

Hyper-Laplacian regularized multi-view subspace clustering with low-rank tensor constraint[☆]

Gui-Fu Lu^{*}, Qin-Ru Yu, Yong Wang, Ganyi Tang

School of Computer Science and Information, Anhui Polytechnic University, WuHu, Anhui 241000, China

ARTICLE INFO

Article history:

Received 6 October 2019

Received in revised form 2 February 2020

Accepted 21 February 2020

Available online 25 February 2020

Keywords:

Multi-view features

Subspace clustering

Manifold regularization

Low-rank tensor representation

ABSTRACT

In this paper, we propose a novel hyper-Laplacian regularized multiview subspace clustering with low-rank tensor constraint method, which is referred as HLR-MSCLRT. In the HLR-MSCLRT model, the subspace representation matrices of different views are stacked as a tensor, and then the high order correlations among data can be captured. To reduce the redundancy information of the learned subspace representations, a low-rank constraint is adopted to the constructed tensor. Since data in the real world often reside in multiple nonlinear subspaces, the HLR-MSCLRT model utilizes the hyper-Laplacian graph regularization to preserve the local geometry structure embedded in a high-dimensional ambient space. An efficient algorithm is also presented to solve the optimization problem of the HLR-MSCLRT model. The experimental results on some data sets show that the proposed HLR-MSCLRT model outperforms many state-of-the-art multi-view clustering approaches.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, multiple heterogeneous features are often used in machine learning and computer vision fields. Generally, the multiple heterogeneous features are collected from multiple sources or generated from various feature extractors. In multi-view learning (Berkhin, 2002; Chao, Sun, & Bi, 2018; Gu & Li, 2018; Liu, Lai, Ou, Zhang, & Zheng, 2020; Sun, 2013; Xu, Tao, & Xu, 2013; Zhao et al., 2019), each kind of feature is a particular view. Then, to describe the data accurately and improve performance of algorithms, researchers often employ multiple views and many multi-view learning algorithms, which can use the information of all kinds of features or multiple views, have been proposed.

For unsupervised learning, clustering methods (Berkhin, 2002; Elhamifar & Vidal, 2013; Liu et al., 2013; Ng, Jordan, & Weiss, 2002; Zhao, Chen et al., 2019) have been used comprehensively to explore the underlying structure of data. In the past decades, many clustering methods have been developed and the most famous methods may be spectral clustering (SPC) (Ng et al., 2002), sparse subspace clustering (SSC) (Elhamifar & Vidal, 2013), and low-rank representation (LRR) (Liu, Lin et al., 2013).

The above clustering methods, however, are single-view clustering ones. To use in the multi-view clustering scenery, these clustering methods must concatenate features from the different

views into a single union. Then these clustering approaches cannot make an effective use of the multi-view information. To make full use of multi-view information, a lot of multi-view clustering approaches have been developed recently (Cao, Zhang, Fu, Liu, & Zhang, 2015; Xia, Pan, Du, & Yin, 2014; Xie et al., 2018; Zhang, Fu, Liu, Liu, & Cao, 2015).

Generally, the performances of multi-view clustering (MVC) approaches (Bickel & Scheffer, 2004; Kumar, Rai, & Hal Daumé, 2011; Yin, Gao, Xie, & Guo, 2019; Yin, Wu, & Wang, 2018), which can utilize the complementary information of objects from different views, are better than those of single-view approaches.

Sa (2005) proposed a two-view clustering method, called as SM-D. The SM-D algorithm first creates a bipartite graph, which is based on minimizing-disagreement criterion, to connect the two-view features. Then the conventional spectral clustering method is used to obtain the final clustering result. Xia et al. (2014) proposed a robust multi-view spectral clustering (RMSC) method, in which a transition probability matrix from each single view is constructed firstly, and then these matrices are used to recover a shared matrix. To remove the noise from different views, low-rank and sparse decomposition is also used in the RMSC method. The above methods are both graph-based ones, which are closely related to multiple kernel learning (MKL) technique. Tzortzis and Likas (2012) proposed a kernel-based weighted multi-view clustering method, in which views are first considered as given kernel matrices and then a weighted combination of these kernels is learned to partition the data.

By interchanging the partition information among different views, Bickel and Scheffer (2004) proposed a MVC method, which is a co-training style method. Based on spectral clustering, Kumar

[☆] This research is supported by NSFC of China (No. 61976005, 61572033, 61772277); the Major Project of Natural Science Research in Colleges and Universities of Anhui Province, China (Grant Number: KJ2019ZD15).

^{*} Corresponding author.

E-mail address: luguifu_tougao@163.com (G.-F. Lu).

and III (2011) proposed another MVC method, which uses the spectral embedding from one view to train the similarity graph used for the other view. Cao et al. (2015) proposed a diversity-induced multi-view subspace clustering (DiMSC) method, which uses the Hilbert Schmidt independence criterion (HSIC) as a diversity term to explore the complementary information among views. Recently, Zhao et al. (Zhao, Wang, Miao, Xu, & Zhang, 2018; Zhao et al., 2017; Zhao, Wang, Zuo, Shen and Shao, 2020) proposed some multi-view discriminant analysis methods to make use of the information in the different views. However, these methods are supervised methods since they utilize the discriminant information, and then these methods cannot be used in the MVC scenery.

The goal of subspace learning based MVC approaches, which are based on the assumption that all the views are generated from a latent subspace, is to capture shared latent subspace and then conduct clustering. Blaschko and Lampert (2008) proposed a subspace learning based MVC method, in which canonical correlation analysis (CCA) is used to project the multi-view high-dimensional data onto a low-dimensional subspace. White (Guo, 2013) developed a convex reformulation of multi-view subspace learning method, in which the squared loss used in CCA is replaced by robust loss, which enforces conditional independence between views.

Although the above mentioned MVC methods show good performances in MVC, they cannot make fully use of the properties of multi-view data. Generally, the higher order correlation (Peng, Li, & Shao, 2008) underlying the multi-view data cannot be utilized since these methods only focus on capturing the pairwise correlations between different views. In fact, the real-world data are ubiquitously in multi-dimension, which are often referred as tensors. Then, for multi-view data, the clustering performance is not optimal if the correlations in original spatial structure are not captured. To address this issue, Zhang et al. (2015) developed a low-rank tensor constrained multi-view subspace clustering (LT-MSC) method to explore the complementary information between different views. In the LT-MSC method, the subspace representation matrices of different views are regarded as a tensor. Besides, low-rank tensor constraint is used to capture the high order correlations underlying multi-view data. Based on the recently proposed tensor singular value decomposition (t-SVD) (Kilmer, Braman, Hao, & Hoover, 2013; Zhang & Xia, 2018; Zhou, Lu, Lin, & Zhang, 2018), Xie et al. (2018) proposed a t-SVD based multi-view subspace clustering (t-SVD-MSC) method. By using the augmented Lagrangian method, the minimization problem of t-SVD-MSC can be efficiently solved with low computational complexity.

The clustering performances of the traditional MVC method are desirable, however, these methods only consider global consensus among multiple views and ignore the view-specific local geometrical structure. In the real world, data often reside on low-dimensional manifolds embedded in a high-dimensional ambient space. Then, to discover the underlying manifold structure hidden in the high-dimensional data, some manifold learning methods have been developed, e.g., Isomap (Tenenbaum, Silva, & Langford, 2000), locally linear embedding (LLE) (Roweis & Saul, 2000), and Laplacian Eigenmap (LE) (Belkin & Niyogi, 2003). By introducing a regularization term based on the manifold structures of data, Yin, Gao, and Lin (2016) proposed a novel Laplacian regularized low-rank representation method (HRLRR), in which a nearest neighbor graph is used to model the local geometric structures of data. Then, the performance of LRR is further improved.

