SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

May 2022, Vol. 65 152105:1–152105:16
https://doi.org/10.1007/s11432-020-3369-8

Incomplete multi-view clustering via local and global co-regularization

Jiye LIANG[1*], Xiaolin LIU[1], Liang BAI[1], Fuyuan CAO[1] & Dianhui WANG[2,3]

[1]*Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan* 030006*, China;*
[2]*Artificial Intelligence Research Institute, China University of Mining and Technology, Xuzhou* 221116*, China;*
[3]*Department of Computer Science and Computer Engineering, La Trobe University, Melbourne* VIC 3086*, Australia*

Received 7 December 2020/Revised 4 May 2021/Accepted 30 August 2021/Published online 18 April 2022

Abstract The incompleteness of multi-view data is a phenomenon associated with real-world data mining applications, which brings a huge challenge for multi-view clustering. Although various types of clustering methods, which try to obtain a complete and consensus clustering result from a latent subspace, have been developed to overcome this problem, most methods excessively rely on views-public instances to bridge the connection with view-private instances. When lacking sufficient views-public instances, existing methods fail to transmit the information among incomplete views effectively. To overcome this limitation, we propose an incomplete multi-view clustering algorithm via local and global co-regularization (IMVC-LG). In this algorithm, we define a new objective function that is composed of two terms: local clustering from each view and global clustering from multiple views, which constrain each other to exploit the local clustering information from different incomplete views and determine a global consensus clustering result, respectively. Furthermore, an iterative optimization method is proposed to minimize the objective function. Finally, we compare the proposed algorithm with other state-of-the-art incomplete multi-view clustering methods on several benchmark datasets to illustrate its effectiveness.

Keywords incompleteness, multi-view clustering, local clustering, global clustering, co-regularization

Citation Liang J Y, Liu X L, Bai L, et al. Incomplete multi-view clustering via local and global co-regularization. Sci China Inf Sci, 2022, 65(5): 152105, https://doi.org/10.1007/s11432-020-3369-8

1 Introduction

Clustering is a classical, unsupervised learning technique [1–4], which has been extensively studied in several fields [5–8]. Clustering aims at grouping a set of unlabeled instances into meaningful clusters, and it requires instances in the same cluster to be highly similar and be dissimilar in different clusters. Recently, increasing attention has been given to clustering multi-view data with records made up of multiple heterogeneous representations from different sources [9–12]. This task is of significant practical relevance in several fields, such as scene recognition [13, 14], natural language semantic segmentation [15], biomedical analysis [16], and recommendation system [17].

However, in real-world applications, incompleteness is a phenomenon associated with multi-view data [18–21]. Figure 1 shows topic detection and tracking (TDT) as an example. News reports can be obtained in several different formats, such as text, pictures, or videos. It is obvious that different media provide diverse aspects of news. However, in reality, agencies prefer partial presentations, which may include some statements absent from views. In other words, not every instance in all the views is sampled. Thus, multi-view data with missing instances of views is referred to as incomplete multi-view data. The goal of traditional multi-view clustering methods is to reconcile the divergences by integrating the knowledge among multiple views or adapting information from one view to another so that a coincident clustering result is formed. Nevertheless, the presence of an incomplete setting generally disables the traditional multi-view clustering methods. The missing instance patterns become even more complex

* Corresponding author (email: ljy@sxu.edu.cn)

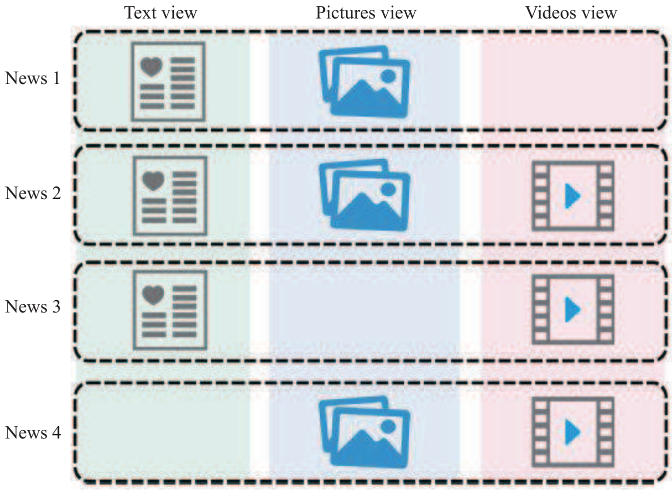

Figure 1 An example of incomplete multi-view data.

for the data with more views. In such a situation, it is difficult to learn a hidden, common representation for all of the incomplete views. Therefore, leveraging the complementary knowledge hidden in different views and reducing the impact of missing instances are the most challenging problems for incomplete multi-view clustering.

Recently, there has been an increasing interest in incomplete multi-view clustering. The related algorithms can be roughly classified into three categories.

(1) Nonnegative matrix factorization (NMF)-based methods aim to learn a latent common subspace in which views-public instances are enforced to have the same representations. As a pioneering work, partial multi-view clustering (PVC) [22] has been introduced, which learns the representations of both view-private instances and views-public instances simultaneously. All instances are uniformly represented in a potential subspace. Incomplete multi-modality grouping (IMG) [23] can be seen as an extension of PVC, which requires generated subspaces to conform to a global manifold structure. Multiple incomplete view clustering (MIC) [24] uses weighted NMF with $L_{2,1}$ regularization to obtain a more robust consensus representation. NMF-based methods emphasize the consistency of latent subspace without considering the local structure of each view; thus, the local cluster structure of each view may be destroyed in the latent space.

(2) Graph-based methods aim to learn a complete graph (or similarity matrix) from different views that reveals the relationships of all instances. Graph-based incomplete multi-view clustering (GIMC) [25] is the first attempt at constructing a complete graph for partitioning incomplete multi-view data. First, GIMC constructs a graph for each view. Subsequently, it fuses all the constructed graph matrices to generate a consensus graph matrix. Finally, it transforms the clustering problem into a graph segmentation problem. Partial multi-view clustering via auto-weighting similarity completion (PMVC-ASC) [26] jointly learns the consensus similarity matrix by exploring complementary information among multiple distinct feature sets. The anchor-based partial multi-view clustering (APMC) [27] method utilizes anchors to reconstruct instance-to-instance relationships for clustering. Incomplete multi-view clustering via structured graph learning (SGL) [28] combines the view-private instances and views-public instances to excavate an ideal structured graph, and then SGL utilizes graph segmentation technology to divide all the instances. Adaptive sample-level graph combination for partial multi-view clustering (ASGC-PMVC) [29] learns a latent subspace using adaptive graph fusion, which can seamlessly integrate the complementary and consistent information from multiple views and enhance the clustering performance on incomplete multi-view data. Additionally, adaptive graph completion-based incomplete multi-view clustering (AGCIMC) [30] develops a joint framework for graph completion and consensus representation learning. Generalized incomplete multi-view clustering with flexible locality structure diffusion (GIMC-FLSD) [31] proposes a graph-regularized framework for structure preserving and representation learning. Incomplete multi-view spectral clustering with adaptive graph learning (IMSC-AGL) [32] integrates graph construction and consensus representation learning into a joint optimization framework for incomplete multi-view clustering. Unified tensor framework for incomplete multi-view clustering and missing-view inferring (IMVTSC-MVI) [33] seeks to recover missing views and explores the full information of such

