Unified Tensor Framework for Incomplete Multi-view Clustering and Missing-view Inferring

Jie Wen,[†1] **Zheng Zhang,**[†*1,2] **Zhao Zhang,**[*3] **Lei Zhu,**[4] **Lunke Fei,**[5] **Bob Zhang,**[6] **Yong Xu**[1,2]

[1]Shenzhen Key Laboratory of Visual Object Detection and Recognition, Harbin Institute of Technology, Shenzhen, Shenzhen 518055, China
[2]Peng Cheng Laboratory, Shenzhen 518055, China
[3]School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230006, China
[4]School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China
[5]School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China
[6]PAMI Research Group, Dept. of Computer and Information Science, University of Macau, Taipa, Macau
jiewen_pr@126.com, darrenzz219@gmail.com, cszzhang@gmail.com, leizhu0608@gmail.com, flksxm@126.com, bobzhang@um.edu.mo, yongxu@ymail.com

Abstract

In this paper, we propose a novel method, referred to as incomplete multi-view tensor spectral clustering with missing-view inferring (IMVTSC-MVI) to address the challenging multi-view clustering problem with missing views. Different from the existing methods which commonly focus on exploring the certain information of the available views while ignoring both of the hidden information of the missing views and the intra-view information of data, IMVTSC-MVI seeks to recover the missing views and explore the full information of such recovered views and available views for data clustering. In particular, IMVTSC-MVI incorporates the feature space based missing-view inferring and manifold space based similarity graph learning into a unified framework. In such a way, IMVTSC-MVI allows these two learning tasks to facilitate each other and can well explore the hidden information of the missing views. Moreover, IMVTSC-MVI introduces the low-rank tensor constraint to capture the high-order correlations of multiple views. Experimental results on several datasets demonstrate the effectiveness of IMVTSC-MVI for incomplete multi-view clustering.

Introduction

Multi-view clustering (MVC) which attempts to group a set of data points by adaptively exploring the information of multiple views without the guidance of label information, has witnessed great development in the past decades (Yang and Wang 2018; Zhang et al. 2018; Gao et al. 2020; Tang et al. 2020; Li et al. 2020b; Huang, Kang, and Xu 2020; Zhan et al. 2018). Generally, conventional MVC performs model learning and clustering based on the assumption that all views are fully observed and strictly aligned. However, it is common that some views are unavailable in many practical applications, which forms incomplete multi-view data (Xu, Tao, and Xu 2015; Liu et al. 2018; Wen et al. 2020b;

Wang et al. 2019a; Zhang et al. 2020, 2019). For instance, not all webpages are composed of links, images, and texts in web document analysis. Clustering on such incomplete multi-view data is called incomplete MVC (IMVC) (Hu and Chen 2019a; Wen et al. 2020a). The missing views not only lead to information loss, but also increase the difficulties of the excavation of complementary information. These factors make IMVC a challenging problem.

For IMVC, Trivedi et al. proposed a two-step kernel learning based approach, which first recovers the incomplete kernel and then performs the kernel canonical correlation analysis for clustering (Trivedi et al. 2010). Since this method needs one complete view as a reference for the missing view completion, it is inapplicable to practical applications. In recent years, many more flexible multiple kernel learning based methods have been proposed to address the above issue (Liu et al. 2019c,b, 2020, 2019a; Wang et al. 2019b). For instance, incomplete multiple kernel k-means with mutual kernel completion (IMKKM-IK-MKC) designs a joint framework for incomplete kernel completion and clustering (Liu et al. 2019c). Besides of these methods, several graph learning and matrix factorization (MF) based methods have also been proposed for IMVC. Representatively, incomplete multi-view spectral clustering with adaptive graph learning (IMVSC_AGL) explores the certain self-representation information learned from the available instances to obtain the consensus representation (Wen, Xu, and Liu 2020). Different from the graph learning based methods which transfer the view missing problem from the feature space to graph space, MF based methods focus on obtaining the consensus representation shared by all views from the original feature space. Partial multi-view clustering (PMVC) is one of the pioneer works, which jointly decomposes all available views into one consensus representation whose uniformity is guaranteed by the connections among paired views (Li, Jiang, and Zhou 2014). Based on PMVC, incomplete multi-modal grouping (IMG) (Zhao, Liu, and Fu 2016), partial multi-view subspace clustering (Xu et al. 2018), and IMVC via graph regularized matrix factorization (IMVC_GRMF)

(Wen et al. 2019b) seek to improve the clustering performance by introducing the manifold structure based graph constraint. Multiple incomplete multi-view clustering (MIC) (Shao, He, and Philip 2015), online multi-view clustering (OMVC) (Shao et al. 2016), doubly aligned incomplete multi-view clustering (DAIMC) (Hu and Chen 2019a), and one-pass incomplete multi-view clustering (OPIMC) (Hu and Chen 2019b) designed more flexible weighted MF based models for IMVC, which commonly fill the missing views with the corresponding average instance or zero vector and then impose the prior view-available information as the diagonal matrix on the joint matrix factorization. The core idea of these weighted MF based methods is also to explore the certain paired connections among those aligned views to guarantee the semantic consistency of the common representation across views.

Recently, some IMC methods attempted to solve the view missing and view unaligned problems by recovering the missing views. For instance, Wen et al. proposed a unified embedding alignment framework (UEAF) based on the orthogonal MF (Wen et al. 2019a). Two generative adversarial networks (GANs) based deep methods were also proposed for missing view recovery and IMC (Xu et al. 2019; Wang et al. 2018). However, these methods are inflexible and only applicable to the special incomplete data, where UEAF requires that the feature dimensions of all views are larger than the cluster number and the two deep methods are only suitable to the case with sufficiently aligned samples whose views are all fully observed.

Recovering the missing views based on the available views is interesting and intuitive to address the IMC problem. It not only naturally addresses the incomplete problem, but also has the potential to utilize the hidden information of the missing views to enhance the clustering performance. To this end, we propose a novel missing-view recovery based IMC framework, called IMVTSC-MVI, to address the challenging problem on MVC with missing views. Specifically, IMVTSC-MVI designs a new low-rank representation model to learn the similarity graphs of all views and recover the missing instances simultaneously. Then, to guarantee the semantic consistency of the multiple views, a consensus constraint is imposed on the graphs. Moreover, a tensor constraint is introduced to IMVTSC-MVI to explore the high-order correlations of multiple views. By integrating the above constraints into a joint low-rank tensor learning framework, IMVTSC-MVI can sufficiently explore the inter-view and intra-view information for missing-view recovery and similarity graph learning. The main contributions of this paper are summarized as follows:

1) IMVTSC-MVI provides a very novel and effective model for missing view recovery and incomplete multi-view clustering.

2) Different from the existing methods that only explore information across different views, IMVTSC-MVI sufficiently explores the intra-view and inter-view information by integrating the consensus graph constraint and tensor rank constraint into the representation based graph learning model.

Notations and preliminaries

In this paper, the matrix and tensor matrix are denoted by $A \in R^{n_1 \times n_2}$ and $\mathcal{A} \in R^{n_1 \times n_2 \times n_3}$, respectively. The nuclear norm of matrix $A \in R^{m \times n}$ is denoted as $\|A\|_* = \sum_i \sigma_i$, where σ_i denotes the ith singular value of matrix A. The l_F norm of tensor $\mathcal{A}$ is defined as $\|\mathcal{A}\|_F = \sqrt{\sum_{i,j,k} (\mathcal{A}_{i,j,k})^2}$, where $\mathcal{A}_{i,j,k}$ is the (i,j,k)th element. $I \in R^{n \times n}$ denotes the identity matrix. The identity tensor is denoted by $\mathcal{I} \in R^{n \times n \times n_3}$ whose superdiagonal elements $\mathcal{I}_{i,i,i} = 1$ ($1 \leq i \leq \min(n, n_3)$). The transpose tensor of $\mathcal{A} \in R^{n_1 \times n_2 \times n_3}$ is $\mathcal{A}^T \in R^{n_2 \times n_1 \times n_3}$, which is the transposing of all frontal slices. 1 denotes a vector whose elements are all 1.