In this paper, to address the issue of LT-MSC, that is, it cannot make use of the view-specific local geometrical structure, we propose a novel hyper-Laplacian regularized multiview subspace clustering with low-rank tensor constraint method (HLR-MSCLRT). First, the subspace representation matrices of different

views are stacked as a tensor, and then the high order correlations among data can be captured. Second, a low-rank constraint is adopted to the constructed tensor to reduce the redundancy information of the learned subspace representations. Third, a hyper-Laplacian graph regularization term is used to preserve the local geometry structure embedded in a high-dimensional ambient space. An efficient algorithm is also presented to solve the optimization problem of the HLR-MSCLRT model. The experimental results on some data sets show the effectiveness of the proposed HLR-MSCLRT model.

The remainder of the paper is organized as follows. In Section 2, we introduce some related works. In Section 3, we propose the hyper-Laplacian regularized multiview subspace clustering with low-rank tensor constraint method (HLR-MSCLRT) and the efficient procedure for solving HLR-MSCLRT. We report the experiment results in Section 4. Finally, we conclude the paper in Section 5.

2. Related works

2.1. Low-rank tensor constrained multi-view subspace clustering (LT-MSC)

An M -way (or M -mode) tensor is a multilinear structure in $R^{I_1 \times I_2 \times \dots \times I_M}$. The tensor nuclear norm (Liu, Musialski, Wonka, & Ye, 2009; Liu, Musialski, Wonka and Ye, 2013) is defined as

$$\|\mathbf{Z}\|_* = \sum_{m=1}^M \xi_m \|Z_{(m)}\|_* \quad (1)$$

where ξ_m is constant and satisfies $\xi_m > 0$ and $\sum_{m=1}^M \xi_m = 1$. $\mathbf{Z} \in R^{n_1 \times n_2 \times \dots \times n_M}$ is an M -order tensor, and $Z_{(m)}$ is the matrix by unfolding the tensor $\mathbf{Z}$ along the m th mode defined as

$$\text{unfold}(\mathbf{Z}) = Z_m \in R^{I_m \times (I_1 \times \dots \times I_{m-1} \times I_{m+1} \times \dots \times I_M)} \quad (2)$$

$\|\cdot\|_*$ denotes the nuclear norm of a tensor or a matrix. The matrix nuclear norm and tensor nuclear norm are, respectively, the sum of singular values of a matrix and a convex combination of the nuclear norms of all matrices unfolded along each mode. The objective function of LT-MSC is defined as

$$\begin{aligned} \min_{Z^{(v)}, E^{(v)}} \quad & \|\mathbf{Z}\|_* + \lambda \|E\|_{2,1} \\ \text{s.t.} \quad & X^{(v)} = X^{(v)} Z^{(v)} + E^{(v)}, v = 1, 2, \dots, V, \\ & \mathbf{Z} = \Psi(Z^{(1)}, Z^{(2)}, \dots, Z^{(V)}), \\ & E = [E^{(1)}; E^{(2)}; \dots; E^{(V)}] \end{aligned} \quad (3)$$

where $X^{(v)}$, $E^{(v)}$ and $Z^{(v)}$ are, respectively, the data matrix, the reconstruction error matrix and the subspace representation matrix for the v th view. $\Psi(\cdot)$ constructs the tensor $\mathbf{Z}$ by merging the different representation $Z^{(v)}$ to a 3-order tensor, the dimensionality of which is $N \times N \times V$. $E = [E^{(1)}; E^{(2)}; \dots; E^{(V)}]$ is formed by vertically concatenating together along the column of errors corresponding to each view. $\|\cdot\|_{2,1}$ denotes the $l_{2,1}$ -norm.

2.2. Hypergraph preliminaries

Given a hypergraph $G = (V, E, W)$, V denotes a set of vertices, E denotes the hyperedge set e of V such that $\cup_{e \in E} e = V$, and the weight $w(e)$, which is positive and the element of weight matrix $\mathbf{W}$, is associated with each hyperedge e . An incidence matrix $H \in R^{|V| \times |E|}$ represents the relationship between the vertices and the hyperedges and its entry is defined as

$$h(\mathbf{v}_i, \mathbf{e}_j) = \begin{cases} 1, & \text{if } \mathbf{v}_i \in \mathbf{e}_j \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Let $d(\mathbf{e}_j) = \sum_{\mathbf{v}_i \in V} h(\mathbf{v}_i, \mathbf{e}_j)$ be the degree of a hyperedge $\mathbf{e}_j$ and D_E be the diagonal degree matrix whose diagonal entries correspond to the degree of each hyperedge. Similarly, let D_V be the diagonal matrix whose diagonal entries correspond to the degree of each vertex. The unnormalized hyper-Laplacian matrix is (Zhou, Huang, & Schölkopf, 2006) defined as

$$L_h = D_V - HWD_E^{-1}H^T \quad (5)$$

3. Hyper-Laplacian regularized multiview subspace clustering with low-rank tensor constraint

In this section, a novel hyper-Laplacian regularized multiview subspace clustering with low-rank tensor constraint method (HLR-MSCLRT) is first proposed, in which the local manifold structures of given data are considered. Then, we present an efficient procedure for solving HLR-MSCLRT.

3.1. Problem formulation

The objective function of HLR-MSCLRT can be formulated as

$$\begin{aligned} \min_{Z^{(v)}, E^{(v)}} & \|Z\|_* + \lambda_1 \|E\|_{2,1} + \lambda_2 \sum_{v=1}^V \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right) \\ \text{s.t. } & X^{(v)} = X^{(v)} Z^{(v)} + E^{(v)}, v = 1, 2, \dots, V, \\ & Z = \Psi(Z^{(1)}, Z^{(2)}, \dots, Z^{(V)}), \\ & E = [E^{(1)}; E^{(2)}; \dots; E^{(V)}] \end{aligned} \quad (6)$$

where $L_h^{(v)}$ is the view-specific hyper-Laplacian matrix which is build on $Z^{(v)}$. The hyper-Laplacian regularized term, i.e., $\sum_{v=1}^V \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right)$, is based on the assumption that if two data points are close in the intrinsic geometry of the data space, their mappings are also close to each other in a new space. By introducing a hyper-Laplacian regularization term, we can effectively keep the local geometry structure embedded in a high-dimensional ambient space.

By using Eqs. (1), (6) can be converted into

$$\begin{aligned} \min_{Z^{(v)}, E^{(v)}} & \sum_{m=1}^M \gamma_m \|Z_{(m)}\|_* + \|E\|_{2,1} + \gamma \sum_{v=1}^V \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right) \\ \text{s.t. } & X^{(v)} = X^{(v)} Z^{(v)} + E^{(v)}, v = 1, 2, \dots, V, \\ & Z = \Psi(Z^{(1)}, Z^{(2)}, \dots, Z^{(V)}), \\ & E = [E^{(1)}; E^{(2)}; \dots; E^{(V)}] \end{aligned} \quad (7)$$

where $\gamma_m = \frac{\xi_m}{\lambda_1} > 0$ and $\gamma = \frac{\lambda_2}{\lambda_1}$.