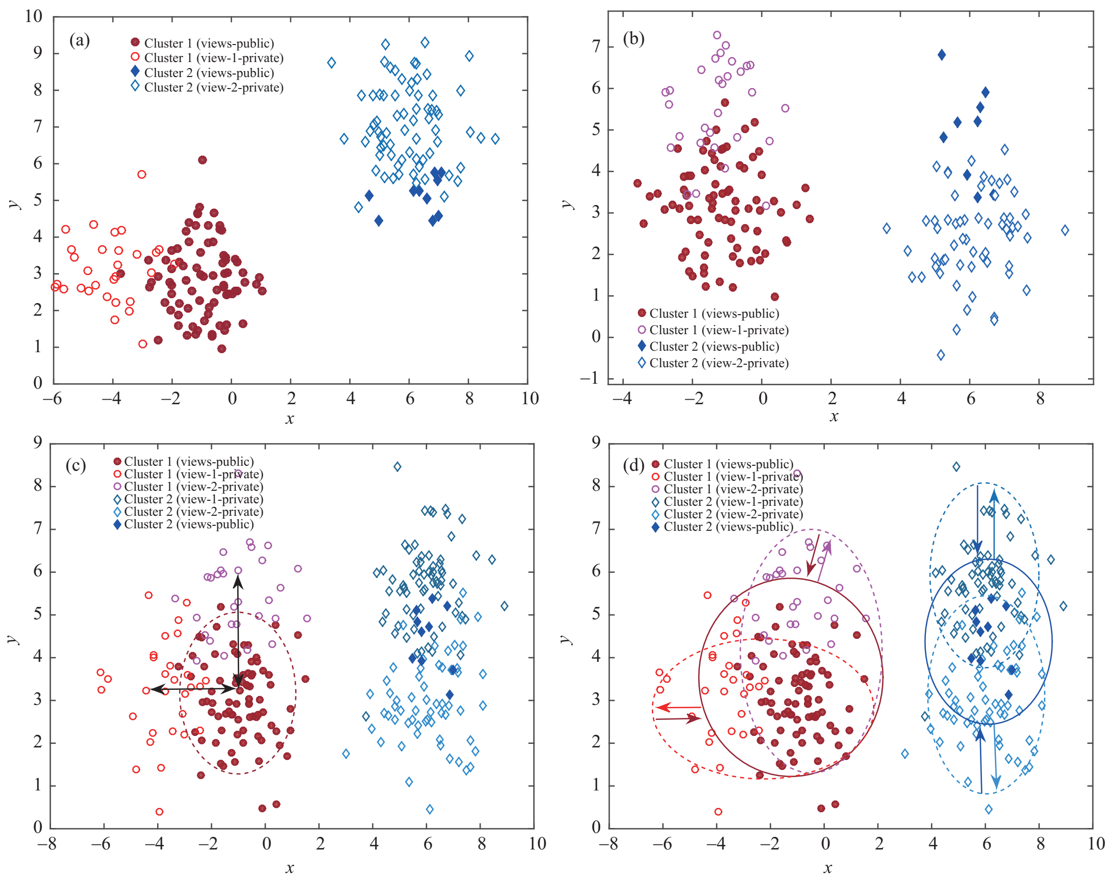

Figure 2 The synthetic dataset is represented by two incomplete views and adopts different methods to partition instances into two clusters. (a) View 1; (b) view 2; (c) existing methods; (d) our method.

recovered views with available views for data clustering. Compared with NMF-based methods, graph-based methods can effectively exploit the relationships among views. However, because some views are suffering from missing instances, these methods face difficulties in constructing a complete graph that connects all the instances and reveals all the local structures.

(3) Multi-kernel-based methods redefine incomplete multi-view clustering as a joint optimization problem. The clustering process and kernel matrix imputation are alternately performed until convergence. The recently proposed multiple kernel k-means with incomplete kernels (MKKM-IK) [34] lays the foundation of incomplete multi-kernels clustering. Different from MKKM-IK, localized incomplete multiple kernel k-means (LI-MKKM) [35] only requires the similarity of a sample to its k-nearest neighbors to align with their ideal similarity values. However, the comparatively intensive computational and storage complexities limit the practical implementation of the algorithms mentioned above. Late fusion incomplete multi-view clustering (LF-IMVC) [36] has been proposed to integrate the incomplete base clustering results generated by incomplete views on the basis of multiple kernel k-means clustering. This method significantly improves the efficiency of clustering. However, a main disadvantage of multi-kernel based methods is that they are sensitive to the quality of the preconstructed kernels. Besides, the filled kernel matrix largely damages the local structure of each incomplete view.

Although the abovementioned methods are effective in some cases and have shown encouraging performances in processing incomplete multi-view data, most of them excessively depend on views-public instances to bridge the connection with view-private instances. Figure 2 shows the synthetic dataset represented by two incomplete views. The goal of incomplete view clustering is to divide all instances from two incomplete views (a) and (b) into two meaningful partitions, i.e., red cluster and blue cluster. If an instance exists only in one view but is absent in another, we regard such an instance as a view-private instance, which is displayed by a hollow point. The instance that exists in both two views can be regarded as a views-public instance, which is displayed by a solid point. It can be seen from Figure 2(c) that the existing methods can effectively transmit the structural information among incomplete views

Table 1 Summary of notations

Notation	Description
n	The number of all the instances
n_v	The number of observed instances in the v-th view
m	The number of views
c	The number of clusters
$\boldsymbol{X}^v \in \mathbb{R}^{n_v \times d_v}$	Data matrix of observed instances in the v-th view
$\boldsymbol{M}^v \in \mathbb{R}^{n_v \times n}$	Indicator matrix of the v-th view
$\boldsymbol{U}^v \in \mathbb{R}^{n_v \times c}$	Local clustering matrix for the v-th view
$\boldsymbol{U}^* \in \mathbb{R}^{n \times c}$	Global clustering matrix for all the instances
$\boldsymbol{S}^v \in \mathbb{R}^{n_v \times n_v}$	Affinity matrix of the v-th view
α^v	Adaptive weight for the v-th view
r and λ	Hyper-parameters

when there are sufficient views-public instances in the views. Except for the blue cluster, when there are few views-public instances, it is difficult for the existing methods to utilize such a rare number of views-public instances to reflect the local structure of each view. In other words, the existing methods cannot effectively transmit structural information among views. In extreme cases, when there are no views-public instances as intermediaries in multiple views, the existing methods fail to address incomplete views. Therefore, we propose an incomplete multi-view clustering algorithm via local and global co-regularization (IMVC-LG) by fully considering the mutual effect between local clustering and global clustering. Although our method includes no obvious views-public information, the global clustering structure can be exploited as potential common information. As shown in Figure 2(d), each dotted line represents the local clustering information of the view, and the solid line represents the global clustering information after integrating different local information. Global clustering results act as latent common information, which can guide local clustering to maximize the agreement on multiple distinct views. Local structural information act as the individual information of each view, which can be utilized to describe the data comprehensively and accurately. Moreover, with the help of an adaptive weighting strategy, the proposed method can effectively balance the importance of different views and reduce the negative impact of unbalanced incomplete views. Based on the above analysis, we define a new objective function that is made up of two terms, i.e., local clustering from each view and global clustering from multiple views, which constrain each other to exploit the local clustering information from different incomplete views and determine a global consensus clustering result, respectively. Furthermore, an iterative optimization method is developed for IMVC-LG to demonstrate its solvability. Finally, comparisons with other state-of-the-art incomplete multi-view clustering methods on benchmark datasets illustrate the effectiveness of our proposed method.

The rest of this paper is organized as follows: Section 2 presents our proposed method and its optimization algorithm in detail. The experimental evaluations and discussions are presented in Section 3. Section 4 presents the conclusions of this study.

2 Methodology

2.1 Problem formulation

Before we describe the method, we formulate the incomplete multi-view clustering problem first. Assume there is a multi-view dataset with incomplete instances in each view. The dataset contains n instances $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n]$, and let $\{\boldsymbol{X}^1, \boldsymbol{X}^2, \ldots, \boldsymbol{X}^m\}$ be the data matrices of m views. $\boldsymbol{X}^v = [\boldsymbol{x}_1^v, \boldsymbol{x}_2^v, \ldots, \boldsymbol{x}_{n_v}^v] \in \mathbb{R}^{n_v \times d_v} (n_v < n)$ represents the data in the v-th view, which contains n_v observed instances with d_v features. Like most existing methods, it is assumed that there is at least one view available for each instance. The purpose of incomplete multi-view clustering is to divide the original n instances into c clusters. The primary notations are summarized in Table 1.

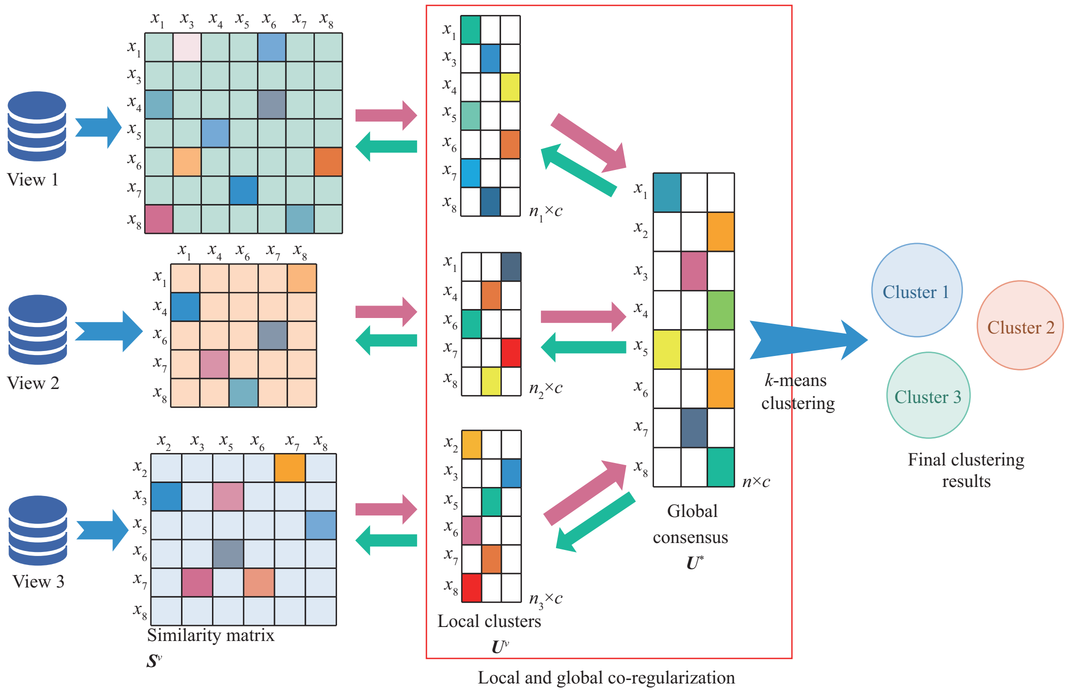

Figure 3 The description of IMVC-LG. First, we consider global structural information as potential common information. Then, we consider the effect of local clustering structures and their divergences on the effectiveness of multi-view clustering. Finally, we utilize the mutual constraints of local clustering and global clustering to obtain the final clustering result.

