13, pp. 647-652

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
\[ \Gamma(U^v, U) = \left\| \frac{K_U}{\| K_U \|_F} - \frac{K_{U^v}}{\| K_{U^v} \|_F} \right\|_F^2, \]  
\[ (10) \]

where \( K_U \) and \( K_{U^v} \) are the kernel graphs of representations \( U \) and \( U^v \), respectively. Inspired by [5, 15], we exploit the linear kernel to construct the kernel graph, i.e. \( K(x, x') = x x' \); thus, we have \( K_U = U U^T \) and \( \| K_U \|_F = c \). By overlooking the scaling factor, we can rewrite (10) as follows:

\[ \Gamma(U^v, U) \propto \langle c - \text{Tr}(U^v(U^v)^T U U^T) \rangle. \]
\[ (11) \]

Overall, the objective function to learn the consensus low-dimensional representation of all views can be presented as the following formula:

\[ \min_{U^v} \sum_v \left( \text{Tr}(U^v(U^v)^T U^v) + \lambda \langle c - \text{Tr}(U^v(U^v)^T U U^T) \rangle \right) \]
\[ \text{s.t. } U^v(U^v)^T = I, \quad U^T U = I. \]
\[ (12) \]

### 3.1.3 Endow different weights to distinct views and overall objective function

Taking into the consideration the fact that the discriminative powers of different views maybe extremely unequal due to the distinct missing rates and meanings of feature, we assign a weight to each view to represent the significance of each view, which can be formulated as follows:

\[ \min_{U^v, \alpha^v} \sum_v (\alpha^v) \left( \text{Tr}(U^v(U^v)^T U^v) + \lambda \langle c - \text{Tr}(U^v(U^v)^T U U^T) \rangle \right) \]
\[ \text{s.t. } U^v(U^v)^T = I, \quad U^T U = I, \quad \sum_v \alpha^v = 1, \quad \alpha^v \geq 0. \]
\[ (13) \]

where the symbol \( \gamma_v \) refers to the objective function of the \( v \)th view, \( \alpha^v \) is adopted to weight distinct significance of different views, and \( r > 1 \) is the super-parameter to smooth the contribution. By endowing each view with a suitable weight, the proposed method is able to learn a more compact and discriminative representation. Overall, the objective function is formulated as follows:

\[ \min_{U^v, \alpha^v} \sum_v (\alpha^v) \left( \text{Tr}(U^v(U^v)^T U^v) + \lambda \langle c - \text{Tr}(U^v(U^v)^T U U^T) \rangle \right) \]
\[ \text{s.t. } U^v(U^v)^T = I, \quad U^T U = I, \quad \sum_v \alpha^v = 1, \quad \alpha^v \geq 0. \]
\[ (14) \]

### 3.2 Model optimisation

The constrained problem (14) is difficult to solve; thus, we update the unknown variables iteratively to achieve the local optimum:

**Step 1:** Update the consensus representation \( U \): (14), concerning \( U \) can be rewritten as follows with \( U^v \) and \( \alpha^v \) fixed:

\[ \max_{U^v \in \mathcal{C}} \text{Tr} \left(U^v(U^v)^T \sum_v \alpha^v (U^v(U^v)^T U^v)^{\frac{1}{2}} \right) \]
\[ (15) \]

This is a typical eigenvalue decomposition problem, and the optimal solution to \( U \) is the eigenvector set corresponding to \( c \) biggest eigenvalues of matrix \( \sum_v \alpha^v (U^v(U^v)^T U^v)^{\frac{1}{2}} \).

**Step 2:** Update representation \( U^v \): By fixing \( U \) and \( \alpha^v \), the objective function concerning \( U^v \) can be rewritten as follows:

\[ \min_{U^v \in \mathcal{C}} \text{Tr}(U^v(U^v)^T U^v) - \lambda \text{Tr}(U^v(U^v)^T U U^T) \]
\[ \Leftrightarrow \max_{U^v \in \mathcal{C}} \text{Tr}(U^v(U^v)^T U^v(U^v)^T U^v) \]
\[ (16) \]

Equation (16) is also a typical eigenvalue decomposition problem, and the eigenvector set corresponding to the first \( c \) biggest eigenvalues of matrix \( (U^v(U^v)^T - L^v) \) is selected as the optimal solution to variable \( U^v \).

