Neurocomputing 118 (2013) 79–86

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Multi-view hypergraph learning by patch alignment framework

Chaoqun Hong [a], Jun Yu [b,c,*], Jonathan Li [b,c], Xuhui Chen [a]

[a] Faculty of Computer Science, Xiamen University of Technology, Xiamen, Fujian 361024, China
[b] Department of Computer Science, Xiamen University, Xiamen, Fujian 361005, China
[c] Key Laboratory for Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry of Education, Xiamen University, Xiamen, Fujian 361005, China

ARTICLE INFO

Article history:
Received 29 September 2012
Received in revised form
4 January 2013
Accepted 5 February 2013
Communicated by M. Wang
Available online 28 February 2013

Keywords:
Dimensionality reduction
Hypergraph
Multi-view learning
Patch alignment framework

ABSTRACT

Graph-based methods are currently popular for dimensionality reduction. However, most of them suffer from over-simplified assumption of pairwise relationships among data. Especially for multi-view data, different relationships from different views are hard to be integrated into a single graph. In this paper, we propose a novel semi-supervised dimensionality reduction method for multi-view data. First, we assume the hyperedges in hypergraph as patches and apply hypergraph to the patch alignment framework. Second, the weights of the hyperedges are computed with statistics of distances between neighboring pairs and the patches from different views are integrated. In this way, we construct Multi-view Hypergraph Laplacian matrix and we get the dimensionality-reduced data by solving the standard eigen-decomposition to obtain the projection matrix. The experimental results demonstrate the effectiveness of the proposed method on retrieval performance.

1. Introduction

Images or objects could be represented by several types of features in the related researches on computer vision. These features include color, shape, contour, texture and so on. Actually, varied features describe different properties of the same image. Since every single type of features could not completely describe one image, researchers have proposed learning methods by combining different types of features. These methods are called multi-view learning methods. Although combining multiple feature is not always beneficial [1], multi-view learning has attracted plenty of attention [2–4].

A great deal of efforts have been carried out to get better multi-view learning methods which are used in applications such as classification, retrieval, clustering and feature selection. However, the features representing images are usually in a high-dimensional space. This leads to the so-called "curse of dimensionality" problem. In this way, the consumption of both time and space in the learning process are influenced by the high-dimensional features.

Researchers have also devoted themselves into solving the "curse of dimensionality" problem by using dimensionality reduction approaches. Traditional well-known dimensionality reduction approaches include principle component analysis (PCA) [5]. It is unsupervised and does not consider the connectivity among different view. Therefore, PCA is not suitable for dimensionality reduction of multi-view data. Linear discriminant analysis (LDA) is another widely used approach [6]. It is supervised, but the global linearity of LDA prohibits its effectiveness for non-linear distributed measurements. Other researchers proposed manifold learning based dimensionality reduction approaches, including locally linear embedding (LLE) [7], ISOMAP [8], Laplacian eigenmaps (LE) [9], Hessian eigenmaps (HLLE) [10], and local tangent space alignment (LTSA) [11]. Among these approaches, Laplacian eigenmaps is a graph-based approach. It represents the images as vertices and the links between each pair of them as edges. If two vertices are connected by an edge, they may share some similar characteristics and they are called neighbors. In this way, a correlation graph can be constructed. Therefore, we can easily conclude that the key problem of graph-based dimensionality reduction approaches is how to construct the correlation graph. Most of the researches on graph-based dimensionality reduction approaches focus on this problem, such as Laplacian Regularization [12], Normalized Laplacian Regularization [13,14], Local Learning Regularization [15] and Markov random walk explanation [16]. It could also be combined with sparsity-based model to conduct semi-supervised learning [17]. Graph-based idea could describe different features in a unified form [18–20]. In this way, it has also been extended to dimensionality reduction for multi-view data [21]. The training results of these methods are matrices describing the structures of the correlation graphs. These matrices are called Laplacian matrices. However, graph-based approaches for multi-view learning usually encounter two problems.