2.2 The proposed IMVC-LG algorithm

In this subsection, we define a new objective function as follows:

$$\min_{\boldsymbol{U}^*,\{\boldsymbol{U}^v\}_{v=1}^m} \mathcal{J} = \sum_{v=1}^{m} L_{\boldsymbol{U}^v} + G_{\boldsymbol{U}^*}, \tag{1}$$

which is composed of two terms, i.e., local clustering results from each view and global clustering results from multiple views. By fully considering the mutual effect of local clustering and global clustering, the terms are constrained by each other to exploit the local clustering information from different incomplete views and determine a global consensus clustering result, respectively. The general framework of IMVC-LG is shown in Figure 3.

2.2.1 *Local clustering*

In this paper, we focus on spectral clustering because it is a flexible clustering algorithm [37]. Before performing spectral clustering, it is necessary for the proposed method to construct an undirected weighted graph with an affinity matrix $\boldsymbol{S}^v \in \mathbb{R}^{n_v\times n_v}$ for each view. Each entry s_{ij}^v of the symmetric matrix $\boldsymbol{S}^v$ represents the affinity of a pair of instances. Commonly used similarity construction methods include k-nearest neighbor (KNN) method, cosine similarity measure, and Gaussian kernel function. However, each method has its own disadvantages. For example, cosine similarity cannot consider the local geometric structure of the data. KNN similarity and Gaussian kernel similarity are distance-based measures, which are sensitive to noise and outliers in the data. To capture the correlations hidden in the observed multi-view instances, we adopt the idea of adaptive local structure learning of clustering with adaptive neighbors (CAN) [38] on incomplete multiple views. In other words, we aim to construct a similarity matrix in each view such that a smaller distance $\|\boldsymbol{x}_i^v-\boldsymbol{x}_j^v\|_2^2$ between two instances corresponds to a large probabilistic similarity value s_{ij}^v. The function of local clustering for the v-th view is

$$L_{\boldsymbol{U}^v} = \sum_{i,j=1}^{n_v} \left\|\boldsymbol{x}_i^v-\boldsymbol{x}_j^v\right\|_2^2 s_{ij}^v + \sum_{i}^{n_v} \gamma^v \left\|\boldsymbol{s}_i^v\right\|_2^2 + \operatorname{tr}\left(\boldsymbol{U}^{v\mathrm{T}}\boldsymbol{D}^{v-\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{S}^v}\boldsymbol{D}^{v-\frac{1}{2}}\boldsymbol{U}^v\right). \tag{2}$$

The Laplacian matrix $\boldsymbol{L}_{\boldsymbol{S}^v} \in \mathbb{R}^{n_v \times n_v}$ is calculated by $\boldsymbol{L}_{\boldsymbol{S}^v} = \boldsymbol{D}^v - \boldsymbol{S}^v$, where $\boldsymbol{D}^v$ is a diagonal matrix, and its i-th diagonal element is calculated as $d_{ii}^v = \sum_j [(s_{ij}^v + s_{ji}^v)/2]$. γ^v is the regularization parameter, which can be determined as a constant for each view [38]. Let $\boldsymbol{u}_1^v, \boldsymbol{u}_2^v, \ldots, \boldsymbol{u}_c^v$ denote the first c eigenvalues of $\boldsymbol{L}_{\boldsymbol{S}^v}$ corresponding eigenvectors. Local clustering process normalizes all eigenvectors to have unit length and forms the matrix $\boldsymbol{U}^v = [\boldsymbol{u}_1^v, \boldsymbol{u}_2^v, \ldots, \boldsymbol{u}_c^v]$ by stacking the eigenvectors in columns. The spectral embedding $\boldsymbol{U}^v \in \mathbb{R}^{n_v \times c}$ is an approximately continuous-valued of the discrete cluster assignment result, which can be regarded as a recasted matrix of the original feature matrix. Each row of $\boldsymbol{U}^v$ is a new representation of an instance with lower dimensions and contains more discriminative, local cluster information.

It should be noted that the proposed method is essentially different from CAN. The CAN algorithm is a new version of the spectral clustering algorithm, specifically for addressing single view data, and it cannot effectively handle multi-view data. We extend the part of CAN that adaptively constructs samples' similarity to handle incomplete multi-view data. IMVC-LG obtains clustering results based on the standard spectral clustering algorithm. The two algorithms seem to be very similar, but the optimization method and the computational complexity of IMVC-LG are different from those of CAN.

2.2.2 *Global clustering*

To exploit the latent common information, a global clustering function is proposed to integrate consistent information and complementary information among different views. The problem can be formulated as follows:

$$G_{\boldsymbol{U}^*} = \sum_{v=1}^{m} (\alpha^v)^r F(\boldsymbol{U}^v, \boldsymbol{U}^*), \tag{3}$$

where $\boldsymbol{U}^* \in \mathbb{R}^{n \times c}$ is the global clustering indicator matrix for all instances. An index matrix $\boldsymbol{M}^v \in \mathbb{R}^{n_v \times n}$ is introduced to identify the missing instances of each view from original instances, and its (i, j)-th element m_{ij}^v is defined as follows:

$$m_{ij}^v = \begin{cases} 1, & \text{if } \boldsymbol{x}_i^v \text{ corresponds to the } j\text{-th original instance,} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Instead of directly integrating the local cluster results, $\boldsymbol{U}^v$, various regular constraints are added to $\boldsymbol{U}^*$. In IMVC-LG, because the membership of two instances in an individual local view should be relatively consistent in the global space, we believe the pairwise similarities of instances under the local cluster structure should be as similar as possible in the global consistent structure. In IMVC-LG, local clustering result $\boldsymbol{U}^v$ only contains instances observed from each view. Therefore, the index matrix $\boldsymbol{M}^v$ is utilized to select corresponding instances observed from the complete global clustering result $\boldsymbol{U}^*$ for each view. To exploit the local clustering information from different incomplete views and determine a global consensus clustering result, we propose the following cost function by using the inner product of indicator matrices as a measure of discrepancy between two clustering results:

$$F(\boldsymbol{U}^v, \boldsymbol{U}^*) = \frac{1}{2} \left\| \boldsymbol{U}^v \boldsymbol{U}^{v\mathrm{T}} - (\boldsymbol{M}^v \boldsymbol{U}^*)(\boldsymbol{M}^v \boldsymbol{U}^*)^{\mathrm{T}} \right\|_{\mathrm{F}}^2. \tag{5}$$

Because $\boldsymbol{U}^{v\mathrm{T}} \boldsymbol{U}^v = \boldsymbol{I}_c$, $\boldsymbol{U}^{*\mathrm{T}} \boldsymbol{U}^* = \boldsymbol{I}_c$, Eq. (5) can be rewritten as

$$F(\boldsymbol{U}^v, \boldsymbol{U}^*) = c - \operatorname{tr}\left(\boldsymbol{U}^v \boldsymbol{U}^{v\mathrm{T}} \left(\boldsymbol{M}^v \boldsymbol{U}^*\right)\right) (\boldsymbol{M}^v \boldsymbol{U}^*)^{\mathrm{T}}. \tag{6}$$

In real-world applications of multi-view data, different views generally have different physical meanings and discriminant information. Specifically, for incomplete multi-view cases, the discriminative information of different views has huge differences due to the different feature dimensions and numbers of available instances. Therefore, it is unreasonable to treat all incomplete views equally during global clustering. To this end, α^v is utilized to balance the significance of the v-th view. α^v is a positive weight, and the parameter $r > 1$ is used to control the smoothness of the distribution of the weights. In this way, the method can adaptively identify the complementary effects of different local discrimination for model training.

Subsequently, we need to minimize the sum of dissimilarities between all local cluster indicator matrices and global consensus matrix, to obtain the final clustering results.

2.2.3 *Unified objective function of IMVC-LG*

To preserve the semantic correlations between the above two components, the objective function integrates the above functions, i.e., Eqs. (2) and (3) into one unified incomplete multi-view clustering model as follows:

$$\begin{aligned}\min_{\boldsymbol{U}^*,\{\boldsymbol{U}^v\}_{v=1}^m} \sum_{v=1}^{m}\sum_{i,j=1}^{n_v} \Big\{ \left\|\boldsymbol{x}_i^v-\boldsymbol{x}_j^v\right\|_2^2 s_{ij}^v + \gamma^v \left\|\boldsymbol{s}_i^v\right\|_2^2 + \lambda \operatorname{tr}\left(\boldsymbol{U}^{v\mathrm{T}}\boldsymbol{D}^{v-\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{S}^v}\boldsymbol{D}^{v-\frac{1}{2}}\boldsymbol{U}^v\right) \\ + (\alpha^v)^r \left(c-\operatorname{tr}\left(\boldsymbol{U}^v\boldsymbol{U}^{v\mathrm{T}}(\boldsymbol{M}^v\boldsymbol{U}^*)(\boldsymbol{M}^v\boldsymbol{U}^*)^{\mathrm{T}}\right)\right)\Big\} \\ \text{s.t. } \forall v,i,\ s_{ii}^v=0,\ s_{ij}^v \geqslant 0,\ \mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1,\ \sum_{v=1}^{m}\alpha^v=1,\ \alpha^v\geqslant 0,\ \boldsymbol{U}^{v\mathrm{T}}\boldsymbol{U}^v=\boldsymbol{I}_c,\ \boldsymbol{U}^{*\mathrm{T}}\boldsymbol{U}^*=\boldsymbol{I}_c.\end{aligned} \tag{7}$$

2.3 Optimization of IMVC-LG

Next, an iterative optimization method is proposed to minimize the objective function, i.e., Eq. (7). IMVC-LG optimizes the objective function with respect to one variable while fixing others until it converges. Similarly, IMVC-LG divides the overall optimization into two processes: the optimization of local clustering and the optimization of global clustering.