The optimization problem Eq. (7) can be solved via the inexact augmented Lagrange multiplier (ALM) (Lin, Liu, & Su, 2011). To adopt the ALM method to our problem, i.e., (7), our objective function should be separable. Then, we introduce auxiliary variables G_m to replace $Z_{(m)}$ and have

$$\begin{aligned} \min_{Z^{(v)}, E^{(v)}} & \sum_{m=1}^M \gamma_m \|G_m\|_* + \|E\|_{2,1} + \gamma \sum_{v=1}^V \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right) \\ \text{s.t. } & P_m Z = G_m, m = 1, 2, \dots, M \\ & X^{(v)} = X^{(v)} Z^{(v)} + E^{(v)}, v = 1, 2, \dots, V, \\ & Z = \Psi(Z^{(1)}, Z^{(2)}, \dots, Z^{(V)}), \\ & E = [E^{(1)}; E^{(2)}; \dots; E^{(V)}] \end{aligned} \quad (8)$$

where $\mathbf{z}$ and $\mathbf{g}_m$ are, respectively, the vectorizations of the tensor $\mathbf{Z}$ and the matrix G_m . P_m , which is a permutation matrix used to align the corresponding elements between $Z_{(m)}$ and G_m , is the alignment matrix corresponding to the mode- m unfolding. By

using the ALM method, Eq. (8) can be reformulated as

$$\begin{aligned} L(Z^{(1)}, \dots, Z^{(V)}; E^{(1)}, \dots, E^{(V)}; G_1, \dots, G_M) = \\ \|E\|_{2,1} + \sum_{m=1}^M \left(\gamma_m \|G_m\|_* + \langle \alpha_m, P_m \mathbf{z} - \mathbf{g}_m \rangle + \frac{\mu}{2} \|P_m \mathbf{z} - \mathbf{g}_m\|_F^2 \right) + \\ \sum_{v=1}^V \left(\gamma \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right) + \langle Y_v^T, X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)} \rangle \right. \\ \left. + \frac{\mu}{2} \|X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)}\|_F^2 \right) \end{aligned} \quad (9)$$

where α_m and Y_v are the Lagrange multipliers, μ is the penalty parameter.

3.2. Optimization

The optimization process of Eq. (9) could be separated into the following steps.

$Z^{(v)}$ -subproblem: When $E^{(v)}$ and G_m are fixed, $Z^{(v)}$ can be obtained by solving the following subproblem:

$$\begin{aligned} Z^{(v)*} = \arg \min_{Z^{(v)}} & \sum_{m=1}^M \left(\langle \alpha_m, P_m \mathbf{z} - \mathbf{g}_m \rangle + \frac{\mu}{2} \|P_m \mathbf{z} - \mathbf{g}_m\|_F^2 \right) + \\ & \sum_{v=1}^V \left(\gamma \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right) + \langle Y_v^T, X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)} \rangle \right. \\ & \left. + \frac{\mu}{2} \|X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)}\|_F^2 \right) \\ = \arg \min_{Z^{(v)}} & \sum_{m=1}^M \left(\langle \Omega^v(\alpha_m), \Omega^v(P_m \mathbf{z} - \mathbf{g}_m) \rangle + \frac{\mu}{2} \|\Omega^v(P_m \mathbf{z} - \mathbf{g}_m)\|_F^2 \right) \\ & + \sum_{v=1}^V \left(\gamma \text{tr} \left(Z^{(v)} L_h^{(v)} Z^{(v)T} \right) + \langle Y_v^T, X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)} \rangle \right. \\ & \left. + \frac{\mu}{2} \|X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)}\|_F^2 \right) \end{aligned} \quad (10)$$

where $\Omega^v(\cdot)$ represents the operation which selects the elements corresponding to the v th view and then reshapes these elements into a matrix. According to Eq. (10), $Z^{(v)}$ can be computed as:

$$\begin{aligned} Z^{(v)*} = & \left(\sum_{m=1}^M B_m^{(v)} - \sum_{m=1}^M A_m^{(v)} + X^{(v)T} X^{(v)} - X^{(v)T} E^{(v)} \right. \\ & \left. + X^{(v)T} Y_v \right) \left(MI + X^{(v)T} X^{(v)} + 2\gamma L_h^{(v)} \right)^{-1} \end{aligned} \quad (11)$$

where $A_m^{(v)} = \Omega^v(\alpha_m)$ and $B_m^{(v)} = \Omega^v(\mathbf{g}_m)$.

$\mathbf{z}$ -subproblem: After all $Z^{(v)}$ s of multiple views are computed, we can update $\mathbf{z}$ by first constructing the tensor $\mathbf{Z}$ and then vectorizing $\mathbf{Z}$.

E -subproblem: When $Z^{(v)}$ and G_m are fixed, E can be obtained by solving the following subproblem:

$$\begin{aligned} E^* = \arg \min_E & \|E\|_{2,1} \\ & + \sum_{v=1}^V \left(\langle Y_v^T, X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)} \rangle \right. \\ & \left. + \frac{\mu}{2} \|X^{(v)} - X^{(v)} Z^{(v)} - E^{(v)}\|_F^2 \right) \\ = \arg \min_E & \frac{1}{\mu} \|E\|_{2,1} + \frac{1}{2} \|E - F\|_F^2 \end{aligned} \quad (12)$$

where F is built through vertically concatenating the matrices $X^{(v)} - X^{(v)} Z^{(v)} + Y^{(v)}$ together along column. Eq. (12) has the

following solution (Liu, Lin et al., 2013):

$$E_{:,i}^* = \begin{cases} \frac{\|F_{:,i}\|_2 - \frac{1}{\mu}}{\|F_{:,i}\|_2} \|F_{:,i}\|_2, & \|F_{:,i}\|_2 > \frac{1}{\mu} \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

where $F_{:,i}$ denotes the i th column of the matrix F .

Y_v -subproblem: The Lagrange multipliers Y_v can be updated as:

$$Y_v^* = Y_v + (X^{(v)} - X^{(v)}Z^{(v)} - E^{(v)}) \quad (14)$$

Algorithm 1: HLR-MSCLRT.

Input: Multiple types of feature matrices: $X^{(1)}, X^{(2)}, \dots, X^{(V)}$, parameters γ_m 's, γ , and the cluster number K

Output: Clustering result C .

1: Initialize : $Z^{(1)} = \dots = Z^{(V)} = 0$; $E^{(1)} = \dots = E^{(V)} = 0$;
 $Y_1 = \dots = Y_V = 0$; $\alpha_1 = \dots = \alpha_M = 0$; $G_1 = \dots = G_M = 0$;

$\mu = 10^{-6}$; $\rho = 1.1$; $\varepsilon = 10^{-8}$; $\mu_{\max} = 10^{10}$

Construct the matrices H , L and M ;

2: While not converged do

3: Construct hyper-Laplacian matrices $\{L_h^{(v)}\}_{v=1}^V$ from
 $\{Z^{(v)}\}_{v=1}^V$ by using Eq.(5);

// for Z

4: Update $\{Z^{(v)}\}_{v=1}^V$ according to Eq.(11);

// for E

5: Update $\{E^{(v)}\}_{v=1}^V$ according to Eq.(12);

// for Y

6: Update $\{Y_v\}_{v=1}^V$ according to Eq.(14);

7: Update $\mathbf{z}$;

8: For each of M modes do

Update G_m according to Eq.(15);

Update $\mathbf{g}_m$ according to Eq.(16);

Update α_m according to Eq.(17);

End

9: Update μ by $\mu = \min(\rho\mu; \mu_{\max})$;

10: Check the convergence conditions:

$$\|X^{(v)} - X^{(v)}Z^{(v)} - E^{(v)}\|_{\infty} < \varepsilon \quad \text{and}$$

$$\|P_m \mathbf{z} - \mathbf{g}_m\|_{\infty} < \varepsilon$$

end

11: Let $S = \frac{1}{V} \sum_{v=1}^V |Z^{(v)}| + |Z^{(v)^T}|$;

12: Apply spectral clustering on matrix S ;

Output: Clustering result C ;

G_m -subproblem: G_m can be updated as:

$$G_m^* = \arg \min_{G_m} \|G_m\|_* + \langle \alpha_m, P_m \mathbf{z} - \mathbf{g}_m \rangle + \frac{\mu}{2} \|P_m \mathbf{z} - \mathbf{g}_m\|_F^2 \\ = \text{prox}_{\beta_m}^{\text{tr}}(\Omega_m(P_m \mathbf{z} - \mathbf{g}_m)) \quad (15)$$

where $\Omega_m(\cdot)$ represents the operation which reshapes a vector to a corresponding matrix according to the m th mode unfolding, and $\text{prox}_{\beta_m}^{\text{tr}}(\cdot)$ denotes the spectral soft-threshold operation, i.e., $\text{prox}_{\beta_m}^{\text{tr}}(L) = U \max(S - \beta_m, 0) V^T$, where $L = USV^T$ is the singular value decomposition (SVD) of L and $\beta_m = \frac{\gamma_m}{\mu}$.