**Step 3:** Update weights \( \alpha^v (v = 1, ..., l) \): The optimisation problem concerning \( \alpha^v \) is degraded to the following formula:

\[ \min_{\alpha^v} \sum_v (\alpha^v) \gamma_v \text{s.t. } \sum_v \alpha^v = 1, \quad \alpha^v \geq 0. \]
\[ (17) \]

where \( \gamma_v = \text{Tr}(U^v(U^v)^T U^v) + \lambda \langle c - \text{Tr}(U^v(U^v)^T U U^T) \rangle \).

The solution to (17) is formulated as follows:

\[ \alpha^v = (\gamma_v^{\frac{1}{2}})^{\frac{1}{2}} \left( \sum_v (\gamma_v)^{\frac{1}{2}} \right)^{-1}. \]
\[ (18) \]

Algorithm 1 (Fig. 1) summarises the optimisation procedure wholly. After the consensus representation \( U \) is acquired, \( k \)-means algorithm is utilised on \( U \) to acquire the final clustering result.

### 4 Experiment

#### 4.1 Experimental settings

##### 4.1.1 Datasets

The summarisation of datasets used in this experiment is shown in Table 1. More details are illustrated as follows: (i) BUAA-VisirFace (BUAA) dataset [https://github.com/hdzhao/IMG/tree/master/data.][16]: BUAA dataset includes two image modalities regarded as two views of a person, i.e. visual image and near-infrared image. A subset of 90 images related to 10 volunteers is utilised in the experiment, and each image adopted has been resized into a 10 × 10 matrix, and then vectorised ahead of time. (ii) Caltech101 dataset [17]: The whole Caltech dataset is composed of 101 object classes. Following the experiment setting in [15], we adopt a subset [https://www.dropbox.com/s/ulvatoo8gepcdfk/Caltech101-7.mat.] dubbed as Caltech7 containing 1474 samples of 7 classes to carry out our experiment. In our experiment, GIST feature and local binary patterns (LBP) feature are regarded as two views of each sample. (iii) BBCSport dataset [18]: The full dataset comprises 737 documents classified to five kinds of ball games. A subset [https://github.com/GPMVC Dummy/GPMVC/tree/master/ partialMV/PVC/recreateResults/data] containing 116 instances from 5 classes with four views in each instance is exploited in the experiment. (iv) 3 Sources dataset: This dataset is an incomplete multi-view text dataset. As a subset [https://github.com/GPMVDummy/GPMVC/tree/master/partialMV/PVC/recreateResults/data] of the full dataset, 169 instances related to 6 topics are composed of all three sources, which is used in our experiment.

Notably, two kinds of incomplete multi-view datasets are constructed as follows: (i) For the BUAA and Caltech7 datasets, we construct several incomplete two views datasets with 10, 30, 50, 70, and 90% randomly selected samples as paired samples, and then half of unpaired samples discard the first view, with another half discarding the second view. (ii) For the BBCSport and 3 Sources datasets, we randomly remove 10, 30, and 50% samples from each view, ensuring that no sample lost all views simultaneously.

##### 4.1.2 Compared methods and performance metrics

To prove the advantages of our method, we select several state-of-the-art methods including best single view (BSV) [8], Concat [8], PVC [7], graph regularised PVC (GPMVC) [19], incomplete multi-modality grouping (IMG) [8], multi-incomplete-view clustering (MIC) [11], doubly aligned incomplete multi-view clustering (DAIMC) [10], and online multi-view clustering (OMVC) [20]. To assess the performance of each method, three metrics were adopted in our experiment, i.e. the clustering accurate (ACC), normalised mutual information (NMI), and purity (PUR) [21].

There are several points worth noting as follows: (i) For the compared methods excluding BSC and Concat, grid search strategy
Experimental results and analysis

Experimental results of distinct methods on two kinds of incomplete multi-view datasets are displayed in Tables 2 and 3, in which the best results are denoted as bold. We discover the following points from the experimental results. (i) In the experiment, BSV and Concat perform worst than others in most circumstances. They both fill in the missing instances with the average features of the corresponding views, which is harmful to the performance. In contrast with above strategy, our method, which aims to estimate the proximity relations related to absent instance, achieves good performance, which further illustrates the rationality of our method. (ii) From the experimental results, we can find that GPMVC and our method always perform better than PVC, which maybe attributed to the graph embedding technique that GPMVC and the proposed method utilised. This indicates that learning the intrinsic structure can help us to achieve better performance. (iii) For all datasets, our method outperforms the others in most situations, especially on the 3 Sources dataset. The phenomena strongly demonstrate the effectiveness of our estimation for missing proximity relations in affinity graph with proximity relations of all views.