The tensor singular value decomposition operation (t-SVD) and tensor nuclear norm are defined as follows (Wu, Lin, and Zha 2019; Kilmer et al. 2013).

Definition 1 (t-SVD): The SVD of tensor $\mathcal{A} \in R^{n_1 \times n_2 \times n_3}$ is defined as: $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$, where $\mathcal{U} \in R^{n_1 \times n_1 \times n_3}$ and $\mathcal{V} \in R^{n_2 \times n_2 \times n_3}$ are orthogonal tensor, *i.e.*, $\mathcal{U}^T * \mathcal{U} = \mathcal{U} * \mathcal{U}^T = \mathcal{I}$ and $\mathcal{V}^T * \mathcal{V} = \mathcal{V} * \mathcal{V}^T = \mathcal{I}$. $\mathcal{S} \in R^{n_1 \times n_2 \times n_3}$ is a f-diagonal tensor whose entire frontal slices are diagonal matrices. '*' denotes the 't-product'.

Definition 2 (t-SVD based tensor nuclear norm): For a tensor $\mathcal{A} \in R^{n_1 \times n_2 \times n_3}$, the t-SVD based tensor nuclear norm is defined as: $\|\mathcal{A}\|_{\circledast} = \sum_{k=1}^{n_3} \left\| \mathcal{A}_f^{(k)} \right\|_* = \sum_{i=1}^{\min(n_1, n_2)} \sum_{k=1}^{n_3} \left| \mathcal{S}_f^{(k)}(i,i) \right|$, where $\mathcal{A}_f$ denotes the fast Fourier transformation (FFT) of tensor $\mathcal{A}$ along the third dimension, *i.e.*, $\mathcal{A}_f = fft(\mathcal{A}, [], 3)$. $\left\| \mathcal{A}_f^{(k)} \right\|_*$ denotes the nuclear norm of the kth frontal slice of tensor $\mathcal{A}_f$. $\mathcal{S}_f^{(k)}(i,i)$ is the ith singular value of $\mathcal{A}_f^{(k)}$ calculated by SVD on $\mathcal{A}_f^{(k)}$, *i.e.*, $\mathcal{A}_f^{(k)} = \mathcal{U}_f^{(k)} \mathcal{S}_f^{(k)} \mathcal{V}_f^{(k)T}$.

The proposed method

Recovering the missing views is significant for addressing the IMC problem. However, it is impossible to obtain reasonable missing views by just borrowing the information from the other views in the existing methods, especially for the case with less aligned information among views. If we can simultaneously explore the inter-view information and intra-view information of different views, we may obtain quite good results. To this end, a very new method, *i.e.*, IMVTSC-MVI, is proposed to infer the missing views and partition the incomplete multi-view data in this section. The pipeline of IMVTSC-MVI is shown in Fig. 1.

Model of IMVTSC-MVI

In fields of clustering, graph learning based methods are popular and have received lots of attention in recent years (Nie et al. 2016; Wan and Meila 2016; Wang, Wu, and Kittler 2020; Wang et al. 2020b; Xie et al. 2020). For a given set of data points, it seeks to construct the intrinsic graph whose every element represents a kind of similarity of the corresponding two points for clustering. In this branch, low-rank representation is widely considered for obtaining such
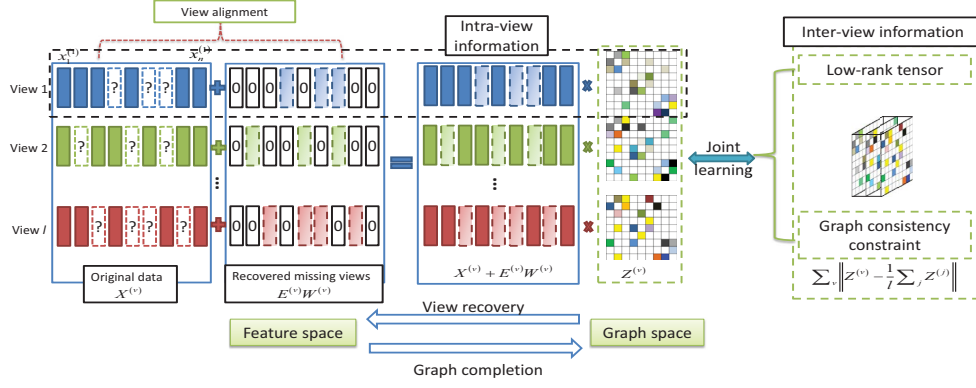
Figure 1: The pipeline of IMVTSC-MVI.

a graph (Vidal and Favaro 2014; Feng et al. 2014; Liu et al. 2012; Chen, Wu, and Kittler 2020). For multi-view data, a naive low-rank representation based graph learning model can be formulated as follows (Zhang et al. 2015):

$$\min_{\{Z^{(v)}, B^{(v)}\}_{v=1}^l} \sum_{v=1}^l \left\| Z^{(v)} \right\|_* + \lambda_v \psi \left(B^{(v)} \right)$$
$$s.t.\ X^{(v)} = X^{(v)} Z^{(v)} + B^{(v)}$$

(1)

where $X^{(v)} \in R^{m_v \times n}$ denotes the set of data points in the vth view, m_v denotes the feature dimension of the vth view, n denotes the number of data points. $Z^{(v)} \in R^{n \times n}$ is the similarity graph of the vth view to learn. $B^{(v)} \in R^{m_v \times n}$ denotes the reconstruction errors, $\psi \left(B^{(v)} \right)$ is a function to simulate different kinds of noises. $\lambda_v > 0$ is a penalty parameter to balance the importance of the corresponding item.

This naive model requires all views to be complete. Besides this, it ignores the complementary information among different views. To address these issues, we propose the following tensor based incomplete multi-view low-rank representation model, which seeks to recover the missing views and obtains the intrinsic graphs of all views simultaneously:

$$\min_{\{E^{(v)}, Z^{(v)}, B^{(v)}\}_{v=1}^l} \sum_{v=1}^l \left(\lambda_1 \left\| E^{(v)} \right\|_F^2 + \lambda_2 \left\| B^{(v)} \right\|_1 \right) + \|\mathcal{Z}\|_{\circledast}$$
$$s.t.\ X^{(v)} + E^{(v)} W^{(v)} = \left(X^{(v)} + E^{(v)} W^{(v)} \right) Z^{(v)} + B^{(v)},$$
$$Z^{(v)} \geqslant 0, diag \left(Z^{(v)} \right) = 0, Z^{(v)} 1 = 1$$

(2)

where $X^{(v)} \in R^{m_v \times n}$ denotes the set of available instances and missing instances of the vth view, in which the missing instances are filled as a zero vector. $E^{(v)} \in R^{m_v \times n_v}$ denotes the missing instances to recover, where n_v is the number of missing instances in the vth view. Error matrix $B^{(v)} \in R^{m_v \times n}$ is used to compensate the sparse noises. Nonnegative constraint $Z^{(v)} \geqslant 0$ explicitly enforces the representation coefficients to be the graph weights that denote the similarity degree of the corresponding two samples. $diag \left(Z^{(v)} \right) = 0$ is introduced to avoid the negative contribution of samples in representing themselves. $Z^{(v)} 1 = 1$

is introduced to guarantee that all instances can be connected with at least one instance. Tensor $\|\mathcal{Z}\|_{\circledast}$ is constructed by $\left\{ Z^{(v)} \right\}_{v=1}^l$. Introducing the tensor low-rank constraints $\|\mathcal{Z}\|_{\circledast}$ is beneficial to capture the high-level correlations among different views (Wu, Lin, and Zha 2019; Xu et al. 2020). $W^{(v)} \in R^{n_v \times n}$ is a prior matrix defined as follows, which indicates the view available and missing information of the vth view:

$$W_{i,j}^{(v)} = \begin{cases} 1, & if\ the\ ith\ missing\ instance\ is\ x_j^{(v)} \\ 0, & otherwise \end{cases}$$

(3)

where $x_j^{(v)}$ denotes the jth instance of the vth view.