$\mathbf{g}_m$ -subproblem: Similar to updating $\mathbf{z}$, we update $\mathbf{g}_m$ by directly reshaping G_m , i.e.,

$$\mathbf{g}_m^* \leftarrow G_m \quad (16)$$

α_m -subproblem: Similar to updating Y_v , the Lagrange multipliers α_m can be updated as

$$\alpha_m^* \leftarrow \alpha_m + P_m \mathbf{z} - \mathbf{g}_m \quad (17)$$

Now the procedure for solving Eq. (9) is summarized in Algorithm 1.

3.3. Computational complexity

In the proposed method, the construction of $\{L_h^{(v)}\}_{v=1}^V$ and the computations of $Z^{(v)}$, E and G_m are the four main computation-intensive steps. As for the hyper-Laplacian, it takes $O(VN^2 \log(N))$ for all the views. As for $Z^{(v)}$, it takes $O(VN^3)$ for all views. As for E , it takes $O(VN^2)$ in each iteration. As for G_m , it also takes $O(VN^3)$ for all views. Then, the total computational complexity of the proposed methods is

$$O(K(VN^2 \log(N) + VN^2)) + O(N^3) \quad (18)$$

where K is the number of iterations and $O(N^3)$ is the cost of spectral clustering.

3.4. Convergence analysis

In Lin, Chen, and Ma (2009) and Lin et al. (2011), the authors have proved the strong convergence of the ALM method with two blocks. However, when the number of block is bigger than 2, the convergence property of the ALM method is still unknown. The objective function of (6) is not smooth, and in HLR-MSCLRT, there are blocks $\{Z^{(v)}\}_{v=1}^V$, $\{E^{(v)}\}_{v=1}^V$ and $\{G_m\}_{m=1}^M$, it is difficult to prove the global convergence of HLR-MSCLRT theoretically.

However, we can prove HLR-MSCLRT converges the local optimal point of the objective function, i.e., Eq. (7). This is because: (1) each update of different variables is convex and has closed-form solution. Then, we can conclude the objective function is monotonically decreasing towards a stationary point. (2) The objective is bounded from below. Therefore, the proposed optimization approach converges the local optimal point of the objective function. In Section 4.5, Fig. 9 shows the convergence curves of the proposed method on Yale database and we can find that the proposed procedure for solving HLR-MSCLRT converges fast.

4. Experimental results

In this section, we will compare the performances of our proposed HLR-MSCLRT with some state-of-the-art methods. Specifically, the compared methods include the following ones:

SPC_{best} (Ng et al., 2002): The standard spectral clustering algorithm with the most informative view;

LR_{best} (Liu, Lin et al., 2013): LRR algorithm with the most informative view;

DiMSC (Cao et al., 2015): Diversity-induced multi-view subspace clustering;

Co-Reg SPC (Kumar et al., 2011): Spectral clustering algorithm which uses the co-training manner within the spectral clustering framework;

LT-MSC (Zhang et al., 2015): low-rank tensor constrained multi-view subspace clustering;

In the above methods, the first two methods are the single view ones and the other three methods are multi-view ones.

4.1. Dataset descriptions

In our experiments, five image dataset, i.e., ORL, Yale, Extended YaleB, COIL-20 and MSRC-v1, which are widely used in face and image clustering, are used to test the performances of different method.

ORL: The ORL face database consists of a total of 400 face images, of a total of 40 people (10 samples per person). For



Fig. 1. Images of one person in ORL.



Fig. 2. Images of one person in Yale.



Fig. 3. Images of one person in Extended YaleB.



Fig. 4. Some example images in COIL-20 image database.



Fig. 5. Some example images in MSRC-v1 image database.

some subjects, the images were taken at different times, varying the lighting, facial expressions (open/closed eyes, smiling/not smiling) and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, front position (with tolerance for some side movement). In our experiments, each image in ORL database was manually cropped and resized to 64×64 . Some example images of one person are shown in Fig. 1.

Yale face image database: The Yale face database contains 165 gray scale images of 15 individuals, each individual has 11 images. The images demonstrate variations in lighting condition, facial

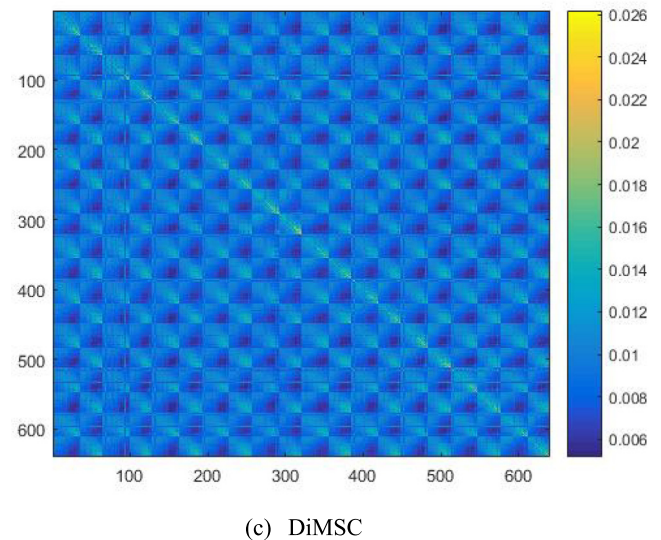
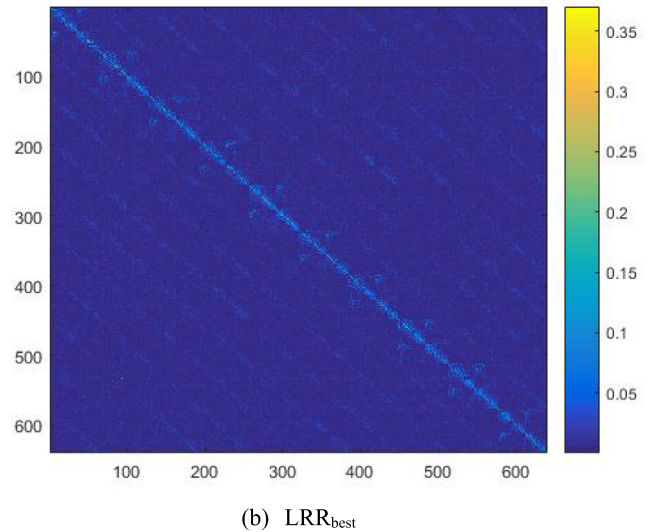
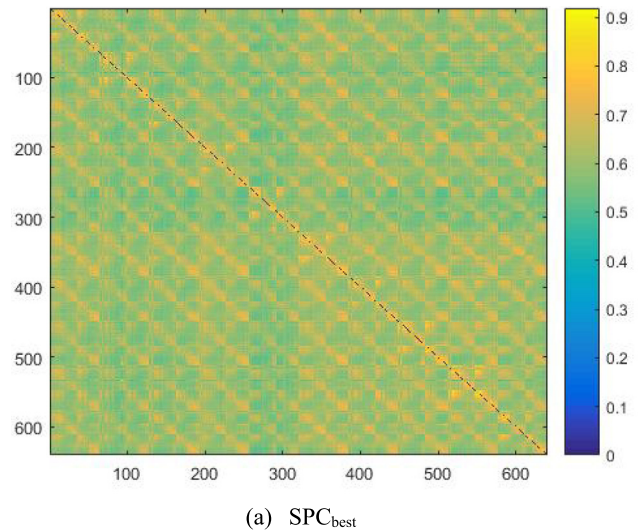
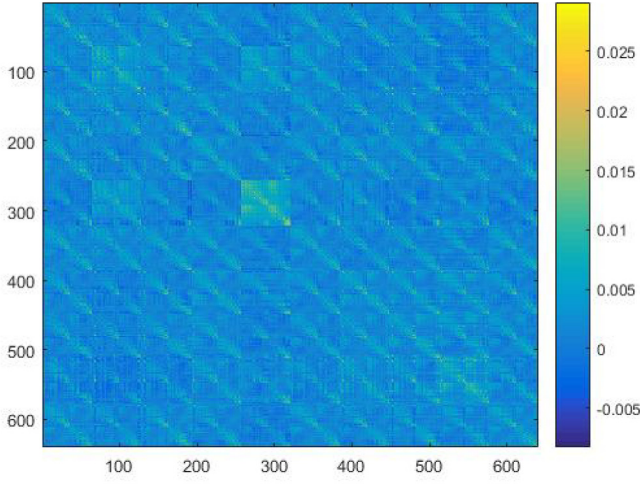
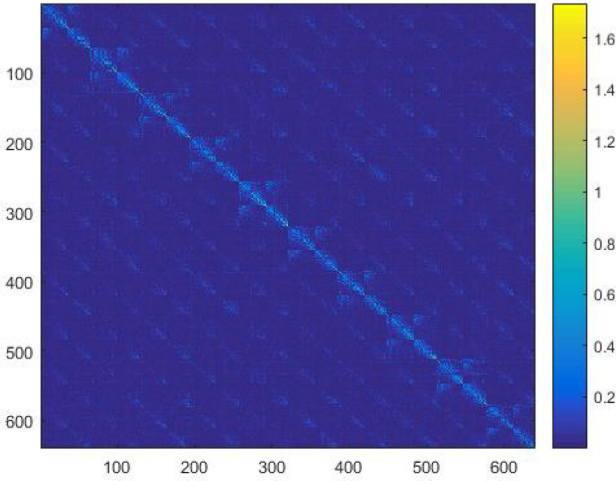