Fig. 1  Algorithm 1: IMSC_PRE [solving (14)]

Table 1  Summary of the used datasets

| Dataset       | #Class | #View | #Sample | #Features |
|---------------|--------|-------|---------|-----------|
| BUAA          | 10     | 2     | 90      | 100/100   |
| Caltech7      | 7      | 2     | 1474    | 512/928   |
| 3 Sources     | 6      | 3     | 169     | 3560/3631/3608 |
| BBCSport      | 5      | 4     | 116     | 1991/2063/2113/2158 |

Table 2  Mean ACCs (%), NMIs (%), and PURs (%) of distinct methods on BUAA and Caltech7 datasets. Note: Bold values denote the best results

| Dataset       | Methods/Paired rate | 0.1  | 0.3  | 0.5  | 0.7  | 0.1  | 0.3  | 0.5  | 0.7  | 0.1  | 0.3  | 0.5  | 0.7  |
|---------------|---------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| BUAA          | BSV                 | 48.33| 56.96| 64.26| 70.81| 43.10| 53.03| 61.78| 69.91| 50.19| 58.66| 65.96| 72.35|
|               | Concat              | 45.62| 46.61| 47.46| 52.34| 51.22| 51.95| 52.43| 56.51| 47.99| 49.41| 49.68| 54.81|
|               | PVC                 | 57.41| 66.46| 70.01| 75.92| 61.35| 67.07| 71.97| 78.70| 60.58| 68.79| 71.98| 77.95|
|               | GPMVC               | 58.98| 68.75| 74.28| 78.28| 62.12| 70.25| 74.33| 81.63| 61.26| 71.66| 75.36| 81.71|
|               | IMG                 | 53.95| 67.39| 76.14| 79.36| 54.72| 67.53| 76.74| 82.83| 57.32| 68.65| 78.23| 82.42|
|               | MIC                 | 57.77| 70.00| 68.22| 74.89| 59.92| 72.59| 69.56| 74.11| 61.11| 73.33| 70.44| 75.33|
|               | DAIMC               | 57.56| 71.55| 79.78| 80.67| 60.16| 74.63| 79.13| 83.32| 60.44| 74.00| 79.33| 82.00|
|               | OMVC                | 60.59| 62.89| 65.74| 68.37| 60.69| 63.99| 64.32| 66.91| 62.15| 65.57| 67.48| 69.78|
|               | ours                | 60.33| 73.10| 80.31| 86.55| 63.79| 74.91| 80.96| 85.81| 64.16| 75.34| 81.81| 87.10|
| Caltech7      | BSV                 | 42.66| 40.97| 39.83| 42.92| 29.04| 32.11| 35.13| 38.83| 32.06| 75.61| 79.18| 83.02|
|               | Concat              | 36.83| 31.74| 36.36| 43.38| 33.82| 34.44| 34.56| 38.15| 38.83| 72.06| 79.18| 83.02|
|               | PVC                 | 43.46| 43.96| 44.46| 44.76| 38.99| 40.26| 40.17| 41.60| 43.17| 83.85| 83.92| 85.17|
|               | GPMVC               | 45.57| 47.19| 46.99| 46.99| 40.05| 40.96| 41.83| 42.61| 83.30| 84.02| 84.19| 86.23|
|               | IMG                 | 42.05| 42.36| 42.23| 41.17| 32.38| 33.29| 35.05| 35.96| 77.54| 79.60| 80.28| 81.65|
|               | MIC                 | 43.93| 48.34| 49.85| 50.05| 33.77| 37.59| 40.15| 40.70| 80.80| 81.19| 83.69| 84.35|
|               | DAIMC               | 39.40| 45.52| 50.46| 48.79| 31.55| 36.11| 41.22| 41.20| 79.38| 81.89| 84.53| 84.36|
|               | OMVC                | 41.31| 41.49| 43.36| 44.53| 30.89| 31.54| 34.41| 37.12| 78.39| 80.46| 81.69| 82.53|
|               | ours                | 57.27| 56.89| 55.93| 57.30| 43.95| 45.68| 45.90| 47.30| 84.68| 85.15| 85.20| 85.55|

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
4.3 Analysis of the parameter sensitivity

Fig. 2(a) shows the trend of PUR variation with the smooth parameter $r$ on BUAA with 50 and 70% paired rates, and Fig. 2(b) shows the trend of PUR variation with the trade-off parameter $\lambda$ on 3 Sources dataset with 10 and 30% missing rates. It can be observed that the clustering performance indicators are relatively stable in some interval ranges, which means that our method is insensitive to the super-parameters to some extent.