In fact, $W^{(v)}$ can be regarded as a mapping matrix to transform the recovered missing views $E^{(v)}$ to their corresponding locations in $X^{(v)}$. In this way, $X^{(v)} + E^{(v)} W^{(v)}$ can be regarded as a complete view, with all views being naturally aligned for multi-view learning.

Notably, different views not only contain complementary information, but also share the semantic consistency, where such consistency guarantees that all views have the same cluster decision to those data points (Wen et al. 2020b; Li et al. 2020a). Therefore, it is important to preserve such semantic consistency in multi-view clustering. To this end, we introduce a semantic consistency based constraint on the graphs. As a result, the final model is expressed as follows:

$$\min_{\{E^{(v)}, Z^{(v)}, B^{(v)}\}_{v=1}^l} \sum_{v=1}^l \left(\begin{array}{c} \lambda_1 \left\| E^{(v)} \right\|_F^2 + \lambda_2 \left\| B^{(v)} \right\|_1 \\ + \lambda_3 \left\| Z^{(v)} - \frac{1}{l} \sum_{k=1}^l Z^{(k)} \right\|_F^2 \end{array} \right) + \|\mathcal{Z}\|_{\circledast}$$
$$s.t.\ X^{(v)} + E^{(v)} W^{(v)} = \left(X^{(v)} + E^{(v)} W^{(v)} \right) Z^{(v)} + B^{(v)},$$
$$Z^{(v)} \geqslant 0, diag \left(Z^{(v)} \right) = 0, Z^{(v)} 1 = 1$$

(4)

where λ_1, λ_2, and λ_3 are positive penalty parameters.

After obtaining the similarity graphs $\left\{ Z^{(v)} \right\}_{v=1}^l$, IMVTSC-MVI performs the spectral clustering (Elhamifar and Vidal 2013) on the average graph (*i.e.*,

$Z^* = \frac{1}{l} \sum_{v=1}^{l} Z^{(v)}$) to obtain the final clustering results.

Solution to IMTSC-MVI

Alternating direction method of multipliers (ADMM) is adopted to optimize the non-convex problem (4) (Boyd et al. 2011). To make problem (4) separable, several auxiliary variables $\left\{P^{(v)}\right\}_{v=1}^{l}$ are introduced as follows:

$$\min_{\Psi} \sum_{v=1}^{l} \begin{pmatrix} \lambda_1 \left\| E^{(v)} \right\|_F^2 + \lambda_2 \left\| B^{(v)} \right\|_1 \\ +\lambda_3 \left\| Z^{(v)} - \frac{1}{l} \sum_{k=1}^{l} Z^{(k)} \right\|_F^2 \end{pmatrix} + \|\mathcal{P}\|_{\circledast}$$

$$s.t. \ X^{(v)} + E^{(v)} W^{(v)} = \left(X^{(v)} + E^{(v)} W^{(v)} \right) Z^{(v)} + B^{(v)},$$

$$Z^{(v)} = P^{(v)}, Z^{(v)} \geqslant 0, diag\left(Z^{(v)} \right) = 0, Z^{(v)} I = I \tag{5}$$

where $\mathcal{P} \in R^{n \times n \times l}$ is a 3-order tensor collected by all $\left\{P^{(v)}\right\}_{v=1}^{l}$. $\Psi = \left\{E^{(v)}, Z^{(v)}, B^{(v)}, P^{(v)}\right\}_{v=1}^{l}$.

The augmented Lagrangian function of problem (5) is:

$$L\left(\Psi\right) = \sum_{v=1}^{l} \begin{pmatrix} \lambda_1 \left\| E^{(v)} \right\|_F^2 + \lambda_2 \left\| B^{(v)} \right\|_1 \\ +\lambda_3 \left\| Z^{(v)} - \frac{1}{l} \sum_{k=1}^{l} Z^{(k)} \right\|_F^2 \\ +\frac{\mu}{2} \left\| Y^{(v)} - Y^{(v)} Z^{(v)} - B^{(v)} + \frac{C^{(v)}}{\mu} \right\|_F^2 \\ +\frac{\mu}{2} \left\| Z^{(v)} - P^{(v)} + \frac{A^{(v)}}{\mu} \right\|_F^2 \end{pmatrix} + \|\mathcal{P}\|_{\circledast} \tag{6}$$

where $Y^{(v)} = X^{(v)} + E^{(v)} W^{(v)}$, Ψ denotes the set of all variables to compute, $\mu > 0$ is a penalty parameter, $\left\{A^{(v)}\right\}_{v=1}^{l}$ and $\left\{C^{(v)}\right\}_{v=1}^{l}$ are Lagrange multiplier. Then, all variables can be obtained by iteratively solving the corresponding sub-problems as follows (Fang et al. 2020; Chen et al. 2020):

Step 1: $Z^{(v)}$ can be obtained by solving the following sub-problem:

$$Z^{(v)} = \underset{Z^{(v)} I = I, Z^{(v)} \geqslant 0, diag\left(Z^{(v)}\right)=0}{\arg\min} L\left(Z^{(v)}\right) \tag{7}$$

where

$$L\left(Z^{(v)}\right) = \frac{\mu}{2} \left\| Y^{(v)} - Y^{(v)} Z^{(v)} - B^{(v)} + \frac{C^{(v)}}{\mu} \right\|_F^2$$

$$+\lambda_3 \left\| Z^{(v)} - \frac{1}{l} \sum_{k=1}^{l} Z^{(k)} \right\|_F^2 + \frac{\mu}{2} \left\| Z^{(v)} - P^{(v)} + \frac{A^{(v)}}{\mu} \right\|_F^2.$$

By setting the partial derivative $\partial L\left(Z^{(v)}\right)/\partial Z^{(v)} = 0$, we can obtain that:

$$\tilde{Z}^{(v)} = \left(\frac{2\lambda_3(l-1)^2}{l^2} I + \mu I + \mu Y^{(v)T} Y^{(v)} \right)^{-1}$$

$$\left(2\frac{l-1}{l^2} \lambda_3 \sum_{k=1,k\neq v}^{l} Z^{(k)} + \mu P^{(v)} - A^{(v)} + H^{(v)} \right) \tag{8}$$

where $H^{(v)} = Y^{(v)T}\left(\mu\left(Y^{(v)} - B^{(v)}\right) + C^{(v)}\right)$. Then, we can solve the following problem to obtain the optimal $Z^{(v)}$:

$$Z^{(v)} = \underset{Z^{(v)} I = I, Z^{(v)} \geqslant 0, diag\left(Z^{(v)}\right)=0}{\arg\min} \left\| Z^{(v)} - \tilde{Z}^{(v)} \right\|_F^2 \tag{9}$$

The optimal solution to problem (9) is (Nie et al. 2016):

$$Z_{i,j}^{(v)} = \begin{cases} 0, & i = j \\ \left(\tilde{Z}_{i,j}^{(v)} + \eta \right)_+, & i \neq j \end{cases} \tag{10}$$

where $(a)_+$ is a function to transform the negative elements in vector a into 0 and preserve the non-negative elements. η is calculated as:

$$\eta = \frac{1}{n-1} - \frac{1}{n-1} \sum_{j=1,j\neq i}^{n} \tilde{Z}_{i,j}^{(v)} \tag{11}$$