2.3.1 *Optimization of local clustering*

By fixing all the other variables, optimizing $\boldsymbol{U}^v$ for the v-th view amounts to minimizing the following function:

$$\begin{aligned}&\min_{\boldsymbol{U}^v} \operatorname{tr}\left(\boldsymbol{U}^{v\mathrm{T}}\boldsymbol{L}^v\boldsymbol{U}^v\right) \\ &\text{s.t. } \boldsymbol{U}^{v\mathrm{T}}\boldsymbol{U}^v=\boldsymbol{I}_c.\end{aligned} \tag{8}$$

Eq. (8) is equivalent to solving the standard spectral clustering objective with a modified Laplacian matrix $\boldsymbol{L}^v=\lambda\boldsymbol{D}^{v-\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{S}^v}\boldsymbol{D}^{v-\frac{1}{2}}-(\alpha^v)^r(\boldsymbol{M}^v\boldsymbol{U}^*\boldsymbol{U}^{*\mathrm{T}}\boldsymbol{M}^{v\mathrm{T}})$. The optimal solution of (8) is the eigenvectors corresponding to the c smallest eigenvalues of $\boldsymbol{L}^v$. IMVC-LG normalizes all the eigenvectors to have a unit length and forms the matrix $\boldsymbol{U}^v$ by stacking the eigenvectors in columns.

$\boldsymbol{S}^v$ can be calculated for each view independently because the similarity matrices do not have relationships among each other. Let $d_{x_{ij}}^v=\|\boldsymbol{x}_i^v-\boldsymbol{x}_j^v\|_2^2$, $d_{u_{ij}}^v=\|\frac{\boldsymbol{u}_i^v}{\sqrt{d_{ii}^v}}-\frac{\boldsymbol{u}_j^v}{\sqrt{d_{jj}^v}}\|_2^2$, and $\boldsymbol{d}_i^v$ denotes as a vector with the j-th element as $d_{ij}^v=d_{x_{ij}}^v+\lambda d_{u_{ij}}^v$.

$$\begin{aligned}&\min_{\boldsymbol{s}_i^v}\left\|\boldsymbol{s}_i^v+\frac{1}{2\gamma^v}\boldsymbol{d}_i^v\right\|_2^2 \\ &\text{s.t. } s_{ii}^v=0,\ s_{ij}^v\geqslant 0,\ \mathbf{1}^{\mathrm{T}}\boldsymbol{s}_i^v=1.\end{aligned} \tag{9}$$

The closed-form solution of (9) can be achieved by an efficient algorithm presented in [38].

2.3.2 *Optimization of global clustering*

By fixing all the other variables, optimizing $\boldsymbol{U}^*$ amounts to minimizing the following function:

$$\begin{aligned}&\max_{\boldsymbol{U}^*} \operatorname{tr}\left(\boldsymbol{U}^{*\mathrm{T}}\boldsymbol{L}^*\boldsymbol{U}^*\right) \\ &\text{s.t. } \boldsymbol{U}^{*\mathrm{T}}\boldsymbol{U}^*=\boldsymbol{I}_c.\end{aligned} \tag{10}$$

Eq. (10) is equivalent to solving the standard spectral clustering objective with a modified Laplacian matrix $\boldsymbol{L}^*=\sum_{v=1}^m(\alpha^v)^r(\boldsymbol{M}^{v\mathrm{T}}\boldsymbol{U}^v\boldsymbol{U}^{v\mathrm{T}}\boldsymbol{M}^v)$. The optimal solution of (10) is the eigenvectors corresponding to the c largest eigenvalues of $\boldsymbol{L}^*$. IMVC-LG normalizes all eigenvectors to have a unit length and forms the matrix $\boldsymbol{U}^*$ by stacking the eigenvectors in columns.

By fixing all the other variables, α^v acts as an adaptive weight for each view, whose optimization process can be simplified into the following problem:

$$\begin{aligned} &\min_{\alpha^v} \sum_{v=1}^{m} (\alpha^v)^r F(\boldsymbol{U}^v, \boldsymbol{U}^*) \\ &\text{s.t.} \ \sum_{v=1}^{m} \alpha^v = 1, \ \alpha^v \geqslant 0, \end{aligned} \tag{11}$$

where $r > 1$, the closed-form solution of (11) can be rewritten as follows:

$$\alpha^v = \frac{(F(\boldsymbol{U}^v, \boldsymbol{U}^*))^{1/(1-r)}}{\sum_{v=1}^{m} (F(\boldsymbol{U}^v, \boldsymbol{U}^*))^{1/(1-r)}}. \tag{12}$$

Algorithm 1 briefly summarizes the iterative optimization procedures of the proposed method.

Algorithm 1 IMVC-LG

Input: Dataset $\{\boldsymbol{X}^1, \boldsymbol{X}^2, \ldots, \boldsymbol{X}^m\}$ with m incomplete views, the number of clusters c, hyperparameters λ and r.
Output: Indicator matrix of final clustering $\boldsymbol{U}^* \in \mathbb{R}^{n \times c}$.

```
Initialize:
Utilizing (9) to generate the similarity matrix for each view by d^v_ij = d^v_x_ij;
Initialize the weight for each view α^v = 1/m;
Initialize U^v by solving standard spectral clustering algorithm;
Initialize U* by solving (10);
repeat
    Update U^v by solving (8);
    Update S^v by solving (9);
    Update U* by solving (10);
    Update α^v by using (12);
until converges;
Apply k-means to U* to obtain the final clustering results.
```

3 Experiments

In this section, we compare the proposed method (IMVC-LG) with other state-of-the-art incomplete multi-view clustering methods on benchmark datasets to illustrate the effectiveness of the proposed method. Then, we analyze the impact of parameters and present the results about convergence behavior.

3.1 Datasets

Three complete multi-view datasets and a natural incomplete multi-view dataset were used for evaluation: Newsgroups dataset[1)] (NGs), BBC dataset[2)] (BBC), handwritten digit dataset[3)] (HW), and three sources dataset[4)] (3Sources). The details are summarized in Table 2. It is worth emphasizing that the first three datasets are originally complete; in other words, each instance is sampled in all the views. The 3Sources dataset includes naturally incomplete multi-view data. The missing rates of three views in the 3Sources dataset are 15.38%, 29.33%, and 27.40%.

Multi-view data include scaling differences among different views. To eliminate these differences and improve clustering performance, it is necessary to normalize the dataset. If the comparative experiment does not clearly declare the method of normalization, each instance is normalized to unit L_2 norm in this paper, such as [27].

3.2 Baseline methods

We compare the performances of the following methods.

1) http://lig-membres.imag.fr/grimal/data.html.
2) http://mlg.ucd.ie/datasets/segment.html.
3) http://archive.ics.uci.edu/ml/datasets/Multiple+Features.
4) http://mlg.ucd.ie/datasets/3sources.html.

Table 2 Statistics of the datasets

Dataset	NGs	BBC	HW	3Sources
# d_1	2000	4659	216	3560
# d_2	2000	4633	76	3631
# d_3	2000	4665	64	3068
# d_4	–	4684	6	–
# d_5	–	–	240	–
# d_6	–	–	47	–
# view	3	4	6	3
# instances	500	685	2000	416
# class	5	5	10	6

• Single view spectral clustering. Best single view (BSV) [5] is a local clustering method that uses the most informative view. Notably, the informative view has the best separate spectral clustering performance among all multiple views. Hence, we ran standard spectral clustering on each incomplete view and reported the best clustering result.

• Incomplete two-view clustering based on NMF. PVC [24] is a traditional, pioneered work dealing with partial multi-view setting. PVC seeks to find a latent common subspace for both public and private instances; therefore, it is a views-public instances based method; IMG [23] is a method following the idea of PVC. IMG explores a global structure of partial multi-view data by utilizing a graph Laplacian term. Incomplete multi-view clustering via graph regularized matrix factorization (IMC-GRMF) [39] is an improved method of PVC. Similar to PVC, the matrix factorization technique is exploited in IMC-GRMF to learn common latent representation, in which the representations corresponding to those samples with all views are regularized to be consistent. These three methods mentioned above only work for two-view data; therefore, we evaluated all two-view combinations and reported the best results.