* Corresponding author. Tel.: +86 18959285679.
E-mail address: yujun@xmu.edu.cn (J. Yu).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.02.017

- First, they always assume the relationships among images are pairwise. For example, the method by Yan et al. disposes most of the parameters, but the performance is still limited by the simplified assumption of pairwise relationships [22].
- Second, combining the correlations embedded in different views is difficult. The similarities of different features are measured by different criteria. Most of the approaches require a recursive refining process to get a reasonable combination of criteria.

To avoid the over-simplified assumption of pairwise relationships among data, researchers further proposed hypergraph. Hypergraph representation is becoming more and more popular and now widely used in many applications, such as classification [23], image segmentation [24] and video object segmentation [25]. Unlike a simple graph that has an edge between two vertices, a set of vertices is connected by a hyperedge in a hypergraph. Each hyperedge is assigned a weight. When constructing hypergraph, computing the weights of the hyperedges is critical. It significantly influence the description power of the features after dimensionality reduction. For example, the weight of each hyperedge is simply set to 1 in [26]. In [27], the weight of a hyperedge is calculated by summing up the pairwise affinities within the hyperedge. In practice, we are usually faced with a large number of hyperedges, and these hyperedges have different effects. For example, in the handwriting digit classification [26], a set of hyperedges is generated for each pixel; thus, some hyperedges are redundant. Therefore, weighting or selecting hyperedges will help improve classification performance.

In this paper, we propose a novel dimensionality reduction method for multi-view data based on patch alignment framework, which is named as Multi-view Hypergraph Learning (MHL). The contribution of our method is two-fold.

- We introduce hypergraph construction to the part optimization in patch alignment framework. This process is based on a real-valued form of combinatorial optimization problem in constructing hypergraph. The weights of hyperedges for the whole alignment are computed by statistics of distances between neighboring pairs.
- We apply the novel hypergraph to semi-supervised dimensionality reduction of multi-view data. The hyperedges computed with data of different views are integrated together and the dimensionality-reduced data are embedded in the integrated Laplacian matrix.

The rest of this paper is organized as follows. In Section 2, we review the work related to our research including hypergraph and patch alignment framework. In Section 3, the definitions of hypergraph are summarized. In Section 4, we introduce the proposed dimensionality reduction method. Theoretical derivation, algorithm details and analysis are all contained. In Section 5, we show the experimental results by comparing the proposed method with the previous methods. Finally, we show our conclusion the paper in Section 6.

2. Related work

2.1. Hypergraph

In traditional machine learning problem settings with graph-based idea or subspace using manifold assumption, the relationships among objects are usually assumed to be pairwise [28–32]. These objects and their relationships can be described by graphs. However, one edge links only two vertices in traditional graph-based representations. If more than two objects share the same characteristics, more than one edge is needed [26]. To avoid this problem, the hypergraph representation is proposed [33]. Different from the traditional graph-base representation, one edge is able to connect more than two vertices in the hypergraph representation. In other words, vertices connected by an edge are thought as a subset of vertices in the graph. In this way, the hypergraph representation is much more descriptive and powerful than traditional graph representations. Hypergraph representation is now widely used in many applications, such as classification [23,34,35], image segmentation [24], video object segmentation [25], tag-based image search [37–40] and retrieval [41,42].

2.2. Patch alignment framework

Patch alignment framework was proposed by Zhang et al. [43]. It unifies spectral analysis based dimensionality reduction approaches, including LLE/NPE/ONPP, ISOMAP, LE/LPP, LTSA/LLTSA, HLLE, PCA and LDA. It is proposed as a powerful analysis and development tool for dimensionality reduction. It consists of two stages. In the part optimization stage, different approaches have different optimization criteria over patches, each of which is built by one measurement associated with its related ones. In the whole alignment stage, all part optimizations are integrated to form the final global coordinate for all independent patches based on the alignment trick, originally used by Zhang and Zha. Different algorithms were shown to construct whole alignment matrices in an almost identical way, but vary with patch optimizations. Based on patch alignment framework, Zhang further proposed Discriminative Locality Alignment (DLA) for dimensionality reduction [44]. It uses KNN to discover relationships among data. DLA is supervised and could also be extended to a semi-supervised approach. In Xia et al.'s work, DLA is used as an unsupervised approach of dimensionality reduction for multi-view data [45]. Yu et al. [36] proposed a semi-supervised patch alignment framework, and applied it to solve the problem of correspondence construction for cartoon animation.