Step 2: $\left\{P^{(v)}\right\}_{v=1}^{l}$ is obtained by solving the following sub-problem:

$$\left\{P^{(v)}\right\}_{v=1}^{l} = \underset{\left\{P^{(v)}\right\}_{v=1}^{l}}{\arg\min} \|\mathcal{P}\|_{\circledast} + \sum_{v=1}^{l} \frac{\mu}{2} \left\| Z^{(v)} - P^{(v)} + \frac{A^{(v)}}{\mu} \right\|_F^2$$

$$\Leftrightarrow \mathcal{P} = \underset{\mathcal{P}}{\arg\min} \|\mathcal{P}\|_{\circledast} + \frac{\mu}{2} \left\| \mathcal{Z} - \mathcal{P} + \frac{\mathcal{A}}{\mu} \right\|_F^2 \tag{12}$$

where $\mathcal{A} \in R^{n \times n \times l}$ is a tensor collected by all $\left\{A^{(v)}\right\}_{v=1}^{l}$. Problem (12) is a typical t-SVD based tensor nuclear norm minimization problem and has the following closed-form solution (Coates, Ng, and Lee 2011; Hu et al. 2016):

$$\mathcal{P} = \mathcal{U} * \varphi_{\tilde{\mu}}\left(\mathcal{S}\right) * \mathcal{V}^T \tag{13}$$

where $\tilde{\mu} = n\mu$, $\mathcal{Z} + \mathcal{A}/\mu = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ is obtained by t-SVD operation (Coates, Ng, and Lee 2011). $\varphi_{\tilde{\mu}}\left(\mathcal{S}\right) = \mathcal{S} * \mathcal{J}$, where $\mathcal{J} \in R^{n \times l \times n}$ is a diagonal tensor whose diagonal elements in the Fourier domain are expressed as $\mathcal{J}_f\left(i,i,j\right) = \max\left(1 - \tilde{\mu} \big/ \mathcal{S}_f^{(j)}\left(i,i\right), 0\right)$.

Step 3: $E^{(v)}$ can be obtained by optimizing the following problem:

$$E^{(v)} = \underset{E^{(v)}}{\arg\min} L\left(E^{(v)}\right) \tag{14}$$

where $L\left(E^{(v)}\right) = \frac{\mu}{2} \left\| Y^{(v)} - Y^{(v)} Z^{(v)} - B^{(v)} + \frac{C^{(v)}}{\mu} \right\|_F^2 + \lambda_1 \left\| E^{(v)} \right\|_F^2$. By setting the partial derivative $\partial L\left(E^{(v)}\right)/\partial E^{(v)} = 0$, we can obtain the following closed-form solution to $E^{(v)}$:

$$E^{(v)} = \mu Q^{(v)} R^{(v)T} \left(\mu R^{(v)} R^{(v)T} + 2\lambda_1 I \right)^{-1} \tag{15}$$

where $Q^{(v)} = X^{(v)} Z^{(v)} - X^{(v)} + B^{(v)} - C^{(v)}\big/\mu$, $R^{(v)} = W^{(v)} - W^{(v)} Z^{(v)}$.

Step 4: $B^{(v)}$ is updated by solving the problem:

$$\min \lambda_2 \left\| B^{(v)} \right\|_1 + \frac{\mu}{2} \left\| Y^{(v)} - Y^{(v)} Z^{(v)} - B^{(v)} + \frac{C^{(v)}}{\mu} \right\|_F^2 \tag{16}$$

Problem (16) has the following closed-form solution:

$$B^{(v)} = \vartheta_{\lambda_2/\mu}\left(Y^{(v)} - Y^{(v)}Z^{(v)} + C^{(v)}\big/\mu\right) \quad (17)$$

where ϑ is the shrinkage operator.

Step 5: Update multipliers $\left\{A^{(v)}\right\}_{v=1}^l$, $\left\{C^{(v)}\right\}_{v=1}^l$ and parameter μ as follows:

$$A^{(v)} = A^{(v)} + \mu\left(Z^{(v)} - P^{(v)}\right) \quad (18)$$

$$C^{(v)} = C^{(v)} + \mu\left(Y^{(v)} - Y^{(v)}Z^{(v)} - B^{(v)}\right) \quad (19)$$

$$\mu = \min\left(\mu\rho, \mu_0\right) \quad (20)$$

where ρ and μ_0 are constants.

In Algorithm 1, we summarize the above computation procedures of objective problem (4).

Computational complexity analysis

In the first step *w.r.t.* $Z^{(v)}$, the major computational cost is the inverse operation in (8), which takes $O(n^3)$ for an n by n matrix. Thus, the total computational complexity in Step 1 is about $O(ln^3)$. In Step 2, the major computational costs are the FFT, inverse FFT, and t-SVD operations. For an $n \times l \times n$ tensor matrix, the total computational complexity of FFT and inverse FFT is about $O(ln^2log(n))$, and that of the t-SVD operation is about $O(l^2n^2)$ (Wu, Lin, and Zha 2019). As a result, the total computational complexity of Step 2 is about $O(ln^2log(n) + l^2n^2)$. Considering that the other steps only have some basic matrix operations, we ignore their computational costs. According to the above analysis, the computational complexity of Algorithm 1 is about $O(\tau(ln^3 + ln^2log(n) + l^2n^2))$, where τ denotes the iteration number.

Algorithm 1 : IMVTSC-MVI (solving (4))

Input: Incomplete multi-view data $\left\{X^{(v)}\right\}_{v=1}^l$ ($X^{(v)} \in R^{m_v \times n}$), index matrix $\left\{W^{(v)}\right\}_{v=1}^l$ ($W^{(v)} \in R^{n_v \times n}$), parameters λ_1, λ_2, λ_3, penalty parameter $\mu = 0.01$, $\mu_0 = 10^8$, $\rho = 1.2$.
Initialization: $E^{(v)} = 0$, $B^{(v)} = 0$, $A^{(v)} = 0$, $C^{(v)} = 0$. $Z^{(v)}$ is initialized as the binary KNN-graph constructed from $X^{(v)}$, where the elements associated with the missing views are set as 0.
while not converged **do**

    1. Update variables $\left\{Z^{(v)}\right\}_{v=1}^l$ using (10);

    2. Update variables $\left\{P^{(v)}\right\}_{v=1}^l$ using (13);

    3. Update variables $\left\{E^{(v)}\right\}_{v=1}^l$ using (15);

    4. Update variables $\left\{B^{(v)}\right\}_{v=1}^l$ using (17);

    5. Update variables $\left\{A^{(v)}\right\}_{v=1}^l$, $\left\{C^{(v)}\right\}_{v=1}^l$, and μ using (18), (19), and (20).
end while
Output: $\left\{Z^{(v)}\right\}_{v=1}^l$

Table 1: Description of the datasets.

Dataset	# Class	# View	# Samples	# Features
Handwritten	10	5	2000	76,216,64,240,47
Caltech7	7	6	1474	48,40,254,1984,512,928
MNIST	10	2	4000	784,784
NUSWIDE	31	5	2170	65,226,145,74,129
NH_face	5	3	4660	6750,2000,3304

Experiments

Experimental settings

Compared methods: Several state-of-the-art IMC methods, including PMVC, IMG, IMVC_GRMF, MIC, OMVC, DAIMC, OPIMC, UEAF, IMKKM-IK-MKC, IMVSC_AGL, and PMVC via consistent GAN (P-MVC_CGAN) (Wang et al. 2018), are compared with the proposed method to validate its effectiveness. Besides this, two baseline methods, BSV (Best Single View) and Concat are also evaluated, where BSV reports the best Kmeans clustering results on all single views. Concat stacks features of all views into one view and then implements Kmeans on it to obtain the clustering results (Zhao, Liu, and Fu 2016).

Datasets: Five multi-view datasets listed in Table 1 are selected. Specifically, 1) **Handwritten** contains 10 classes (*i.e.*, digits '0-9') and five kinds of features, *i.e.*, Fourier coefficients, profile correlations, Karhunen-Love coefficient, pixel averages, and Zernike moments (Asuncion and Newman 2007). 2) **Caltech7** is a subset of the popular Caltech101 object dataset (Fei-Fei, Fergus, and Perona 2004). It is composed of 7 objects, 1474 samples, and six kinds of features, *i.e.*, Gabor, wavelet moments, CENTRIST, HOG, GIST, and LBP (Li et al. 2015). 3) **MNIST** dataset exploited in our work is composed of 4000 samples and two kinds of features, *i.e.*, pixel feature and edge feature, following (Wang et al. 2018). 4) **NUSWIDE** is a popular large-scale object recognition dataset (Chua et al. 2009). In our work, we randomly construct a subset which consists of 31 classes and 70 samples per class for evaluation, where five kinds of features, *i.e.*, color Histogram, color moments, color correlation, edge distribution, and wavelet texture, are extracted as the five views. 5) Notting-Hill face dataset (**NH_face**) is collected from the movie 'Notting-Hill', which is composed of 4660 faces from 5 persons (Wu et al. 2013). Three kinds of features, *i.e.*, LBP, gray pixels features, and Gabor are extracted as three views (Cao et al. 2015).