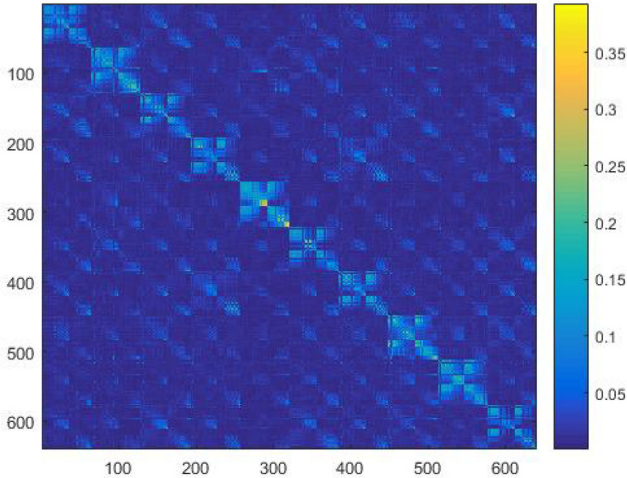
Fig. 6. The affinity matrices obtained by different methods on Extended YaleB database.



(d) Co-Reg SPC



(e) LT-MS



(f) HLR-MSCLRT

Fig. 6. (continued).

expression (normal, happy, sad, sleepy, surprised, and wink). In our experiments, each image in Yale database was resized to 64×64 . Fig. 2 shows sample images of one person.

Extended YaleB: The Extended YaleB dataset consists of 38 individuals and around 64 near frontal images under different illuminations for each individual. Similar to Zhang et al. (2015), we also use the images for the first 10 classes, including 640 frontal face images, which are resized to 32×32 . Fig. 3 shows sample images of one person.

COIL-20: The Columbia Object Image Library (COIL-20) dataset contains 1440 images of 20 object categories. Each category contains 72 images. All the images are resized to 32×32 . Fig. 4 shows sample images in COIL-20 image database.

MSRC-v1: The dataset is comprised of 240 images and it is divided into 9 classes. Similar to Lee and Grauman (2009), 7 classes, which are composed of tree, building, airplane, cow, face, car, bicycle, are selected in our experiments and each class has 30 images. Fig. 5 shows sample images in MSRC-v1 image database.

For ORL, Yale, Extended YaleB and COIL-20 database, similar to Zhang et al. (2015), three types of features, i.e., intensity, LBP (Ojala, Pietikäinen, & Mäenpää, 2002) and Gabor (Lades et al., 1993), are extracted. The standard LBP features are extracted with the sampling size of 8 pixels, and the blocking number of 7×8 . Therefore, the dimensionality of LBP is 3304. The Gabor feature is extracted with one scale $\lambda = 4$ at four orientations $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. Therefore, the dimensionality of Gabor is 6750. For MSRC-v1 database, similar to Lee and Grauman (2009), 256 Local Binary Pattern (LBP), 100 Histogram of Oriented Gradient (HOG), 512 GIST, 1302 CENTRIST, 48 Color Moment (CMT) and 200 SIFT features are extracted to distinguish all of scenes.

On all the five datasets, the M parameters, i.e., γ_m , $m = 1, \dots, M$, are set with equal value. That is, $\gamma_1 = \dots = \gamma_m$. Then, there are two parameters, i.e., γ_m , $m = 1, \dots, M$ and γ , to be tuned in our proposed method. In this paper, the values of γ_m and γ are, respectively, within the range $[0.01, 0.1, 1, 10, 100]$ and $[0.1, 0.9]$. The k -nearest neighbor is used to construct the hyperedges and the value of k is fixed to 5 in all experiments.

4.2. Evaluation metrics

Two popular evaluation metrics, i.e., the clustering accuracy (AC) and the normalized mutual information (NMI), are used to measure the clustering performance of the proposed method. AC is defined as

$$AC = \frac{\sum_{i=1}^n \delta(gnd_i, map(r_i))}{n} \quad (19)$$

where gnd_i is the label given by the datasets, r_i is the label provided by the clustering algorithm, $map(r_i)$ is the permutation mapping function which maps r_i to the equivalent label and $\delta(x, y)$ is the delta function whose value is 1 if $x = y$ and 0 otherwise.

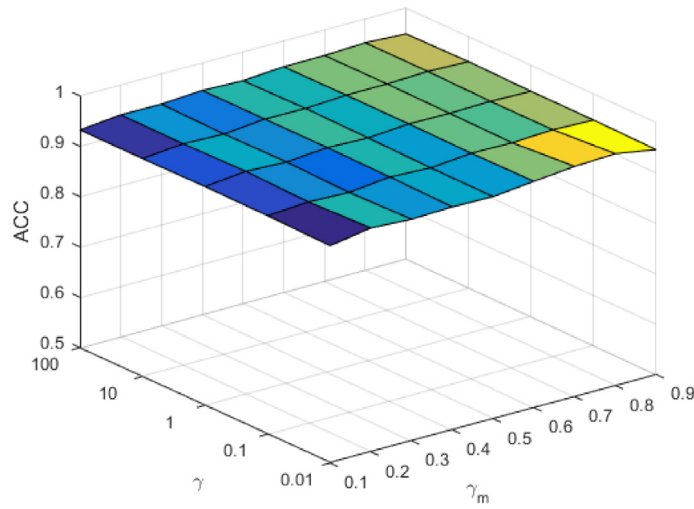
Mutual information (MI), which can measure the similarity between two cluster sets, is widely used in clustering applications. Let C and C' be two sets of clusters, and then MI is defined as

$$MI(C, C') = \sum_{c_i \in C, c'_j \in C'} p(c_i, c'_j) \log \frac{p(c_i, c'_j)}{p(c_i)p(c'_j)} \quad (20)$$