3. Hypergraph for semantic representations

As has been mentioned in the introduction, graph-based representation is widely used in dimensionality reduction algorithms. These algorithms usually assume the relationships among images are pairwise. However, we can only achieve which pairs of images are similar in the graph but we know nothing about the details of the properties they shared. If more than two objects sharing the same properties, more than one edge is required to connect them. Therefore, this assumption is over-simplified and plenty of information is lost while constructing graphs. Assume there are 7 images in the data set. They may contain a flower, a dog or a man, as is shown in Table 1. To describe their relationships, we construct a graph in

Table 1
Images in the data set.

Images	Contains a flower	Contains a dog	Contains a man
I_1	Yes	No	Yes
I_2	Yes	No	No
I_3	No	No	Yes
I_4	No	Yes	Yes
I_5	No	Yes	No
I_6	Yes	Yes	No
I_7	No	No	Yes

Fig. 1. The vertices in this figure are images in the data set. If two images contain the same object, they will be connected by an edge. This representation seems natural, but we miss the information about the objects that images contain. For example, we cannot know which images have trees in this graph. Such information is very useful in dimensionality reduction algorithms for multi-view data, since different views reflect different properties of the images and the relationships on different views lead to different clustering standards. In contrast to the traditional graph, hypergraph gets rid of pairwise assumption. In hypergraph representation, an edge can connect more than two vertices. In this way, an edge actually contains a subset of vertices. Take the same problem in the previous paragraph as an example, images with the objects they share are used to construct the hypergraph shown in Fig. 1. In the hypergraph, the images that contain the same object are clearly grouped. Therefore, we used the hypergraph instead of traditional graph to represent the semantic correlations of the dat.

4. Multi-view hypergraph learning by patch alignment framework

According to the patch alignment framework, the learning process of the proposed dimensionality reduction approach consists of part optimization and global alignment. The flowchart is shown in Fig. 2. In this section, we follow the definitions used in the previous section. The definitions of symbols in the hypergraph are presented in Table 2.

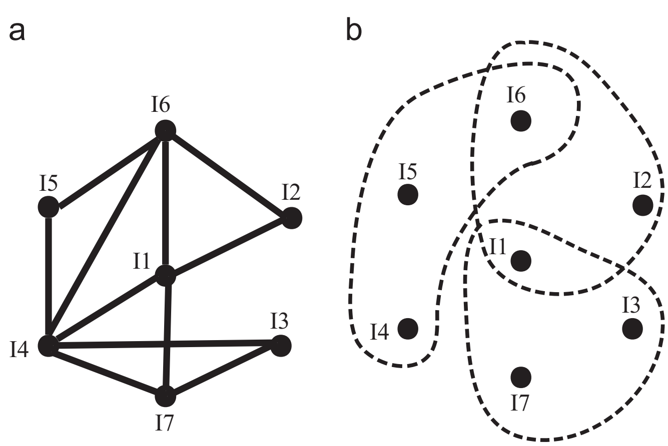

Fig. 1. The comparison of the simple graph and the hypergraph. (a) Traditional graph representing Table 1. (b) Hypergraph representing Table 1.