Incomplete data construction: On MNIST dataset, we randomly select $p\%$ ($p = \{10, 30, 50, 70\}$) samples as paired samples that have full views, and then divide the remaining samples into two groups, followed by removing the first view from one group and removing the second view from the other group. On the other four datsets, we randomly remove $p\%$ instances from every view under the condition that all samples have at least one view.

Evaluation metrics: Clustering accuracy (ACC), normalized mutual information (NMI), and purity are adopted as the evaluation metrics following (Shao et al. 2016).

Experimental results and analysis

Tables 2-3 list the experimental results of different IMC methods on the five datasets. From the experimental results, following points can be observed:

(1) Our method achieves the best performance on these datasets. Compared with the second best method (IMVSC_AGL) on the Caltech7 and Handwritten datasets with a missing-view rate of 30%, our method achieves about 11% and 8% improvement in terms of NMI, respectively.

(2) With the increasing of the missing-view rate, the clustering results of all methods decrease. For instance, when missing-view rate changes from 10% to 70% on the Handwritten dataset, the ACC of BSV and Concat decrease by about 41% and 47%, respectively. This indicates that missing views seriously affect the excavation of complementary information among multiple views.

(3) On the NH_face dataset with a missing-view rate of 50%, the ACC of our method is about 14% higher than that of IMVSC_AGL. This demonstrates that compared with IMVSC_AGL, the proposed method can obtain higher quality graphs that capture the intrinsic structures of all views.

(4) Our method performs better than two view-recovery based methods, *i.e.*, UEAF and PMVC_CGAN. This demonstrates that our method is superior to the two methods on missing-view recovery and hidden information excavation.
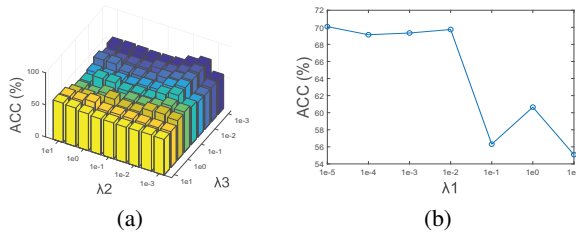

(a)   (b)

Figure 2: ACC (%) of our method w.r.t (a) λ_2 and λ_3 with $\lambda_1 = 1e - 5$ and (b) λ_1 with $\lambda_2 = 5e - 3$ and $\lambda_3 = 5e - 3$ on the Caltech7 dataset with a missing-view rate of 30%.

Parameter analysis

From the experimental results shown in Fig.2, we can observe that the proposed method is sensitive to parameter λ_2 to some extent, where the best performance is obtained when $\lambda_2 = 0.005$. This indicates that selecting a suitable λ_2 for the sparse error item is important in our method. Besides this, one can see that the proposed method is insensitive to parameters λ_1 and λ_3 to some extent, where selecting the two parameters from the range of $[10^{-5}, 10^{-2}]$ and $[10^{-3}, 10^{-2}]$ can guarantee a good performance.

View inferring and graph recovery

View inferring: Fig. 3 shows the missing instances of the two datasets and the corresponding recovered instances obtained by our method. We can observe that our method can well recover the missing-views with the same contour, posture, and expression as the original data. This demonstrates the effectiveness of our method on missing-view inferring.

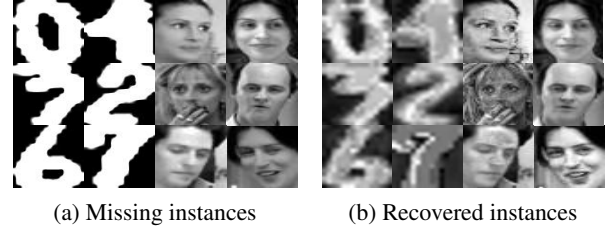(a) Missing instances   (b) Recovered instances

Figure 3: The performance of missing-view recovery, where images of the first two columns and last two columns in (a) are the missing instances of the first view from the Handwritten dataset and the second view from the NH_face dataset, respectively. Images in (b) are the corresponding recovered instances obtained by our method. Notably, 1) Images in (a) are not used in the model learning phase. 2) We have reshaped the original vector based features into images and then re-normalized them for better presentation.

(a)   (b)   (c)
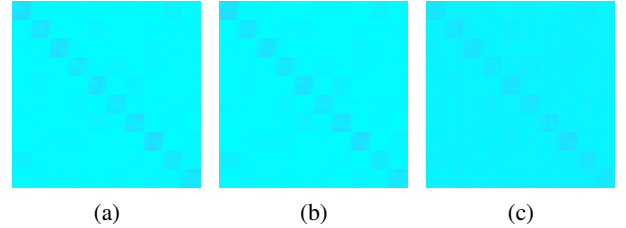
Figure 4: Graphs of the (a) first, (b) third, and (c) fifth views obtained by our method on the Handwritten dataset with a missing-view rate of 30%.

(a)   (b)
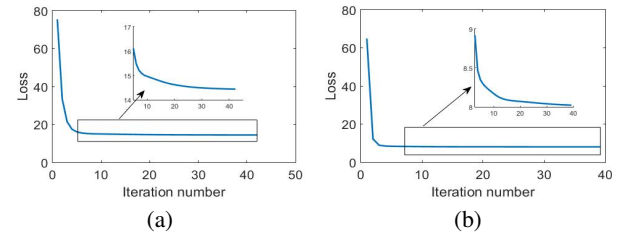
Figure 5: Objective function loss v.s. iteration of the proposed method on the (a) NH_face and (b) Caltech7 datasets with a missing-view rate of 30%.

Graph recovery: Fig.4 shows three affinity graphs obtained by our method on the Handwritten dataset. One can observe that: 1) all graphs have a clear block diagonal structure, where the number the blocks is equal to the cluster number. 2) The connected weights corresponding to the missing views are all well completed. The above phenomenons demonstrate that our method can discover the intrinsic view-consistent graphs from the incomplete data.

Table 2: Experimental results of different methods on the Handwritten, MNIST, and NUSWIDE datasets. Notably, the clustering results of PMVC_CGAN are from the original paper (Wang et al. 2018, 2020a) and '- -' denotes the unreported numbers.