• Incomplete multi-view clustering based on NMF. MIC [22] first fills the missing instances in each incomplete view with the average features, then extends multi-NMF via weighted NMF with $L_{2,1}$ regularization. Doubly aligned incomplete multi-view clustering (DAIMC) [40] applies weighted semi-NMF to learn a common representation for all views, and then performs k-means on the learned representation to obtain the clustering results.

• Incomplete multi-view clustering based on graph-based method. GIMC [25] is the first attempt in constructing a complete graph for partitioning incomplete multi-view data. GIMC constructs a graph for each view, and then integrates all the constructed graph matrices to generate a consensus graph matrix. IMSC-AGL [32] integrates the graph construction and consensus representation learning into a joint optimization framework. It is the first work that exploits the graph learning and spectral clustering techniques to learn the common representation for incomplete multi-view clustering. AGCIMC [30] develops a joint framework for graph completion and consensus representation learning. GIMC-FLSD [31] proposes a graph-regularized framework for structure preserving and representation learning.

• Incomplete multi-view clustering based on multi-kernel-based method. LF-IMVC [36] is a late fusion incomplete multi-view clustering method that effectively and efficiently integrates the incomplete clustering matrices generated by incomplete views. Essentially, LF-IMVC is an ensemble method based on the multiple kernels k-means clustering.

3.3 Experimental setting

To mimic the incomplete multi-view setting, we randomly removed some instances from each view. per%(per $\in \{10, 20, 30, 40, 50\}$) instances were randomly removed from every view to construct the incomplete data with a missing ratio of per%. We ensured that each instance had at least one view remaining.

Experiments were conducted on an Intel i7-7700 CPU personal computer with 48 GB RAM. Moreover, we used the k-means algorithm to achieve the final clustering results. Since k-means is known to be sensitive to initialization, to reduce the effect of the randomness caused by k-means, we repeated the clustering process 20 times with random initialization and reported the average values. We employed the widely used external indices clustering accuracy (ACC), normalized mutual information (NMI), and purity to evaluate clustering performances. We expected the values of these evaluation criterions to be as large as possible.

Table 3 Clustering results (ACC and NMI) of different methods on the NGs dataset with different per%

Method	ACC					NMI				
	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
BSV	0.5712	0.5425	0.4930	0.4120	0.3700	0.3428	0.3227	0.3126	0.2117	0.1722
PVC	0.7134	0.6624	0.6354	0.6230	0.5686	0.5252	0.4440	0.3945	0.3812	0.3634
IMG	0.6904	0.6475	0.6022	0.5828	0.5793	0.4898	0.4588	0.4126	0.3846	0.3440
IMC-GRMF	0.8203	0.7721	0.7563	0.6986	0.6333	0.6172	0.5388	0.5151	0.4438	0.3898
MIC	0.8980	0.8280	0.5740	0.5200	0.4240	0.7528	0.6009	0.3849	0.2824	0.1424
LF-IMVC	0.8480	0.7840	0.6440	0.6100	0.6200	0.6522	0.5404	0.4205	0.3660	0.3823
DAIMC	0.9300	0.8140	0.7560	0.6880	0.5800	0.8091	0.6008	0.5971	0.4474	0.3154
IMSC-AGL	0.9500	0.9240	0.8120	0.8040	0.7700	0.8564	0.7857	0.5741	0.5573	0.4823
GIMC	0.7840	0.7720	0.7480	0.5920	0.5860	0.8222	0.7877	0.7129	0.6229	0.6116
AGCIMC	0.9480	0.8980	0.7001	0.6620	0.5700	0.8323	0.7661	0.5566	0.5404	0.3580
GIMC-FLSD	0.9100	0.8680	0.8180	0.7310	0.7120	0.7797	0.7196	0.6252	0.4404	0.4326
IMVC-LG	**0.9640**	**0.9412**	**0.8960**	**0.8344**	**0.7740**	**0.8884**	**0.8218**	**0.7215**	**0.5921**	**0.5154**

Table 4 Clustering results (ACC and NMI) of different methods on the BBC dataset with different per%

Method	ACC					NMI				
	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
BSV	0.6058	0.5051	0.4307	0.3956	0.3518	0.3799	0.2693	0.2230	0.2030	0.1662
PVC	0.6453	0.6345	0.6322	0.5707	0.5625	0.5104	0.4804	0.4890	0.4483	0.4226
IMG	0.6553	0.6394	0.6083	0.5799	0.5575	0.5021	0.4852	0.4194	0.4099	0.3775
IMC-GRMF	0.7915	0.7688	0.7601	0.6809	0.6395	0.5822	0.5550	0.5575	0.4889	0.4430
MIC	0.7810	0.6745	0.6365	0.6204	0.5051	0.6030	0.5452	0.4347	0.4080	0.2374
LF-IMVC	0.6993	0.6321	0.5577	0.5358	0.5299	0.5138	0.4710	0.4042	0.2923	0.3078
DAIMC	0.8055	0.7524	0.5956	0.5766	0.5561	0.6293	0.5925	0.5021	0.4468	0.4001
IMSC-AGL	0.9080	0.8686	0.8350	0.8161	0.8044	0.7628	0.6921	0.6432	0.5968	0.5713
GIMC	0.5635	0.5577	0.5197	0.5139	0.5095	0.6130	0.6590	0.6375	0.6162	0.5693
AGCIMC	0.8710	0.8554	0.8496	0.7708	0.7693	0.7155	0.6784	0.6551	0.5326	0.5152
GIMC-FLSD	0.7752	0.7623	0.7533	0.7322	0.7044	0.6064	0.5964	0.5722	0.5423	0.5147
IMVC-LG	**0.9168**	**0.8905**	**0.8701**	**0.8453**	**0.8146**	**0.7716**	**0.7291**	**0.6916**	**0.6507**	**0.5936**

Table 5 Clustering results (ACC and NMI) of different methods on the HW dataset with different per%

Method	ACC					NMI				
	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
BSV	0.5508	0.5046	0.4575	0.4363	0.3835	0.5157	0.4890	0.4289	0.3844	0.3243
PVC	0.5954	0.5745	0.5525	0.5320	0.5220	0.5410	0.5121	0.4805	0.4613	0.4485
IMG	0.6337	0.6030	0.5681	0.5428	0.5239	0.5729	0.5332	0.5118	0.4781	0.4612
IMC-GRMF	0.7301	0.6628	0.6221	0.5825	0.5470	0.6780	0.6104	0.5742	0.5359	0.4999
MIC	0.7390	0.6653	0.6014	0.5246	0.5123	0.6481	0.5827	0.5360	0.4835	0.4481
LF-IMVC	0.8955	0.8855	0.8805	0.8065	0.7220	0.8048	0.7877	0.7794	0.6779	0.6256
DAIMC	0.8778	0.8595	0.8855.	0.7782	0.7112	0.7826	0.7716	0.7548	0.6576	0.5815
IMSC-AGL	0.9575	0.9560	0.9394	0.9116	0.8977	0.9170	0.9108	0.8894	0.8601	0.8134
GIMC	0.8915	0.8830	0.8755	0.7790	0.5850	0.8978	0.8875	0.8881	0.8647	0.7785
AGCIMC	0.8671	0.8580	0.8555	0.8310	0.8175	0.9035	0.8938	0.8790	0.8595	0.8510
GIMC-FLSD	0.8752	0.8694	0.8101	0.7362	0.6021	0.7965	0.7746	0.6960	0.6048	0.4874
IMVC-LG	**0.9615**	**0.9586**	**0.9430**	**0.9280**	**0.9025**	**0.9266**	**0.9136**	**0.9003**	**0.8638**	**0.8244**

3.4 Experimental results

Tables 3∼6 and Figure 4 show the performance comparisons on the datasets. We can draw the following observations and discussions.

(1) In most cases, BSV performed the worst in comparison with the other methods. From another perspective, BSV can be viewed as a type of local clustering algorithm. Due to the lack of complementary information from other views, local clustering results are unsatisfactory. With a fixed ratio of incomplete

Table 6 Clustering results (ACC, NMI and Purity) of different methods on the 3Sources dataset

Method	ACC	NMI	Purity
BSV	0.5409	0.3837	0.5649
PVC	0.5725	0.5089	0.6643
IMG	0.5582	0.5018	0.6104
IMC-GRMF	0.8442	0.6728	0.8674
MIC	0.6514	0.6192	0.7428
LF-IMVC	0.6346	0.5418	0.7067
DAIMC	0.8630	0.7113	0.8730
IMSC-AGL	0.8139	0.6901	0.8239
GIMC	0.7332	0.5024	0.6283
AGCIMC	0.7788	0.6836	0.7981
GIMC-FLSD	0.8654	0.7067	0.8774
IMVC-LG	**0.8822**	**0.7544**	**0.8822**