4.4 Convergence analysis

Fig. 3 illustrates the trend of objective function value variation with optimisation iterations on BUAA dataset with 50% paired rate and BBCSport dataset with 30% missing rate. It can be observed from the figures that the objective function is not increasing, and decreases rapidly to the local optimum in the early several iterations. The phenomenon indicates that the proposed method holds a good convergence property.

5 Conclusion

In this paper, we present a novel co-regularised spectral clustering method with proximity relation estimation for arbitrary incomplete multi-view problem. Different from the existing methods, it is the first work to excavate proximity relations of all views to estimate the proximity relations of missing samples in each view during the affinity graph construction period. The co-regularised term helps the method acquire a consensus representation. We also endow each view with the suitable weight, which is adaptively learnt to achieve a discriminative and compact representation of origin data, followed by $k$-means clustering algorithm. Experiments on two types of incomplete multi-view datasets demonstrate the effectiveness of our method in comparison with some state-of-the-art methods. In the future, we will focus on exploring effective incomplete multi-view clustering methods to handle large-scale datasets.

6 Acknowledgment

This work was supported in part by Establishment Funding of Shenzhen Key Laboratory of Visual Object Detection and Recognition under Grant No. ZDSYS20190902093015527, and in part by the National Postdoctoral Program for Innovative Talent under Grant No. BX20190100.

7 References

[1] Yang, Y., Wang, H.: ‘Multi-view clustering: a survey’, Big Data Min. Anal., 2018, 1 (2), pp. 83–107
[2] Zhang, Z., Lu, L., Shen, F., et al.: ‘Binary multi-view clustering’, IEEE Trans. Pattern Anal. Mach. Intell., 2019, 4 (7), pp. 1774–1782

Table 3

| Dataset      | Methods/Missing rate | ACC 0.1 | ACC 0.3 | ACC 0.5 | NMI 0.1 | NMI 0.3 | NMI 0.5 | PUR 0.1 | PUR 0.3 | PUR 0.5 |
|--------------|----------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 3 Sources    | BSV                  | 56.90   | 47.38   | 39.24   | 50.07   | 34.46   | 22.34   | 68.14   | 57.63   | 48.99   |
|              | Concat               | 53.54   | 46.79   | 37.68   | 51.98   | 37.87   | 18.32   | 69.78   | 58.51   | 46.48   |
|              | GPMVC                | 48.24   | 44.50   | 42.01   | 34.82   | 30.44   | 28.15   | 60.47   | 58.58   | 57.40   |
|              | MIC                  | 49.11   | 47.69   | 42.49   | 37.23   | 38.62   | 26.08   | 57.28   | 61.30   | 52.31   |
|              | DAIMC                | 56.33   | 52.43   | 50.73   | 52.98   | 49.07   | 41.64   | 68.99   | 67.21   | 63.56   |
|              | OMVC                 | 43.95   | 41.11   | 39.53   | 36.48   | 28.42   | 24.34   | 59.37   | 48.76   | 45.44   |
|              | ours                 | 81.88   | 72.09   | 58.46   | 69.10   | 63.28   | 48.47   | 83.23   | 79.75   | 70.02   |
| BBCSport     | BSV                  | 58.62   | 51.31   | 44.03   | 43.73   | 31.03   | 21.40   | 65.79   | 55.07   | 47.59   |
|              | Concat               | 70.62   | 58.72   | 33.21   | 61.69   | 38.92   | 18.61   | 80.59   | 63.24   | 37.00   |
|              | GPMVC                | 51.44   | 46.89   | 43.91   | 28.23   | 20.04   | 15.48   | 58.39   | 52.76   | 45.29   |
|              | MIC                  | 51.21   | 46.21   | 46.03   | 29.90   | 25.84   | 24.01   | 55.00   | 51.72   | 52.41   |
|              | DAIMC                | 68.62   | 63.45   | 56.89   | 56.62   | 50.17   | 37.89   | 76.90   | 71.72   | 61.03   |
|              | OMVC                 | 53.33   | 51.38   | 48.79   | 30.64   | 41.57   | 40.63   | 56.49   | 59.20   | 57.47   |
|              | ours                 | 77.41   | 79.29   | 67.62   | 71.99   | 68.32   | 52.71   | 88.05   | 86.41   | 76.07   |