Data	Method	ACC				NMI				Purity			
		0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7	0.1	0.3	0.5	0.7
Handwritten	BSV	68.27±5.66	51.49±2.29	38.24±2.25	27.15±1.31	62.82±3.24	47.01±1.71	32.21±1.00	19.48±0.69	70.72±4.67	53.69±1.54	39.54±2.04	27.76±1.09
	Concat	75.06±3.86	55.48±1.57	42.19±0.99	28.31±0.75	73.08±2.05	51.66±0.99	38.24±1.59	23.50±0.95	77.86±2.79	57.32±1.15	44.21±0.98	30.45±0.80
	MIC	77.59±2.41	73.29±3.41	61.27±3.16	41.34±2.69	70.84±2.08	65.39±2.08	52.95±1.33	34.71±2.11	78.72±2.14	74.31±3.15	62.89±3.08	43.25±2.86
	OMVC	65.04±6.50	55.00±5.06	36.40±4.93	29.80±4.63	56.72±5.05	44.99±4.56	35.16±4.62	25.83±8.37	65.80±6.48	55.89±4.72	38.51±4.87	31.95±5.22
	DAIMC	88.86±0.63	86.73±0.79	81.92±0.88	60.44±6.87	79.78±0.71	76.65±1.07	68.77±0.99	47.10±4.79	88.86±0.63	86.73±0.79	81.92±0.88	61.24±0.42
	OPIMC	80.20±5.40	76.45±5.15	69.50±6.54	56.66±10.06	77.26±3.11	73.74±3.42	66.57±4.18	51.86±7.97	81.19±4.60	78.96±3.37	72.00±6.39	58.16±10.35
	UEAF	85.80±5.62	76.11±7.74	65.39±5.09	61.11±1.41	77.74±3.83	69.37±3.31	55.09±2.05	50.56±1.11	86.07±5.02	76.51±7.17	66.49±4.18	61.60±1.09
	IMKKM-IK-MKC	71.78±1.74	69.07±0.73	66.08±3.25	55.55±1.39	69.43±1.28	65.42±0.61	59.04±2.69	47.36±1.78	75.53±1.10	73.12±0.61	66.58±3.26	56.26±1.07
	IMVSC_AGL	97.15±0.43	95.50±0.48	93.19±0.72	84.08±5.16	93.68±0.53	90.55±0.76	86.39±1.27	75.44±2.68	97.15±0.43	95.50±0.48	93.19±0.72	84.74±3.74
	Ours	**99.89±0.07**	**99.31±0.13**	**99.56±0.15**	**98.84±0.19**	**99.71±0.19**	**98.21±0.35**	**98.84±0.40**	**97.08±0.45**	**99.89±0.07**	**99.31±0.13**	**99.56±0.15**	**98.84±0.19**
MNIST	BSV	33.25±1.79	37.37±0.69	42.76±1.30	47.95±1.36	27.20±0.89	31.39±1.28	37.45±1.42	42.49±1.47	36.28±1.69	40.75±0.96	47.00±1.37	52.62±1.30
	Concat	36.88±1.96	39.24±1.47	43.79±1.71	47.37±1.08	34.48±1.09	33.38±0.54	37.42±1.38	43.17±0.69	41.96±1.30	43.15±0.93	47.72±1.41	52.44±1.04
	PMVC	41.36±2.29	43.42±2.99	44.68±1.23	45.84±1.59	35.46±0.25	38.51±1.63	39.43±1.37	39.83±1.71	46.42±2.06	48.31±2.78	49.07±0.81	49.93±2.09
	IMG	46.34±3.36	47.13±2.24	46.88±1.51	48.31±1.22	39.74±2.42	40.71±2.56	39.87±1.05	44.16±1.09	51.61±3.47	52.12±2.01	51.55±1.75	53.43±1.06
	IMVC_GRMF	49.12±2.46	50.59±2.59	52.37±1.67	52.46±1.59	47.36±0.97	48.18±1.57	50.73±0.98	51.57±1.03	54.59±1.74	56.36±1.81	57.96±1.18	58.24±1.19
	MIC	43.96±2.38	44.42±2.28	44.17±1.37	45.38±2.82	38.77±1.35	40.81±1.28	40.53±0.67	41.61±1.50	48.68±2.06	49.18±1.37	49.05±1.34	50.52±2.18
	OMVC	40.44±2.95	42.23±2.17	40.36±2.20	41.44±3.39	36.21±1.47	36.68±2.16	35.64±1.89	32.25±2.95	45.95±2.31	47.49±2.52	45.02±3.14	45.51±2.93
	DAIMC	45.33±4.12	48.19±1.38	49.25±1.67	49.36±1.87	37.46±3.04	41.09±1.58	43.47±0.82	44.15±0.75	48.70±4.04	52.70±1.76	54.41±1.69	55.13±1.18
	OPIMC	41.40±2.51	48.02±2.63	47.77±3.39	48.71±2.44	34.29±2.33	43.98±1.98	44.63±1.47	45.65±1.15	43.91±2.08	52.95±2.33	53.07±2.86	54.38±2.21
	UEAF	34.75±1.19	49.87±0.91	50.56±1.63	51.07±0.86	32.66±0.74	44.25±1.28	44.31±1.32	45.50±0.99	41.06±0.75	54.86±0.70	55.41±1.48	56.04±0.87
	IMKKM-IK-MKC	47.56±2.18	51.02±0.66	51.72±0.58	52.45±0.41	40.39±1.17	42.76±0.70	43.88±0.54	45.10±0.39	53.08±1.53	56.02±0.72	56.73±0.50	57.42±0.47
	IMVSC_AGL	54.77±2.95	55.44±1.63	55.23±2.61	57.71±1.34	54.10±1.90	55.85±1.17	55.92±2.52	58.68±1.04	60.96±2.29	60.39±1.10	61.17±1.63	63.25±0.50
	PMVC_CGAN	45.17±0.86	48.36±0.71	52.80±0.78	52.02±0.70	39.33±- -	43.22±- -	49.61±- -	48.22±- -	49.09±- -	51.81±- -	58.5±- -	55.45±- -
	Ours	56.64±2.46	56.60±0.84	55.64±0.12	62.65±0.85	56.24±0.83	57.73±0.58	57.18±2.15	59.37±1.10	62.74±0.28	63.39±1.11	62.91±0.94	65.74±0.97
NUSWIDE	BSV	13.16±0.23	11.50±0.18	9.66±0.35	7.72±0.20	15.86±0.23	13.51±0.23	11.09±0.26	8.45±0.24	14.36±0.28	11.49±0.18	10.39±0.23	8.23±0.26
	Concat	16.24±0.28	13.92±0.31	11.35±0.44	10.39±0.34	19.45±0.17	17.29±0.39	13.93±0.52	12.14±0.37	17.83±0.18	15.14±0.29	12.28±0.57	11.12±0.46
	MIC	14.24±0.29	12.21±0.64	10.69±0.54	9.88±0.25	16.44±0.56	14.29±0.40	12.20±0.54	10.89±0.51	15.52±0.26	13.34±0.45	11.57±0.59	10.59±0.34
	OMVC	13.48±0.45	11.98±0.37	11.98±0.37	9.88±0.41	15.86±0.43	14.16±0.32	12.15±0.42	10.75±0.55	14.74±0.36	13.34±0.45	11.57±0.59	10.59±0.34
	DAIMC	14.21±0.46	14.36±0.47	12.26±0.42	10.46±0.56	17.20±0.36	16.48±0.45	13.97±0.51	11.29±1.01	15.65±0.40	14.68±0.35	13.44±0.55	11.39±0.67
	OPIMC	14.98±0.31	14.36±0.47	12.62±1.06	10.55±0.59	16.80±0.61	16.03±0.29	14.15±0.75	11.05±0.29	15.85±0.28	14.97±0.37	13.54±1.14	10.99±0.55
	UEAF	16.19±0.38	15.21±0.35	12.82±0.46	10.71±0.27	19.17±0.30	18.27±0.12	15.65±0.58	12.68±0.11	18.07±0.66	16.44±0.29	14.04±0.49	11.48±0.28
	IMKKM-IK-MKC	13.73±0.19	12.71±0.58	11.49±0.46	10.73±0.21	15.38±0.22	14.63±0.67	13.81±0.43	12.83±0.26	14.76±0.17	13.68±0.63	12.49±0.59	11.51±0.27
	IMVSC_AGL	15.63±0.43	14.42±0.15	13.39±0.67	**11.57±0.71**	18.25±0.32	16.85±0.37	16.12±0.65	13.18±0.72	16.89±0.42	15.43±0.17	14.45±0.52	**12.49±0.73**
	Ours	**16.82±0.35**	**16.21±0.58**	**14.21±0.71**	11.35±0.29	**20.23±0.57**	**19.19±0.41**	**16.86±0.49**	**13.80±0.14**	**18.46±0.44**	**17.57±0.32**	**15.28±0.72**	12.18±0.35

Table 3: Experimental results of different methods on the Caltech7 and NH_face datasets.