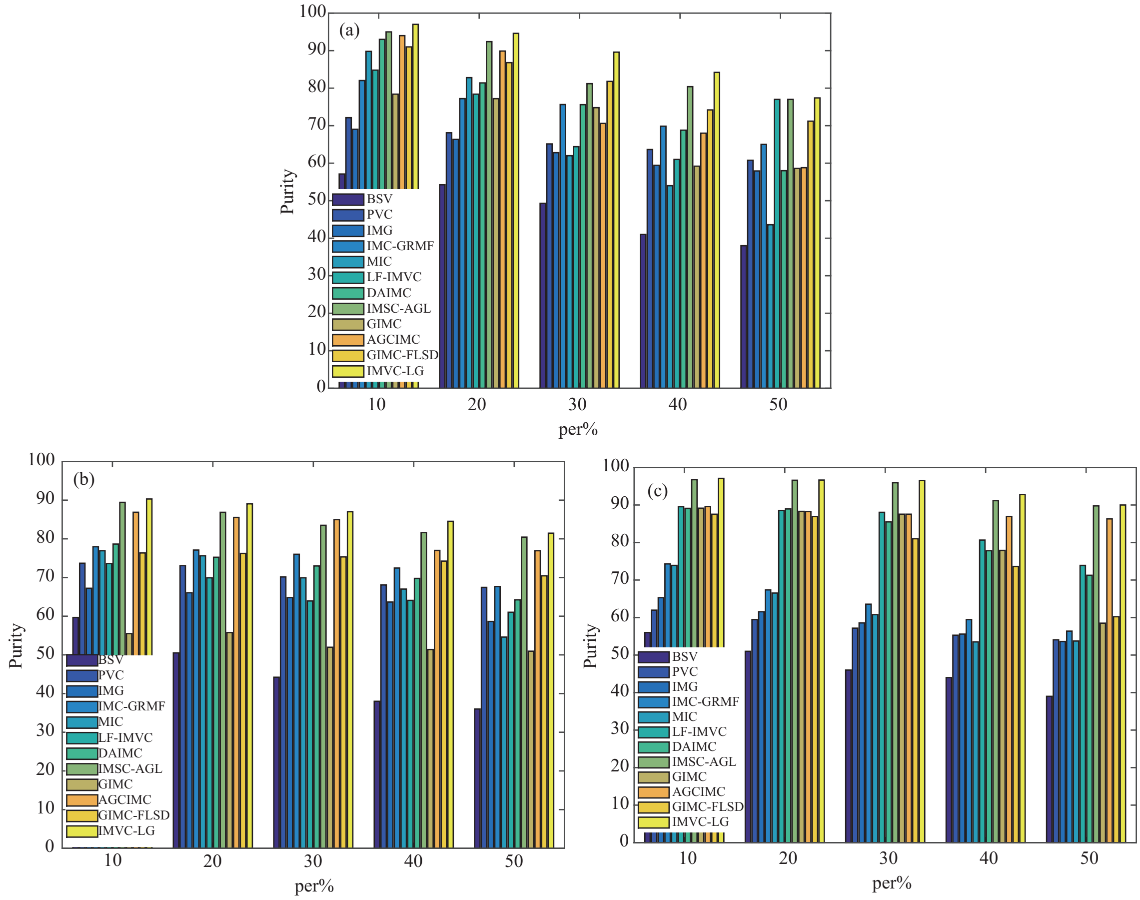

Figure 4 The performances of purity on the (a) NGs, (b) BBC, and (c) HW datasets with different per%.

instances, the performances of incomplete multi-view clustering methods are generally better than that of two-view methods. It can be concluded that exploiting the complementary information of multi-view is an effective approach in dealing with the incomplete problem, and the more views that are utilized, the more satisfying the results will be.

(2) When the missing ratio increased, the clustering performances of all the clustering methods were worse in most cases, which is consistent with intuition. MIC filled the missing instances with the average of values; therefore, the performance of MIC degenerated faster than the other methods. Based on this result, it can be concluded that filling in the missing views with the average of values may not be a useful approach. The features in both inter-class and intra-class may have large variances, and the same assignment will weaken the discrimination among them, especially when the missing rate was large. When

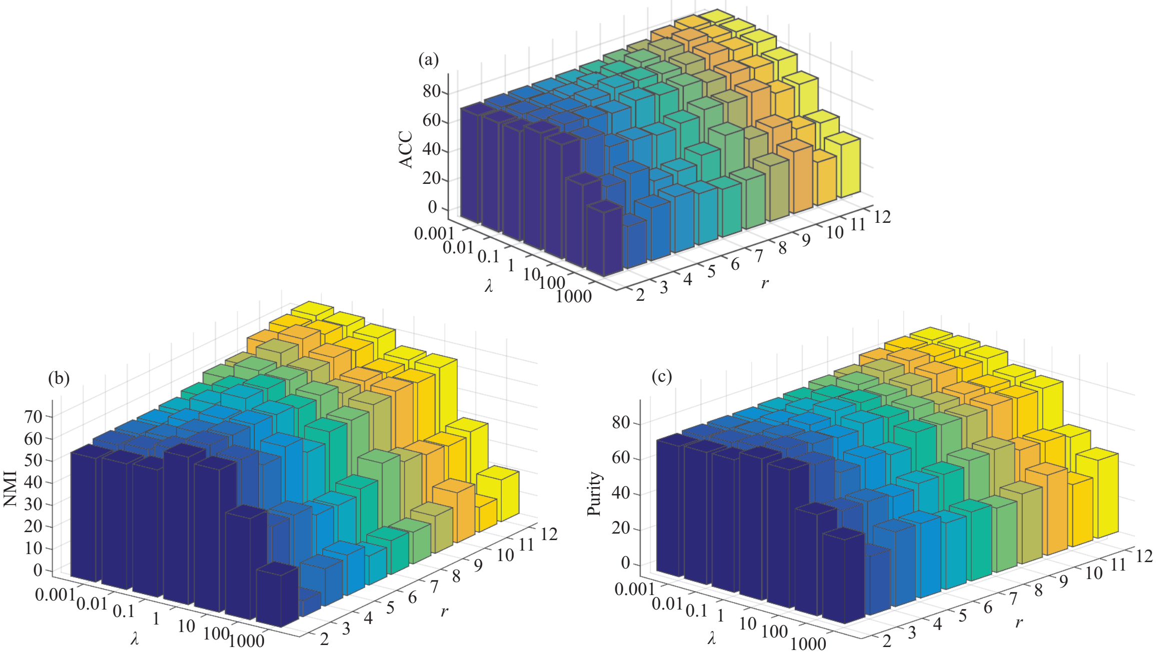

Figure 5 (a) ACC, (b) NMI, and (c) purity versus parameters λ and r of the proposed method on 3Sources dataset.

the missing ratio was relatively large, and the number of views-public instances was relatively small, PVC and its versions performed worse than other methods. We can conclude that views-public based methods only work when the view missing ratio is relatively low. IMVC-LG aimed at utilizing complementary information from multiple views to exploit consistent public structural information among multiple views; therefore, it is minimally affected by the large missing ratio.

(3) In most cases, NMF-based methods were inferior to graph-based and multi-kernel-based methods because that the NMF-based methods are not efficient on nonlinear data. Nevertheless, graph-based and multi-kernel-based clustering algorithms were sensitive to the initial graphs and kernels. IMVC-LG attained better results than them on most datasets. The reason is that IMVC-LG not only considers the local cluster structures of different views but the similarity matrices are also updated in subsequent iterations to avoid the sensitivity of the initial graphs.

(4) The proposed method significantly outperformed the other methods on the above multi-view databases with arbitrary incomplete cases. For instance, the experimental results of our method approximately improved by 20% to 30% on the BBC and NGs datasets in comparison with other methods. On the HW dataset, the clustering performance of our method was the same as that of the latest algorithm, IMSC-AGL. The reason is that our approach not only considers the importance of different views but also exploits the local structure information of incomplete views to learn the global consensus clustering result. Therefore, the promising performance demonstrates the indispensability of two components and further illustrates the effectiveness of our algorithm.

Overall, it can be concluded that IMVC-LG is superior among the compared algorithms on the datasets with incomplete multiple views.

3.5 Parameters and convergence analysis

Experiments were conducted to analyze the sensitivity of the penalty parameters (λ and r). To the best of our knowledge, setting the penalty parameters adaptively for different datasets without manual intervention remains an open problem. In our work, we used the grid searching approach to find the optimal combinations for different datasets. λ was tuned in the range of $\{0.001, 0.01, 0.1, 1, 10, 100, 1000\}$, while r was tuned in the range of $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. As shown in Figures 5(a)–(c), the 3Sources dataset achieved the best performance with the combinations of $\{\lambda = 0.01, r = 10\}$. From Figures 6(a)–(c), it can be observed that the proposed method obtained the best performance when the two parameters were selected with the combinations of $\{\lambda = 0.1, r = 4\}$ on the HW dataset with 30%