4.1. Part optimization

As in [26,24], the combinatorial optimization problem in constructing hypergraph could be relaxed into a real-valued form

$$\arg\min_{f\in R^{|V|}} \frac{1}{2}\sum_{e\in E}\sum_{u,v\subset e}\frac{\omega(e)}{\delta(e)}\left(\frac{f(u)}{\sqrt{d(u)}}-\frac{f(v)}{\sqrt{d(v)}}\right)^2 \tag{1}$$

In the part optimization stage, we define one patch to be the vertices connected by one hyperedge. In this way, the patch in the proposed learning process is defined by

$$\arg\min_{f\in R^{|V|}} \sum_{u,v\subset e}\frac{\omega(e)}{\delta(e)}\left(\frac{f(u)}{\sqrt{d(u)}}-\frac{f(v)}{\sqrt{d(v)}}\right)^2 \tag{2}$$

For one patch, we should compute

$$\sum_{u,v\subset e}\frac{\omega(e)}{\delta(e)}\left(\frac{f(u)}{\sqrt{d(u)}}-\frac{f(v)}{\sqrt{d(v)}}\right)^2 \tag{3}$$

It means that we randomly choose two vertices in the subset of vertices contained by a hyperedge e and sum the value of

$$\frac{\omega(e)}{\delta(e)}\left(\frac{f(u)}{\sqrt{d(u)}}-\frac{f(v)}{\sqrt{d(v)}}\right)^2 \tag{4}$$

Therefore, Eq. (3) could be expanded as

$$\begin{aligned}&\frac{\omega(e)}{\delta(e)}\left(\frac{f(1)}{\sqrt{d(1)}}-\frac{f(2)}{\sqrt{d(2)}}\right)^2+\frac{\omega(e)}{\delta(e)}\left(\frac{f(1)}{\sqrt{d(1)}}-\frac{f(3)}{\sqrt{d(3)}}\right)^2\\&\quad+\cdots+\frac{\omega(e)}{\delta(e)}\left(\frac{f(1)}{\sqrt{d(1)}}-\frac{f(n)}{\sqrt{d(n)}}\right)^2\\&\quad+\frac{\omega(e)}{\delta(e)}\left(\frac{f(2)}{\sqrt{d(2)}}-\frac{f(1)}{\sqrt{d(1)}}\right)^2+\frac{\omega(e)}{\delta(e)}\left(\frac{f(2)}{\sqrt{d(2)}}-\frac{f(3)}{\sqrt{d(3)}}\right)^2\\&\quad+\cdots+\frac{\omega(e)}{\delta(e)}\left(\frac{f(2)}{\sqrt{d(2)}}-\frac{f(n)}{\sqrt{d(n)}}\right)^2\\&\quad+\cdots+\frac{\omega(e)}{\delta(e)}\left(\frac{f(n)}{\sqrt{d(n)}}-\frac{f(1)}{\sqrt{d(1)}}\right)^2+\frac{\omega(e)}{\delta(e)}\left(\frac{f(n)}{\sqrt{d(n)}}-\frac{f(2)}{\sqrt{d(2)}}\right)^2\\&\quad+\cdots+\frac{\omega(e)}{\delta(e)}\left(\frac{f(n)}{\sqrt{d(n)}}-\frac{f(n-1)}{\sqrt{d(n-1)}}\right)^2\end{aligned} \tag{5}$$

In this equation, n is the number of vertices contained in e. In another word, n is the sum of the eth column in matrix H. According to patch alignment framework, the matrix form of Eq. (5) is

$$A_1^T H_{u,e}\frac{\Omega}{D_e}H_{u,e}A_1+A_2^T H_{u,e}\frac{\Omega}{D_e}H_{u,e}A_2+\cdots+A_N^T H_{u,e}\frac{\Omega}{D_e}H_{u,e}A_N \tag{6}$$

$H_{u,e}$ indicates the uth item in the eth column. $H_{u,e}==1$ means that vertex u is in hyperedge e. N is the total number of vertices in

Table 2
The definition of symbols in hypergraph.