Data	Method	ACC			NMI			Purity		
		0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.5
Caltech7	BSV	43.89±1.37	39.06±1.26	38.31±1.68	39.66±2.23	31.63±1.51	26.81±1.38	84.08±1.23	75.25±0.71	68.97±0.49
	Concat	41.25±1.67	40.55±1.89	38.06±0.88	43.48±0.92	37.99±2.17	30.28±0.66	84.91±0.50	82.54±1.12	77.56±0.98
	MIC	44.07±4.97	38.01±2.12	35.80±2.34	33.71±2.66	27.35±1.69	20.44±0.98	78.12±1.76	73.31±0.72	68.26±1.40
	OMVC	40.88±1.54	36.82±1.65	33.28±4.40	28.13±2.54	25.32±1.03	18.76±4.22	79.21±1.77	77.73±1.35	74.05±4.74
	DAIMC	48.29±6.76	47.46±3.42	44.89±4.88	44.61±3.88	38.45±2.88	36.28±2.34	83.32±1.31	76.83±3.23	75.50±1.17
	OPIMC	49.24±2.89	48.34±4.36	44.12±5.85	42.98±1.02	41.54±2.38	35.98±2.77	84.89±0.69	83.70±1.80	80.64±2.06
	UEAF	50.82±4.05	42.71±0.84	36.32±4.22	39.44±2.07	31.07±1.99	24.02±1.37	81.49±1.78	78.26±2.12	76.29±1.93
	IMKKM-IK-MKC	36.54±0.51	34.87±1.53	36.05±0.45	24.09±0.98	23.45±0.52	22.91±0.67	72.98±0.80	73.82±0.53	72.52±1.55
	IMVSC_AGL	54.72±1.01	54.61±2.04	51.78±3.03	44.05±0.97	42.61±1.92	37.35±0.32	84.27±0.81	83.98±0.93	82.31±0.30
	Ours	**64.18±0.49**	**63.19±2.41**	**54.17±2.04**	**55.97±0.71**	**53.84±2.72**	**51.37±1.40**	**89.09±0.52**	**88.69±1.06**	**88.19±1.82**
NH_face	BSV	69.09±4.76	56.82±2.28	46.54±1.90	56.26±4.07	39.29±2.63	26.20±1.09	73.59±2.96	60.13±1.52	50.15±1.28
	Concat	85.87±2.64	63.14±2.78	52.99±1.84	81.46±1.70	59.12±1.14	47.42±1.29	87.39±1.57	87.39±1.57	62.21±1.04
	MIC	78.83±4.07	77.22±0.76	75.77±4.05	73.04±2.78	66.82±0.80	62.84±3.20	82.54±1.66	78.83±0.64	77.40±3.48
	OMVC	75.35±2.11	72.85±3.17	70.61±2.77	68.45±3.22	65.44±2.89	63.34±4.36	80.89±3.05	77.96±2.33	74.52±3.55
	DAIMC	87.42±4.15	85.35±3.44	84.57±3.49	78.37±3.42	74.71±2.91	70.09±5.08	87.03±2.74	85.66±2.91	84.66±3.41
	OPIMC	79.82±8.32	74.57±3.81	71.25±6.27	69.92±6.36	66.87±1.86	64.65±6.94	81.56±5.12	79.02±1.27	78.21±4.01
	UEAF	80.36±0.10	71.22±0.68	64.37±1.13	67.11±0.52	55.52±2.55	47.97±1.50	81.67±0.13	73.32±0.70	68.49±1.21
	IMKKM-IK-MKC	74.34±0.34	75.92±0.93	71.22±1.19	65.21±0.32	66.83±1.24	65.27±1.66	78.96±0.07	79.18±0.16	79.94±1.03
	IMVSC_AGL	87.04±1.84	84.33±2.02	81.21±8.96	78.26±1.73	77.59±0.75	73.34±3.59	87.33±1.22	85.55±1.35	85.38±3.24
	Ours	**95.66±2.84**	**95.79±1.98**	**95.21±2.34**	**91.99±4.87**	**91.68±4.62**	**90.35±3.97**	**95.66±2.84**	**95.79±1.98**	**95.21±2.34**

Convergence analysis

As presented in the previous section, the complex objective function is divided into four convex sub-problems. Optimizing every minimization subproblem will decrease the objective value. This indicates that the loss of objective problem (5) is monotonically decreasing during the updating process of these variables alternately. In Fig.5, we plot the objective function loss v.s. the iteration number on the Caltech7 and NH_face datasets with a missing-view rate of 30%. From the figures, we can also observe that the objective loss is monotonically decreasing and fast converges to a stable point, which demonstrate the good convergence property of the proposed optimization algorithm.

Conclusion

In this paper, we proposed a novel framework, called IMVTSC-MVI, for missing-view inferring and incomplete multi-view clustering. To obtain more reasonable missing-views, IMVTSC-MVI jointly introduced the tensor low-rank representation constraint and semantic consistency based graph constraint. The two approaches not only enable the model to explore the inter-view information and intra-view information, but also bridges the gap between the feature space and manifold space, which are beneficial to obtain the most reasonable missing-views and optimal similarity graphs. Experimental results show that IMVTSC-MVI can well recover the missing views and greatly enhance the clustering performance.

Acknowledgments

References

Asuncion, A.; and Newman, D. 2007. UCI machine learning repository.

Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J.; et al. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning* 3(1): 1–122.

Cao, X.; Zhang, C.; Zhou, C.; Fu, H.; and Foroosh, H. 2015. Constrained multi-view video face clustering. *TIP* 24(11): 4381–4393.

Chen, Z.; Wu, X.-J.; and Kittler, J. 2020. Low-rank discriminative least squares regression for image classification. *Signal Processing* 173: 107485.

Chen, Z.; Wu, X.-J.; Yin, H.-F.; and Kittler, J. 2020. Noise-robust dictionary learning with slack block-diagonal structure for face recognition. *PR* 100: 107118.

Chua, T.-S.; Tang, J.; Hong, R.; Li, H.; Luo, Z.; and Zheng, Y. 2009. NUS-WIDE: a real-world web image database from National University of Singapore. In *ACM CIVR*, 1–9.

Coates, A.; Ng, A.; and Lee, H. 2011. An analysis of single-layer networks in unsupervised feature learning. In *AISTATS*, 215–223.

Elhamifar, E.; and Vidal, R. 2013. Sparse subspace clustering: Algorithm, theory, and applications. *TPAMI* 35(11): 2765–2781.

Fang, X.; Han, N.; Zhou, G.; Teng, S.; Xu, Y.; and Xie, S. 2020. Dynamic double classifiers approximation for cross-domain recognition. *TCYB* .

Fei-Fei, L.; Fergus, R.; and Perona, P. 2004. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. In *CVPR Workshop*, 178–178. IEEE.

Feng, J.; Lin, Z.; Xu, H.; and Yan, S. 2014. Robust subspace segmentation with block-diagonal prior. In *CVPR*, 3818–3825.

Gao, Q.; Xia, W.; Wan, Z.; Xie, D.; and Zhang, P. 2020. Tensor-SVD based graph learning for multi-view subspace clustering. In *AAAI*, volume 34, 3930–3937.

Hu, M.; and Chen, S. 2019a. Doubly aligned incomplete multi-view clustering. In *IJCAI*, 2262–2268.

Hu, M.; and Chen, S. 2019b. One-Pass Incomplete Multi-view Clustering. In *AAAI*, 3838–3845.

Hu, W.; Tao, D.; Zhang, W.; Xie, Y.; and Yang, Y. 2016. The twist tensor nuclear norm for video completion. *TNNLS* 28(12): 2961–2973.

Huang, S.; Kang, Z.; and Xu, Z. 2020. Auto-weighted multi-view clustering via deep matrix decomposition. *PR* 97: 107015.

Kilmer, M. E.; Braman, K.; Hao, N.; and Hoover, R. C. 2013. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIMAX* 34(1): 148–172.

Li, J.; Li, M.; Lu, G.; Zhang, B.; Yin, H.; and Zhang, D. 2020a. Similarity and diversity induced paired projection for cross-modal retrieval. *Information Sciences* 539: 215–228.

Li, J.; Yong, H.; Wu, F.; and Li, M. 2020b. Online Multiview Subspace Learning with Mixed Noise. In *ACM MM*, 3838–3846.