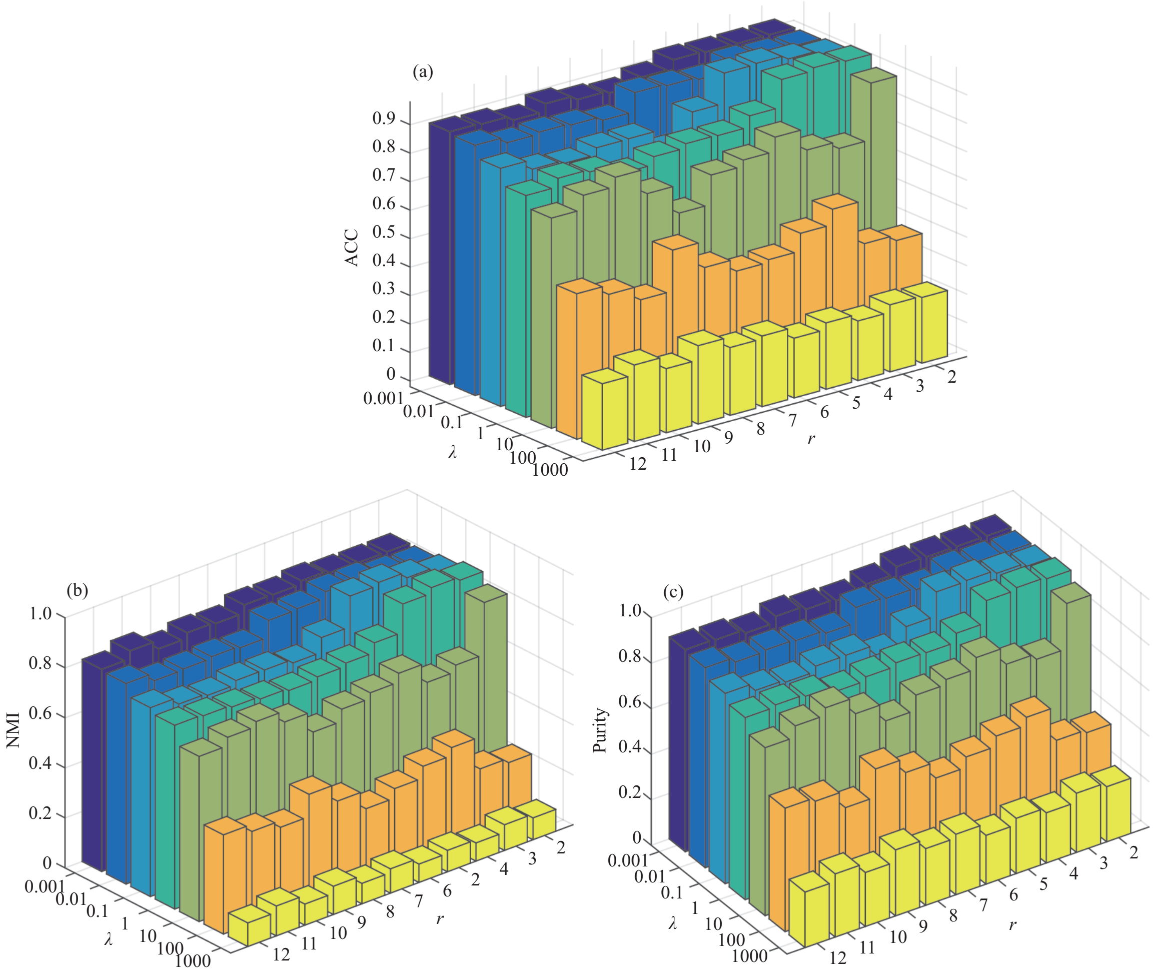

Figure 6 (a) ACC, (b) NMI, and (c) purity versus parameters λ and r of the proposed method on HW dataset with 30% random missing instances of every view.

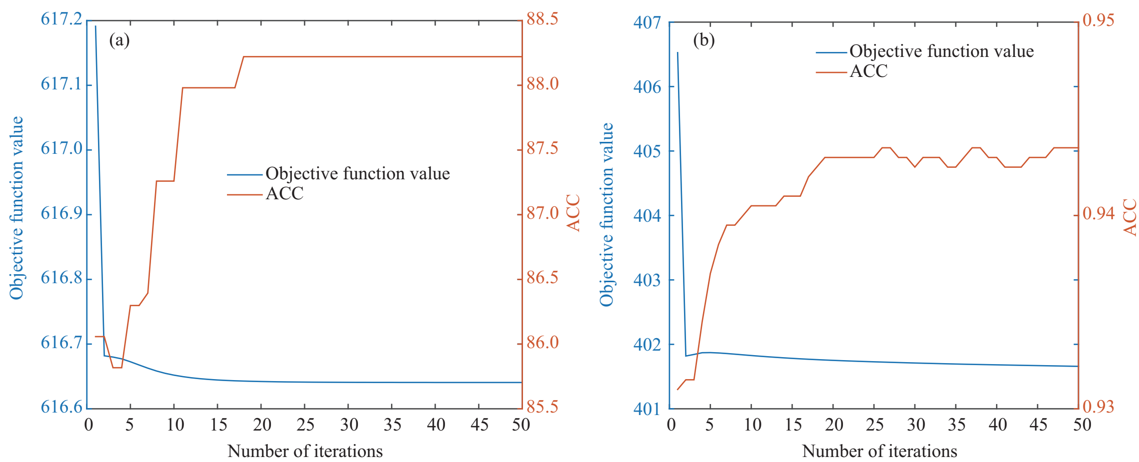

Figure 7 (a) The convergence curve of the proposed method on 3Sources dataset; (b) the convergence curve of the proposed method on HW dataset with 30% random missing instances of every view.

random missing instances of every view. Other datasets have different optimal parameters because their data characteristics are different.

The value of the objective function and ACC versus iteration on two datasets are presented in Figures 7(a) and (b). It is obvious that the objective function curve is monotonically decreasing and the

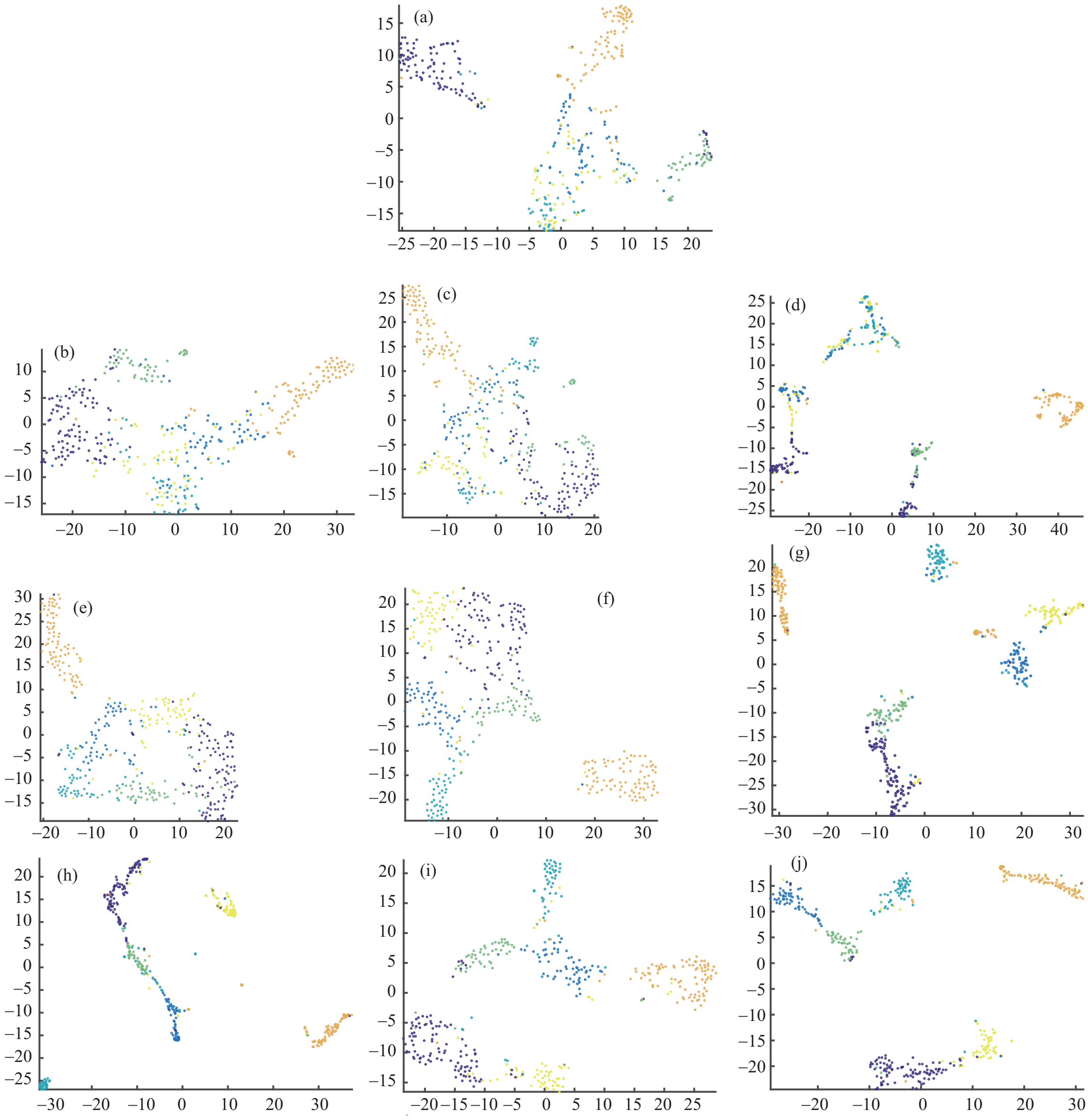

Figure 8 Visualization clustering results on 3Sources dataset. (a) PVC; (b) IMG; (c) IMC-GRMF; (d) MIC; (e) LF-IMVC; (f) DAIMC; (g) IMSC-AGL; (h) AGCIMC; (i) GIMC-FLSD; (j) IMVC-LG.