Symbol	Definition
u, v	Vertices in the hypergraph
e	Edges in the hypergraph
$\omega(e)$	The weight of an edge e
$\delta(e)$	The degree of an edge e. It illustrates how many vertices are connected by e. In traditional graph representation, $\delta(e)=2$
$d(v)$	The degree of a vertex v. It is calculated by summing the weighting values of edges connected to this vertex
D_v	The diagonal matrix containing the vertex degrees
D_e	The diagonal matrix containing the edge degrees
H	In this matrix, $H(v,e)=1$ if $v\in e$
Ω	The diagonal matrix containing the weights of hyperedges
V	The set of vertices
E	The set of edges

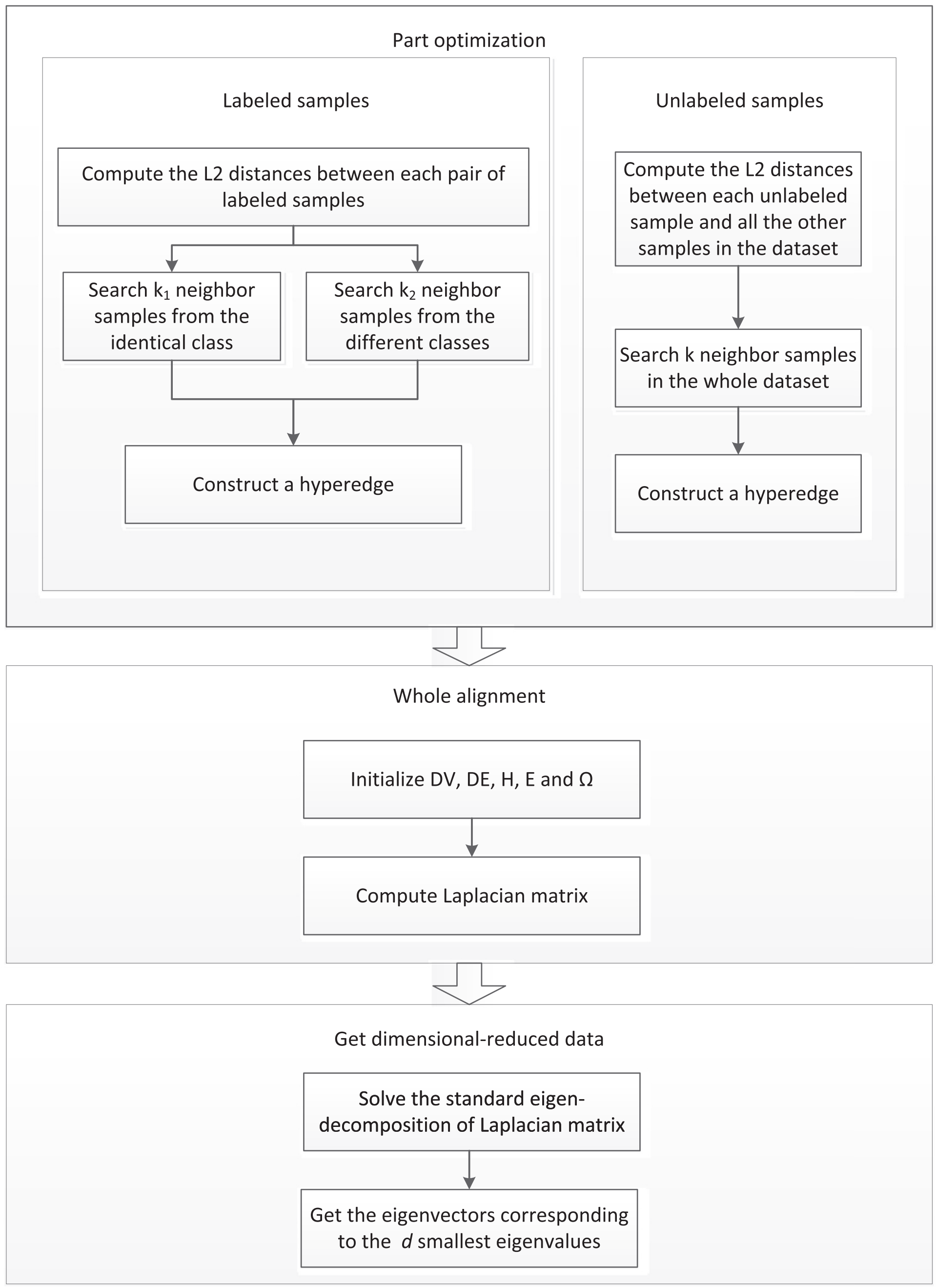

Fig. 2. Flowchart of the proposed method.