Li, S.-Y.; Jiang, Y.; and Zhou, Z.-H. 2014. Partial multi-view clustering. In *AAAI*, 1969–1974.

Li, Y.; Nie, F.; Huang, H.; and Huang, J. 2015. Large-scale multi-view spectral clustering via bipartite graph. In *AAAI*, 2750–2756.

Liu, G.; Lin, Z.; Yan, S.; Sun, J.; Yu, Y.; and Ma, Y. 2012. Robust recovery of subspace structures by low-rank representation. *TPAMI* 35(1): 171–184.

Liu, X.; Li, M.; Tang, C.; Xia, J.; Xiong, J.; Liu, L.; Kloft, M.; and Zhu, E. 2020. Efficient and effective regularized incomplete multi-view clustering. *TPAMI* .

Liu, X.; Wang, L.; Zhu, X.; Li, M.; Zhu, E.; Liu, T.; Liu, L.; Dou, Y.; and Yin, J. 2019a. Absent multiple kernel learning algorithms. *TPAMI* 42(6): 1303–1316.

Liu, X.; Zhu, X.; Li, M.; Tang, C.; Zhu, E.; Yin, J.; and Gao, W. 2019b. Efficient and effective incomplete multi-view clustering. In *AAAI*, 1–8.

Liu, X.; Zhu, X.; Li, M.; Wang, L.; Tang, C.; Yin, J.; Shen, D.; Wang, H.; and Gao, W. 2018. Late fusion incomplete multi-view clustering. *TPAMI* 41(10): 2410–2423.

Liu, X.; Zhu, X.; Li, M.; Wang, L.; Zhu, E.; Liu, T.; Kloft, M.; Shen, D.; Yin, J.; and Gao, W. 2019c. Multiple kernel k-means with incomplete kernels. *TPAMI* .

Nie, F.; Wang, X.; Jordan, M. I.; and Huang, H. 2016. The Constrained Laplacian Rank algorithm for graph-based clustering. In *AAAI*, 1969–1976.

Shao, W.; He, L.; Lu, C.-t.; and Philip, S. Y. 2016. Online multi-view clustering with incomplete views. In *ICBD*, 1012–1017. IEEE.

Shao, W.; He, L.; and Philip, S. Y. 2015. Multiple Incomplete Views Clustering via Weighted Nonnegative Matrix Factorization with $L_2,1$ Regularization. In *ECML PKDD*, 318–334. Springer.

Tang, C.; Liu, X.; Zhu, X.; Zhu, E.; Luo, Z.; Wang, L.; and Gao, W. 2020. CGD: Multi-View Clustering via Cross-View Graph Diffusion. In *AAAI*, 5924–5931.

Trivedi, A.; Rai, P.; Daumé III, H.; and DuVall, S. L. 2010. Multiview clustering with incomplete views. In *NIPS Workshop*, volume 224.

Vidal, R.; and Favaro, P. 2014. Low rank subspace clustering (LRSC). *PRL* 43: 47–61.

Wan, Y.; and Meila, M. 2016. Graph Clustering: Blockmodels and model free results. In *NIPS*, 2478–2486.

Wang, Q.; Ding, Z.; Tao, Z.; Gao, Q.; and Fu, Y. 2018. Partial Multi-view Clustering via Consistent GAN. In *ICDM*, 1290–1295. IEEE.

Wang, Q.; Ding, Z.; Tao, Z.; Gao, Q.; and Fu, Y. 2020a. Generative Partial Multi-View Clustering. *arXiv preprint arXiv:2003.13088* .

Wang, R.; Wu, X.-J.; Chen, K.-X.; and Kittler, J. 2020b. Multiple Riemannian Manifold-valued Descriptors based Image Set Classification with Multi-Kernel Metric Learning. *TBD* .

Wang, R.; Wu, X.-J.; and Kittler, J. 2020. Graph Embedding Multi-Kernel Metric Learning for Image Set Classification with Grassmann Manifold-valued Features. *IEEE TMM* .

Wang, S.; Li, M.; Hu, N.; Zhu, E.; Hu, J.; Liu, X.; and Yin, J. 2019a. K-means clustering with incomplete data. *IEEE Access* 7: 69162–69171.

Wang, S.; Liu, X.; Zhu, E.; Tang, C.; Liu, J.; Hu, J.; Xia, J.; and Yin, J. 2019b. Multi-view clustering via late fusion alignment maximization. In *IJCAI*, 3778–3784. AAAI Press.

Wen, J.; Xu, Y.; and Liu, H. 2020. Incomplete multiview spectral clustering with adaptive graph learning. *TCYB* 50(4): 1418–1429.

Wen, J.; Yan, K.; Zhang, Z.; Xu, Y.; Wang, J.; Fei, L.; and Zhang, B. 2020a. Adaptive Graph Completion Based Incomplete Multi-view Clustering. *TMM* .

Wen, J.; Zhang, Z.; Xu, Y.; Zhang, B.; Fei, L.; and Liu, H. 2019a. Unified embedding alignment with missing views inferring for incomplete multi-view clustering. In *AAAI*, 5395–5400.

Wen, J.; Zhang, Z.; Xu, Y.; and Zhong, Z. 2019b. Incomplete multi-view clustering via graph regularized matrix factorization. In *ECCV Workshops*, 593–608.

Wen, J.; Zhang, Z.; Zhang, Z.; Fei, L.; and Wang, M. 2020b. Generalized Incomplete Multiview Clustering With Flexible Locality Structure Diffusion. *TCYB* .

Wu, B.; Zhang, Y.; Hu, B.-G.; and Ji, Q. 2013. Constrained clustering and its application to face clustering in videos. In *CVPR*, 3507–3514.

Wu, J.; Lin, Z.; and Zha, H. 2019. Essential tensor learning for multi-view spectral clustering. *TIP* 28(12): 5910–5922.

Xie, D.; Gao, Q.; Wang, Q.; Zhang, X.; and Gao, X. 2020. Adaptive latent similarity learning for multi-view clustering. *Neural Networks* 121: 409–418.

Xu, C.; Guan, Z.; Zhao, W.; Wu, H.; Niu, Y.; and Ling, B. 2019. Adversarial Incomplete Multi-view Clustering. In *IJCAI*, 3933–3939. AAAI Press.

Xu, C.; Tao, D.; and Xu, C. 2015. Multi-view learning with incomplete views. *TIP* 24(12): 5812–5825.

Xu, H.; Zhang, X.; Xia, W.; Gao, Q.; and Gao, X. 2020. Low-rank tensor constrained co-regularized multi-view spectral clustering. *Neural Networks* 132: 245–252.

Xu, N.; Guo, Y.; Zheng, X.; Wang, Q.; and Luo, X. 2018. Partial Multi-view Subspace Clustering. In *ACM MM*, 1794–1801. ACM.

Yang, Y.; and Wang, H. 2018. Multi-view clustering: a survey. *Big Data Mining and Analytics* 1(2): 83–107.

Zhan, K.; Niu, C.; Chen, C.; Nie, F.; Zhang, C.; and Yang, Y. 2018. Graph structure fusion for multiview clustering. *TKDE* 31(10): 1984–1993.

Zhang, C.; Cui, Y.; Han, Z.; Zhou, J. T.; Fu, H.; and Hu, Q. 2020. Deep Partial Multi-View Learning. *TPAMI* .

Zhang, C.; Fu, H.; Liu, S.; Liu, G.; and Cao, X. 2015. Low-rank tensor constrained multiview subspace clustering. In *ICCV*, 1582–1590.

Zhang, C.; Han, Z.; Fu, H.; Zhou, J. T.; Hu, Q.; et al. 2019. CPM-Nets: Cross Partial Multi-View Networks. In *NIPS*, 559–569.

Zhang, Z.; Liu, L.; Shen, F.; Shen, H. T.; and Shao, L. 2018. Binary multi-view clustering. *TPAMI* 41(7): 1774–1782.

Zhao, H.; Liu, H.; and Fu, Y. 2016. Incomplete multi-modal visual data grouping. In *IJCAI*, 2392–2398.