ACC curve is increasing until the stable level, which demonstrated that our provided approach is efficient for solving the optimization problem. It can be seen that IMVC-LG will converge after approximately 30 iterations.

3.6 Visualization of clustering results

To illustrate the experimental results more intuitively, we visualize the results of different incomplete multi-view clustering algorithms with the t-SNE [41] method in Figure 8. Taking the PVC method as an example, we performed PVC on the first and second views, and it can be seen from the visualization results that the PVC method cannot separate the blue and yellow clusters well. Besides, when there a large number of views, designing the combination between views is a problem. The MIC method filled in the missing views with the average of values, resulting in the yellow cluster and green cluster being indistinguishable. The results show that the same assignment weakened the discrimination among samples. In most cases, it can be observed that the clustering partitions of graph-based methods are much more obvious than those of NMF-based methods. Figure 8(j) shows the best separability for different clusters, where instances from the same cluster were naturally gathered together, and the gap between different groups is obvious. This demonstrates the effectiveness of our method for incomplete multi-view clustering tasks.

4 Conclusion

In this study, we propose an incomplete multi-view clustering algorithm that simultaneously considers the local information and global information of all the views. The learning processes of the two kinds of information are constrained by each other by integrating two significant principles, i.e., consensus and complementary principles. Considering the mutual effect of local and global clustering via co-regularization, the proposed method can eliminate the clustering divergence between multiple views. Furthermore, an iterative optimization method was developed for IMVC-LG to demonstrate its solvability. Finally, comparisons with other state-of-the-art incomplete multi-view clustering methods on benchmark datasets illustrated the effectiveness of our proposed method. In the future, our research focus will be on using other structural information of the views to reduce the impact of incomplete views on the clustering results.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2020AAA0106100) and National Natural Science Foundation of China (Grant Nos. 62022052, 62006147).

References

1 von Luxburg U. A tutorial on spectral clustering. Stat Comput, 2007, 17: 395–416
2 Rodriguez A, Laio A. Clustering by fast search and find of density peaks. Science, 2014, 344: 1492–1496
3 Yue S H, Wang J S, Tao G, et al. An unsupervised grid-based approach for clustering analysis. Sci China Inf Sci, 2010, 53: 1345–1357
4 Li C Z, Xu Z B, Qiao C, et al. Hierarchical clustering driven by cognitive features. Sci China Inf Sci, 2014, 57: 012109
5 Li Z W, Cheong L F, Yang S G, et al. Simultaneous clustering and model selection: algorithm, theory and applications. IEEE Trans Pattern Anal Mach Intell, 2018, 40: 1964–1978
6 Wang Y S, Fang H Y, Yang D J, et al. Network clustering analysis using mixture exponential-family random graph models and its application in genetic interaction data. IEEE ACM Trans Comput Biol Bioinf, 2019, 16: 1743–1752
7 Oyelade J, Isewon I, Oladipupo O, et al. Data clustering: algorithms and its applications. In: Proceedings of the 19th International Conference on Computational Science and Its Applications (ICCSA), Saint Petersburg, 2019. 71–81
8 Wang J L, Lu Y H, Liu J B, et al. A robust three-stage approach to large-scale urban scene recognition. Sci China Inf Sci, 2017, 60: 103101
9 Bai L, Liang J Y, Cao F Y. Semi-supervised clustering with constraints of different types from multiple information sources. IEEE Trans Pattern Anal Mach Intell, 2021, 43: 3247–3258
10 Zhang C Q, Fu H Z, Hu Q H, et al. Generalized latent multi-view subspace clustering. IEEE Trans Pattern Anal Mach Intell, 2020, 42: 86–99
11 Liu J, Cao F Y, Gao X Z, et al. A cluster-weighted kernel k-means method for multi-view clustering. In: Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI), New York, 2020. 4860–4867
12 Wang H, Yang Y, Liu B. GMC: graph-based multi-view clustering. IEEE Trans Knowl Data Eng, 2020, 32: 1116–1129
13 Lu Z M, Liu C, Zhang Q, et al. Visual analytics for the clustering capability of data. Sci China Inf Sci, 2013, 56: 052110
14 Xu X M, Li K K, Xu C, et al. GDFace: gated deformation for multi-view face image synthesis. In: Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI), New York, 2020. 12532–12540
15 Fei H L, Li P. Cross-lingual unsupervised sentiment classification with multi-view transfer learning. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), 2020. 5759–5771
16 Jiang X P, Hu X H, He T T. Identification of the clustering structure in microbiome data by density clustering on the Manhattan distance. Sci China Inf Sci, 2016, 59: 070104
17 Wang M H, Lin Y J, Yang K P, et at. M2GRL: a multi-task multi-view graph representation learning framework for web-scale recommender systems. In: Proceedings of the 26th ACM Conference on Knowledge Discovery and Data Mining (SIGKDD), 2020. 2349–2358
18 Zhang X N, Song S J, Zhu L, et al. Unsupervised learning of Dirichlet process mixture models with missing data. Sci China Inf Sci, 2016, 59: 012201
19 Zhang C Q, Cui Y J, Han Z B, et al. Deep partial multi-view learning. IEEE Trans Pattern Anal Mach Intell, 2020. doi: 10.1109/TPAMI.2020.3037734
20 Sun S L. A survey of multi-view machine learning. Neural Comput Appl, 2013, 23: 2031–2038
21 Yang Y, Wang H. Multi-view clustering: a survey. Big Data Min Anal, 2018, 1: 83–107
22 Li S Y, Jiang Y, Zhou Z H. Partial multi-view clustering. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI), Québec, 2014. 1968–1974
23 Zhao H D, Liu H F, Fu Y. Incomplete multi-modal visual data grouping. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI), New York, 2016. 2392–2398
24 Shao W X, He L f, Philip S Y. Multiple incomplete views clustering via weighted nonnegative matrix factorization with $L_{2,1}$ regularization. In: Proceedigns of Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML & PKDD), Porto, 2015. 318–334
25 Zhou W, Wang H, Yang Y. Consensus graph learning for incomplete multi-view clustering. In: Proceedings of the 23rd Pacific-Asia Advances in Knowledge Discovery and Data Mining Conference (PAKDD), Macau, 2019. 529–540
26 Min C, Cheng M M, Yu J, et al. Partial multi-view clustering via auto-weighting similarity completion. In: Proceedings of the 13th Chinese Conference on Biometric Recognition (CCBR), Urumqi, 2018. 214–222
27 Guo J, Ye J H. Anchors bring ease: an embarrassingly simple approach to partial multi-view clustering. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI), Hawaii, 2019. 118–125

28 Wu J, Zhuge W Z, Tao H, et al. Incomplete multi-view clustering via structured graph learning. In: Proceeding of the 15th Pacific Rim International Conference on Artificial Intelligence (PRICAI), Nanjing, 2018. 98–112
29 Yang L, Shen C Y, Hu Q H, et al. Adaptive sample-level graph combination for partial multiview clustering. IEEE Trans Image Process, 2020, 29: 2780–2794
30 Wen J, Yan K, Zhang Z, et al. Adaptive graph completion based incomplete multi-view clustering. IEEE Trans Multimedia, 2021, 23: 2493–2504
31 Wen J, Zhang Z, Zhang Z, et al. Generalized incomplete multiview clustering with flexible locality structure diffusion. IEEE Trans Cybern, 2021, 51: 101–114
32 Wen J, Xu Y, Liu H. Incomplete multiview spectral clustering with adaptive graph learning. IEEE Trans Cybern, 2020, 50: 1418–1429
33 Wen J, Zhang Z, Zhang Z, et al. Unified tensor framework for incomplete multi-view clustering and missing-view Inferring. In: Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI), 2021
34 Liu X W, Zhu X Z, Li M M, et al. Multiple kernel k-means with incomplete kernels. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI), San Francisco, 2017. 2259–2265
35 Zhu X Z, Liu X W, Li M M, et al. Localized incomplete multiple kernel k-means. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), Stockholm, 2018. 3271–3277
36 Liu X W, Zhu X Z, Li M M, et al. Late fusion incomplete multi-view clustering. IEEE Trans Pattern Anal Mach Intell, 2019, 41: 2410–2423
37 Bai L, Liang J Y. Sparse subspace clustering with entropy-norm. In: Proceedings of the 37th International Conference on Machine Learning (ICML), Vienna, 2020. 561–568
38 Nie F P, Wang X Q, Huang H. Clustering and projected clustering with adaptive neighbors. In: Proceedings of the 20th ACM International Conference on Knowledge Discovery & Data Mining (SIGKDD), New York, 2014. 977–986
39 Wen J, Zhang X, Xu Y, et al. Incomplete multi-view clustering via graph regularized matrix factorization. In: Proceedings of European Conference on Computer Vision (ECCV), Munich, 2018. 593–608
40 Hu M L, Chen S C. Doubly aligned incomplete multi-view clustering. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI), Stockholm, 2018. 2262–2268
41 Maaten L V D, Hinton G. Visualizing data using t-SNE. J Mach Learn Res, 2008, 9: 2579–2605