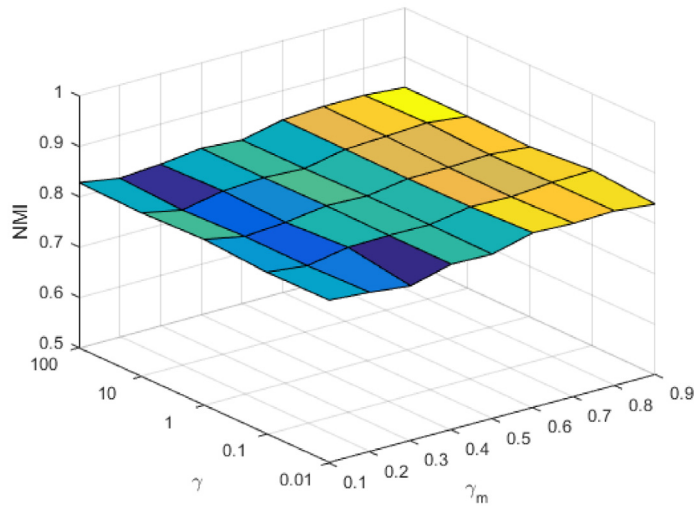
where $p(c_i)$ and $p(c'_j)$ denote the probabilities of a sample belonging to the clusters C and C' , respectively. $p(c_i, c'_j)$ denotes the joint probability that a sample belongs to C and C' simultaneously. Then NMI is defined as

$$NMI(C, C') = \frac{MI(C, C')}{\max(H(C), H(C'))} \quad (21)$$

where $H(C)$ and $H(C')$ are, respectively, the entropies of C and C' .



(a) ACC of HLR-MSCLRT versus the parameters γ_m and γ on the ORL database



(b) NMI of HLR-MSCLRT versus the parameters γ_m and γ on the ORL database

Fig. 7. Parameters tuning in terms of ACC and NMI on ORL and MSRC-v1 datasets.

Table 1
Results (mean $\pm$ standard deviation) on ORL.

Method	SPC _{best}	LRR _{best}	DiMSC
NMI	0.884 $\pm$ 0.003	0.895 $\pm$ 0.007	0.939 $\pm$ 0.006
ACC	0.726 $\pm$ 0.002	0.773 $\pm$ 0.003	0.837 $\pm$ 0.008
Method	Co-Reg SPC	LT-MSC	HLR-MSCLRT
NMI	0.853 $\pm$ 0.003	0.930 $\pm$ 0.002	0.962 $\pm$ 0.007
ACC	0.715 $\pm$ 0.001	0.795 $\pm$ 0.007	0.844 $\pm$ 0.004

Table 2
Results (mean $\pm$ standard deviation) on Yale.

Method	SPC _{best}	LRR _{best}	DiMSC
NMI	0.646 $\pm$ 0.009	0.706 $\pm$ 0.002	0.728 $\pm$ 0.003
ACC	0.634 $\pm$ 0.011	0.703 $\pm$ 0.000	0.703 $\pm$ 0.004
Method	Co-Reg SPC	LT-MSC	HLR-MSCLRT
NMI	0.715 $\pm$ 0.006	0.742 $\pm$ 0.001	0.778 $\pm$ 0.002
ACC	0.668 $\pm$ 0.004	0.721 $\pm$ 0.002	0.739 $\pm$ 0.002

4.3. Experiment results

In this paper, we report the average performances and standard derivations of the compared algorithms, i.e., HLR-MSCLRT, SPC_{best} (Ng et al., 2002), LRR_{best} (Liu, Lin et al., 2013), DiMSC (Cao et al., 2015), Co-Reg SPC (Kumar et al., 2011) and LT-MSC (Zhang et al., 2015) by running these methods 30 times since these methods are generally equipped with k -means.

The detail clustering results on five datasets are reported in Tables 1–5.

From the above experiment results, we can make the following observations:

- (1) On ORL, Yale and MSRC-v1 datasets, the performances of single view methods, i.e., SPC_{best} and LRR_{best} are inferior to those of multi-view ones, i.e., DiMSC, Co-Reg SPC, LT-MSC and HLR-MSCLRT. Then, utilizing the information of all the views can improve the clustering performance. Besides, the clustering performances of the tensor-based methods,

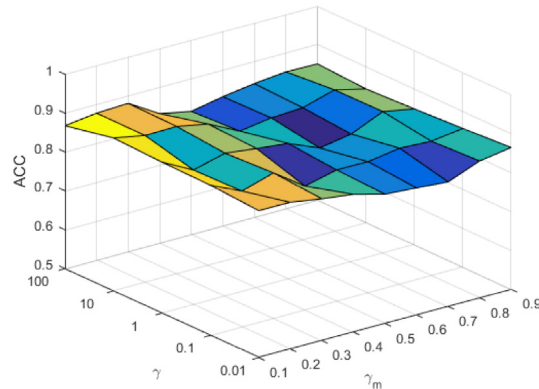
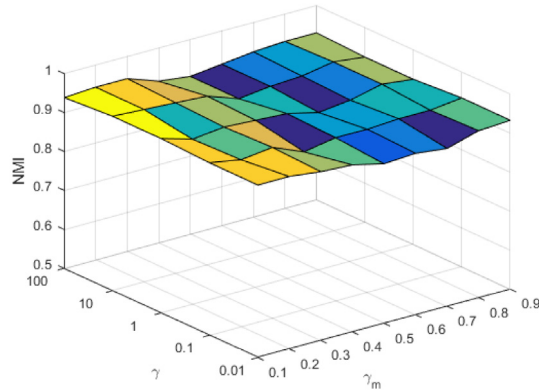
(c) ACC of HLR-MSCLRT versus the parameters γ_m and γ on the MSRC-v1 database(d) NMI of HLR-MSCLRT versus the parameters γ_m and γ on the MSRC-v1 database

Fig. 7. (continued).

Table 3Results (mean $\pm$ standard deviation) on Extended YaleB.

Method	SPC _{best}	LRR _{best}	DiMSC
NMI	0.223 $\pm$ 0.002	0.408 $\pm$ 0.000	0.395 $\pm$ 0.003
ACC	0.281 $\pm$ 0.002	0.447 $\pm$ 0.000	0.468 $\pm$ 0.003
Method	Co-Reg SPC	LT-MSC	HLR-MSCLRT
NMI	0.146 $\pm$ 0.001	0.571 $\pm$ 0.000	0.636 $\pm$ 0.001
ACC	0.239 $\pm$ 0.002	0.562 $\pm$ 0.000	0.613 $\pm$ 0.001

Table 4Results (mean $\pm$ standard deviation) on COIL-20.

Method	SPC _{best}	LRR _{best}	DiMSC
NMI	0.770 $\pm$ 0.009	0.868 $\pm$ 0.006	0.846 $\pm$ 0.008
ACC	0.682 $\pm$ 0.009	0.767 $\pm$ 0.004	0.774 $\pm$ 0.008
Method	Co-Reg SPC	LT-MSC	HLR-MSCLRT
NMI	0.809 $\pm$ 0.009	0.812 $\pm$ 0.004	0.916 $\pm$ 0.005
ACC	0.720 $\pm$ 0.010	0.735 $\pm$ 0.004	0.823 $\pm$ 0.005

i.e., LT-MSC and HLR-MSCLRT, are superior to the other ones. The reason may be that the tensor-based methods can make use of the high order information of data. Compared with LT-MSC, our proposed method, i.e., HLR-MSCLRT, gains significant improvement, which may be attributed to the utilization of view-specific local geometrical structure.