the hypergraph. A_n is

$$A_1 = \left[0 \ \ \frac{f(1)}{\sqrt{d(1)}} - \frac{f(2)}{\sqrt{d(2)}} \ \ \frac{f(1)}{\sqrt{d(1)}} - \frac{f(3)}{\sqrt{d(3)}} \ \ \cdots \ \ \frac{f(1)}{\sqrt{d(1)}} - \frac{f(N)}{\sqrt{d(N)}}\right]$$

$$A_2 = \left[0 \ \ 0 \ \ \frac{f(2)}{\sqrt{d(2)}} - \frac{f(3)}{\sqrt{d(3)}} \ \ \cdots \ \ \frac{f(2)}{\sqrt{d(2)}} - \frac{f(N)}{\sqrt{d(N)}}\right]$$

$$\cdots$$

$$A_n = \left[0 \ \ 0 \ \ 0 \ \ \cdots \ \ \frac{f(n)}{\sqrt{d(n)}} - \frac{f(N)}{\sqrt{d(N)}}\right] \tag{7}$$

Eq. (5) could be rewritten as

$$\frac{1}{2}\left(B_1^T H_{u,e}\frac{\Omega}{D_e}H_{u,e}B_1 + B_2^T H_{u,e}\frac{\Omega}{D_e}H_{u,e}B_2 + \cdots + B_N^T H_{u,e}\frac{\Omega}{D_e}H_{u,e}B_N\right) \tag{8}$$

In this equation, B_n is

$$B_1 = \left[0 \ \ \frac{f(1)}{\sqrt{d(1)}} - \frac{f(2)}{\sqrt{d(2)}} \ \ \frac{f(1)}{\sqrt{d(1)}} - \frac{f(3)}{\sqrt{d(3)}} \ \ \cdots \ \ \frac{f(1)}{\sqrt{d(1)}} - \frac{f(N)}{\sqrt{d(N)}}\right]$$

$$B_2 = \left[\frac{f(2)}{\sqrt{d(2)}} - \frac{f(1)}{\sqrt{d(1)}} \ \ 0 \ \ \frac{f(2)}{\sqrt{d(2)}} - \frac{f(3)}{\sqrt{d(3)}} \ \ \cdots \ \ \frac{f(2)}{\sqrt{d(2)}} - \frac{f(N)}{\sqrt{d(N)}}\right]$$

$$\cdots$$

$$B_n = \left[\frac{f(n)}{\sqrt{d(n)}} - \frac{f(1)}{\sqrt{d(1)}} \ \ \frac{f(n)}{\sqrt{d(n)}} - \frac{f(2)}{\sqrt{d(2)}} \ \ \cdots \ \ \frac{f(n)}{\sqrt{d(n)}} - \frac{f(N)}{\sqrt{d(N)}}\right] \tag{9}$$

Finally, the patch optimization for each hyperedge is defined by

$$\frac{1}{2}\sum_{v \subset e}\frac{F}{DV_v^{1/2}} EH'_e \frac{\Omega}{DE} H_e E' \frac{F}{DV_v^{1/2}} \tag{10}$$

Matrix E is

$$\begin{bmatrix} -\vec{e}^{\,T} \\ I \end{bmatrix} \tag{11}$$

where $\vec{e} = [1, \ldots, 1]^T$, I is an $n \times n$ identity matrix.