Fig. 2

(a) PUR (%) versus $r$ of the proposed method on BUAA with 50 and 70% paired rates, (b) PUR (%) versus $\lambda$ of the proposed method on 3 Sources with 10 and 30% missing rates

Fig. 3

(a) BFVA with 50% paired rate, (b) BBCSport with 30% missing rate

J. Eng., 2020, Vol. 2020 Iss. 13, pp. 647-652
This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)
Li, J., Yong, H., Zhang, B., et al.: ‘A probabilistic hierarchical model for multi-view and multi-feature classification’. Proc. AAAI Conf. Artificial Intelligence, New Orleans, Louisiana, February 2018, pp. 3498–3505

Nie, F., Tian, L., Li, X.: ‘Multi-view clustering via adaptively weighted projections’. Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, London, UK, August 2018, pp. 2022–2030

Kumar, A., Rai, P., Daume, H.: ‘Co-regularized multi-view spectral clustering’. Proc. Advances in Neural Information Processing Systems, Montreal, Canada, December 2011, pp. 1413–1421

Liu, J., Wang, C., Guo, J., et al.: ‘Multi-view clustering via joint non-negative matrix factorization’. Proc. SIAM Int. Conf. Data Mining, Austin, TX, May 2013, pp. 252–260

Li, S.-Y., Jiang, Y., Zhou, Z.-H.: ‘Partial multi-view clustering’. Proc. AAAI Conf. Artificial Intelligence, Quebec City, Canada, July 2014, pp. 1969–1974

Zhao, H., Liu, H., Fu, Y.: ‘Incomplete multi-modal visual data grouping’. Proc. Int. Joint Conf. Artificial Intelligence, New York, July 2016, pp. 2392–2398

Wen, J., Zhong, Z., Xu, Y., et al.: ‘Incomplete multi-view clustering via graph regularized matrix factorization’. Proc. European Conf. Computer Vision Workshop, Munich, Germany, September 2018, pp. 593–608

Hu, M., Chen, S.: ‘Doubly aligned incomplete multi-view clustering’. Proc. Int. Joint Conf. Artificial Intelligence, Stockholm, Sweden, July 2018, pp. 2262–2268

Shao, W., He, L., Philip, S.Y.: ‘Multiple incomplete views clustering via weighted non-negative matrix factorization with $I_{2,1}$ regularization’. Proc. Joint European Conf. Machine Learning and Knowledge Discovery in Databases, Porto, Portugal, September 2015, pp. 318–334

Luxburg, U.V.: ‘A tutorial on spectral clustering’. Stat. Comput., 2007, 17, (4), pp. 395–416

Wen, J., Han, N., Fang, X., et al.: ‘Low-rank preserving projection via graph regularized reconstruction’. IEEE Trans. Cybern., 2019, 49, (4), pp. 1279–1291

Tian, C., Sun, G., Zhang, Q., et al.: ‘Integrating sparse and collaborative representation classifications for image classification’, Int. J. Image Graph., 2017, 17, (2), pp. 786–791

Wen, J., Xu, Y., Liu, H.: ‘Incomplete multiview clustering with adaptive graph learning’. IEEE Trans. Cybern., to be published, doi: 10.1109/TCYB.2018.2884715

Huang, D., Sun, J., Wang, Y.: ‘The BUAA-VisNir face database instructions’. Technical Report IRIP-TR-12-FR-001, School Computer Science Engineering, Beihang University, Beijing, China, 2012

Li, F.-F., Fergus, R., Perona, P.: ‘Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories’, Comput. Vis. Image Understanding, 2004, 106, (1), pp. 59–70

Greene, D., Cunningham, P.: ‘Practical solutions to the problem of diagonal dominance in kernel document clustering’. Proc. Int. Conf. Machine Learning, Pittsburgh, PA, June 2006, pp. 377–384

Rai, N., Negi, S., Chaudhury, S., et al.: ‘Partial multi-view clustering using graph regularized NMF’. Proc. Int. Conf. Pattern Recognition, Cancun, Mexico, December 2016, pp. 2192–2197

Shao, W., He, L., Lu, C.-T., et al.: ‘Online multi-view clustering with incomplete views’. Proc. IEEE Int. Conf. Big Data, Washington, D.C., USA, December 2016, pp. 1012–1017

Fan, L., Zhong, J.: ‘New clustering algorithm for multidimensional data’, J. Digit. Inf. Manag., 2016, 14, (4), pp. 255–262

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)