- (2) On Extended YaleB and COIL20 datasets, our proposed HLR-MSCLRT also gain significant improvement. On Extended YaleB, the matrix-based methods, i.e., SPC_{best}, LRR_{best}, DiMSC and Co-Reg SPC, achieve low performances. The

Table 5Results (mean $\pm$ standard deviation) on MSRC-v1.

Method	SPC _{best}	LRR _{best}	DiMSC
NMI	0.532 $\pm$ 0.001	0.592 $\pm$ 0.000	0.667 $\pm$ 0.001
ACC	0.654 $\pm$ 0.001	0.710 $\pm$ 0.000	0.733 $\pm$ 0.001
Method	Co-Reg SPC	LT-MSC	HLR-MSCLRT
NMI	0.609 $\pm$ 0.000	0.774 $\pm$ 0.000	0.893 $\pm$ 0.000
ACC	0.629 $\pm$ 0.000	0.810 $\pm$ 0.000	0.947 $\pm$ 0.000

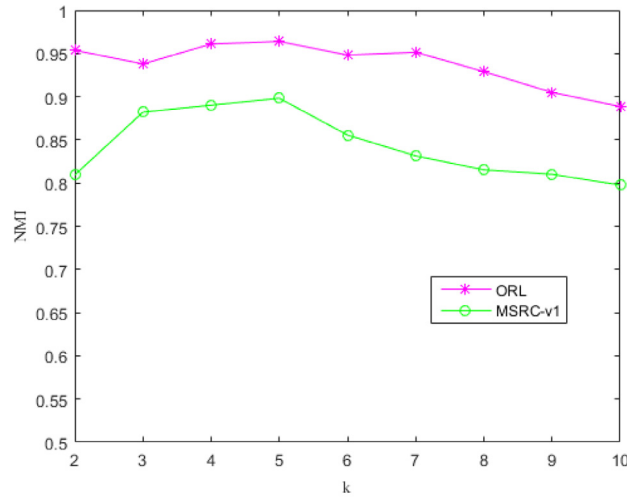
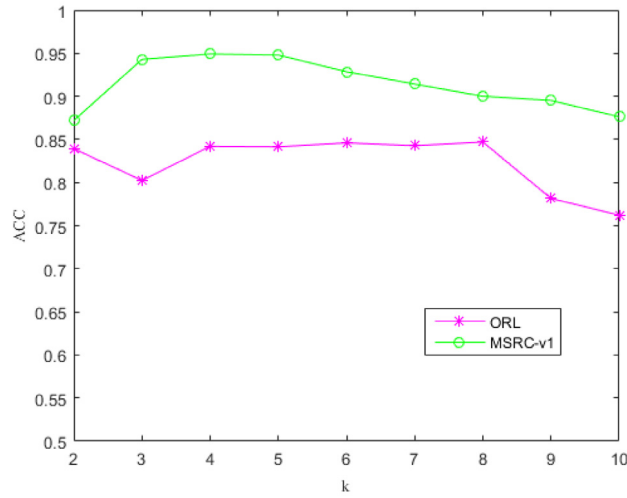
reason may be the large variation of illumination. However, the tensor-based method, i.e., LT-MSC and HLR-MSCLRT, significantly outperform the other methods because of the high order low-rank tensor constraint.

In the following, we will analyze the underlying reasons why the clustering performance of the proposed method is better than the other methods intuitively. In Fig. 6, we show the visualization of the learned affinity matrices. Due to the limitation of space, we only present the results on Extended YaleB.

From Fig. 6, we can easily see that the affinity matrix obtained by HLR-MSCLRT can better reflect the block-structure of data, which can benefit the clustering task. The affinity matrices obtained by the other algorithms, however, are somewhat unsatisfactory for the clustering task.

4.4. Parametric sensitivity

In our proposed HLR-MSCLRT model, there are $M + 1$ parameters that need to be tuned. In fact, there are only two parameters, i.e., γ_m , $m = 1, \dots, M$ and γ , which need to be tuned in our

(a) NMI of HLR-MSCLRT with different values of k on the ORL and MSRC-v1 datasets(b) ACC of HLR-MSCLRT with different values of k on the ORL and MSRC-v1 datasets**Fig. 8.** NMI and ACC with different values of k on ORL and MSRC-v1 datasets.

experiments since the M parameters, i.e., γ_m , $m = 1, \dots, M$, are set with equal value. In our experiments, the values of γ_m and γ are, respectively, within the range $[0.01, 0.1, 1, 10, 100]$ and $[0.1, 0.9]$. In Fig. 7, we show the experiment results of parameters tuning, i.e., γ_m and γ , in terms of ACC and NMI on ORL and MSRC-v1 datasets. In general, the performances of HLR-MSCLRT depend on the tuning of the regularization parameters of γ_m and γ . We can see that the performance on the MSRC-v1 is more sensitive to two regularization parameters than that on the ORL. The reason may be that the MSRC-v1 database contains more kinds of images, which contains the images of tree, building, airplane, cow, face, car and bicycle, than the ORL database, which only contains the face image. However, in general, the clustering performance of HLR-MSCLRT is relatively stable with respect to γ_m and γ .

In our method, the k -nearest neighbor is used to construct the hyperedges and the value of k also influences the performance of the proposed method. In Fig. 8, we show the experiment results with different values of k in terms of ACC and NMI on ORL and MSRC-v1 datasets.

From the experiment results, the clustering performance of HLR-MSCLRT is generally good when $k = 5$. Then the value of k is simply fixed to 5 in our experiments.

4.5. Convergence

According to the convergence conditions in Algorithm 1, i.e., Step 10 in HLR-MSCLRT, we define two error terms, i.e.,

$$RE = \frac{1}{V} \sum_{v=1}^V \|I - Z^{(v)} - P^{(v)}\|_{\infty} \quad (22)$$

and

$$ME = \frac{1}{M} \sum_{m=1}^M \|P_m \mathbf{z} - \mathbf{g}_m\|_{\infty} \quad (23)$$

where RE denotes the reconstruction error and ME denotes the match error.

In Fig. 9, we show the convergence curves of the proposed HLR-MSCLRT method on Yale database. We can find that HLR-MSCLRT converges fast. Usually, the number of optimization iteration is smaller than 40.

5. Conclusions

In this paper, a hyper-Laplacian regularized multiview subspace clustering with low-rank tensor constraint method (HLR-MSCLRT) is developed for clustering. In the proposed model,

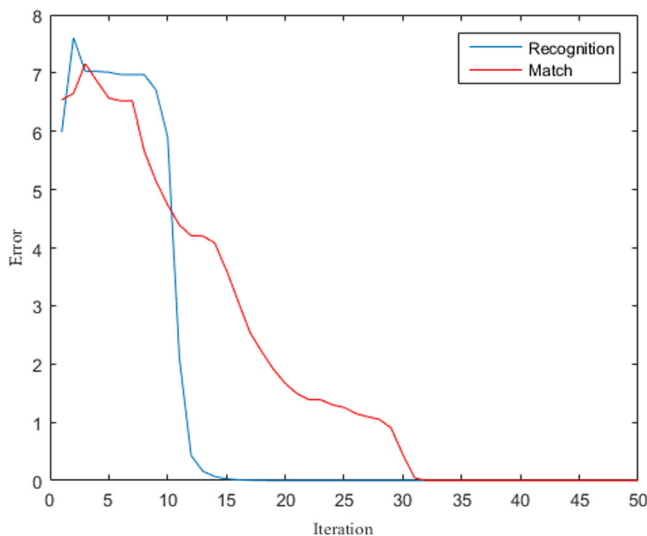


Fig. 9. Convergence curves on Yale database.

tensor is employed to capture the high order correlations among data. Besides, a hyper-Laplacian graph regularization is also used to explore the local geometry structure embedded in a high-dimensional ambient space. The model have been formulated into a unified optimization framework and an efficient procedure to find the solution is also proposed. The experimental results on some data sets show that the proposed HLR-MSCLRT model outperforms some state-of-the-art multi-view clustering approaches.

References

- Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6), 1373–1396.
- Berkhin, P. (2002). Survey of clustering data mining techniques. (pp. 1–56).
- Bickel, S., & Scheffer, T. (2004). Multi-view clustering. In *Proceedings of the fourth IEEE international conference on data mining (ICDM)* (pp. 19–26). Brighton, UK.
- Blaschko, M. B., & Lampert, C. H. (2008). Correlational spectral clustering. In *IEEE Conference on computer vision and pattern recognition* (pp. 1–8). Anchorage, AK, USA.
- Cao, X., Zhang, C., Fu, H., Liu, S., & Zhang, H. (2015). Diversity induced multi-view subspace clustering. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 586–594). Boston, MA, USA.
- Chao, G., Sun, S., & Bi, J. (2018). A survey on multi-view clustering. (pp. 1–17). arXiv preprint [arXiv:1712.06246](https://arxiv.org/abs/1712.06246).
- Elhamifar, E., & Vidal, R. (2013). Sparse subspace clustering: Algorithm, theory, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11), 2765–2781.
- Gu, J., & Li, P. (2018). Multi-view feature selection for heterogeneous face recognition. In *IEEE international conference on data mining (ICDM)* (pp. 983–988).
- Guo, Y. (2013). Convex subspace representation learning from multi-view data. In *proceedings of the twenty-seventh AAAI conference on artificial intelligence* (pp. 387–393). Bellevue, Washington, USA.
- Kilmer, M. E., Braman, K. S., Hao, N., & Hoover, R. C. (2013). Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM Journal on Matrix Analysis and Applications*, 34(1), 148–172.
- Kumar, A., & III, H. D. (2011). A co-training approach for multi-view spectral clustering. In *International conference on machine learning (ICML)* (pp. 393–400). Bellevue, Washington, USA.
- Kumar, A., Rai, P., & Hal Daumé, I. (2011). Co-regularized multi-view spectral clustering. In *Proceedings of the 24th international conference on neural information processing systems (NIPS)* (pp. 1413–1421). Granada, Spain.
- Lades, M., Vorbruggen, J. C., Buhmann, J., Lange, J., Malsburg, C. v. d., Wurtz, R. P., et al. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42(3), 300–311.
- Lee, Y. J., & Grauman, K. (2009). Foreground focus: Unsupervised learning from partially matching images. *International Journal of Computer Vision*, 85(2), 143–166.
- Lin, Z., Chen, M., & Ma, Y. (2009). The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. *UIUC Technica Report, Rep. UIUC-ENG-09-2215*.
- Lin, Z., Liu, R., & Su, Z. (2011). Linearized alternating direction method with adaptive penalty for low rank representation. In *Advances in neural information processing systems* (pp. 612–620). Granada, Spain.
- Liu, Z., Lai, Z., Ou, W., Zhang, K., & Zheng, R. (2020). Structured optimal graph based sparse feature extraction for semi-supervised learning. *Signal Processing*.
- Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., & Ma, Y. (2013). Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35, 171–184.
- Liu, J., Musialski, P., Wonka, P., & Ye, J. (2009). Tensor completion for estimating missing values in visual data. In *Proceedings in international conference on computer vision (ICCV)* (pp. 2114–2121).
- Liu, J., Musialski, P., Wonka, P., & Ye, J. (2013). Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1), 208–220.
- Ng, A. Y., Jordan, M. I., & Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In *Proceedings of the neural information processing systems* (pp. 849–856).
- Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987.
- Peng, W., Li, T., & Shao, B. (2008). Clustering multi-way data via adaptive subspace iteration. In *the 17th ACM conference on information and knowledge management* (pp. 1519–1520). Napa Valley, California, USA.
- Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimension reduction by locally linear embedding. *Science*, 290(5500), 2323–2326.
- Sa, V. R. d. (2005). Spectral clustering with two views. In *ICML workshop on learning with multiple views* (pp. 20–27).
- Sun, S. (2013). A survey of multi-view machine learning. *Neural Computing and Applications*, 23(7–8), 2031–2038.
- Tenenbaum, J. B., Silva, V. d., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(1), 2319–2323.
- Tzortzis, G., & Likas, A. (2012). Kernel-based weighted multi-view clustering. In *Proceedings of the IEEE international conference on data mining* (pp. 675–684). Brussels, Belgium.
- Xia, R., Pan, Y., Du, L., & Yin, J. (2014). Robust multi-view spectral clustering via low-rank and sparse decomposition. In *Proceedings of the twenty-eighth aaai conference on artificial intelligence* (pp. 2149–2155). Québec City, Québec, Canada.
- Xie, Y., Tao, D., Zhang, W., Liu, Y., Zhang, L., & Qu, Y. (2018). On unifying multi-view self-representations for clustering by tensor multi-task minimization. *International Journal of Computer Vision*, 126(11), 1157–1179.
- Xu, C., Tao, D., & Xu, C. (2013). A survey on multi-view learning. (pp. 1–59). 1–59, arXiv preprint [arXiv:1304.5634](https://arxiv.org/abs/1304.5634).
- Yin, M., Gao, J., & Lin, Z. (2016). Laplacian regularized low-rank representation and its applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3), 504–517.
- Yin, M., Gao, J., Xie, S., & Guo, Y. (2019). Multiview subspace clustering via tensorial t-Product representation. *IEEE Transactions on Neural Networks and Learning Systems*, 30(3), 851–864.
- Yin, Q., Wu, S., & Wang, L. (2018). Multiview clustering via unified and view-specific embeddings learning. *IEEE Transactions on Neural Networks and Learning Systems*, 29(11), 5541–5553.
- Zhang, C., Fu, H., Liu, S., Liu, G., & Cao, X. (2015). Low-rank tensor constrained multiview subspace clustering. In *IEEE international conference on computer vision (ICCV)* (pp. 1582–1590). Santiago, Chile.
- Zhang, A., & Xia, D. (2018). Tensor SVD: statistical and computational limits. *IEEE Transactions on Information Theory*, 64(11), 7311–7338.
- Zhao, C., Chen, K., Zang, D., Zhang, Z., Zuo, W., & Miao, D. (2019). Uncertainty-optimized deep learning model for small-scale person re-identification. *Sci. China Inf. Sci.*, 62(12), 1–13.
- Zhao, C., Wang, X., Miao, D., Xu, Y., & Zhang, D. (2018). Maximal granularity structure and generalized multi-view discriminant analysis for person re-identification. *Pattern Recognition*, 79, 79–96.
- Zhao, C., Wang, X., Wong, W. K., Zheng, W., Yang, J., & Miao, D. (2017). Multiple metric learning based on bar-shape descriptor for person re-identification. *Pattern Recognition*, 71, 218–234.
- Zhao, C., Wang, X., Zuo, W., Shen, F., & Shao, L. (2020). Similarity learning with joint transfer constraints for person re-identification. *Pattern Recognition*, 97, 1–10.
- Zhou, D., Huang, J., & Schölkopf, B. (2006). Learning with hypergraphs: Clustering, classification, and embedding. In *Advances in neural information processing systems* (pp. 1601–1608). Canada.
- Zhou, P., Lu, C., Lin, Z., & Zhang, C. (2018). Tensor factorization for low-rank tensor completion. *IEEE Transactions on Image Processing*, 27(3), 1152–1163.