4.2. Whole alignment

In the whole alignment stage of the proposed dimensionality reduction method for multi-view data, we compute the hyperedge for each view. All the hyperedges are used to construct a whole hypergraph. In this way, the optimizations described in the previous subsection will be unified together as a whole one. The optimization for each part are weighted by the weight of the corresponding hyperedge, which is put in the weighting matrix Ω. The computation of the weights is inspired by Huang et al.'s work [27]. In the hypergraph, the weight of a hyperedge is computed by summing the similarity scores of all the pairs of vertices contained in this hyperedge. The similarity score of any pair of vertices is defined by

$$S(u,v) = \exp\left(-\frac{1}{\beta} dist(feat(u), feat(v))\right) \tag{12}$$

In this equation, $feat(u)$ is some type of feature representing vertex u, $dist(feat(u),feat(v))$ is the distance between two features and β is the standard deviation of all distances.

With the hyperedge weighting matrix, the multi-view hypergraph Laplacian can be computed by summing the patch optimization defined in Eq. (10) of all the hyperedges

$$\frac{1}{2}\sum_{e \in E}\sum_{v \in e} \frac{F}{DV_v^{1/2}} EH_e' \frac{\Omega}{DE} H_e E' \frac{F}{DV_v^{1/2}} \tag{13}$$

4.3. Algorithm details

The proposed dimensionality reduction method is semi-supervised for multi-view data. The procedure of MHL is listed as following:

1. For each labeled sample in any view, search k_1 neighbor samples from an identical class and k_2 neighbor samples from different classes. They are used to construct a hyperedge and a patch is built according to Eq. (10).
2. For each unlabeled sample in any view, search k neighbor samples in the whole dataset. They are used to construct a hyperedge and a patch is built according to Eq. (10).
3. The processes in Step 1 and 2 are repeated for all the views.
4. All the hyperedges from different views are used to construct a whole hypergraph. The multiview hypergraph Laplacian matrix is computed according to Eq. (13).
5. Solve the standard eigen-decomposition to obtain the projection matrix, whose vectors are the eigenvectors corresponding to the d smallest eigenvalues. Then we get the dimensionality-reduced data with d dimensions.

Table 3
The space consumption of each matrix.

Matrix	Size
D_v	$\lvert V\rvert \times \lvert V\rvert$
D_e	$\lvert E\rvert \times \lvert E\rvert$
H	$\lvert E\rvert \times \lvert V\rvert$
Ω	$\lvert E\rvert \times \lvert E\rvert$
L	$\lvert E\rvert \times \lvert E\rvert$

4.4. Analysis

We look into MHL in two aspects: space complexity and time complexity. The space consumption of MHL mainly consists of the matrices representing the hypergraph. They are D_v, D_e, H and Ω. They are variable according to the number of images or views. The number of images is equal to the number of vertices $|V|$. According to the constructing procedure of MHL, the number of edges in each view is equal to the number of vertices. Then, the total number of edges is $|E| = |V| \times \mu$, where μ is the number of views. Including the final Laplacian matrix, their sizes are listed in Table 3.

The time complexity of MHL contains three parts. The first part is for the constructions of the matrices by combining different views, i.e., the computation of Ω. The computation of Ω is influenced by the number of nearest neighbors. There are 3 parameters in hypergraph construction to find nearest neighbors. They are listed in Table 4. The time complexity of constructing Ω is

$$O(k_i \times N_{labeled} + k_d \times N_{labeled} + k_u \times N_{unlabeled}) \tag{14}$$

In this equation, $N_{labeled}$ is the number of labeled images while $N_{unlabeled}$ is the number of unlabeled images. The second part is for the computation of Laplacian matrix. This part of time complexity is

$$O\left(\sum_{v \in V_{labeled}} ((k_i + k_d) \times N_{labeled}^2) + \sum_{v \in V_{unlabeled}} (k_u \times |V|^2)\right) \tag{15}$$

The third part is for eigenvalue decomposition. It is $O(|V|^3)$.

5. Experimental results

5.1. Datasets

We test the proposed method on PASCAL VOC2007 database [46]. PASCAL VOC2007 is challenging for its class variations and cluttered backgrounds. To decrease the difficulty, we choose around 1200 images from six easier classes (person, airplane, train, boat, motorbike, and horse) of the VOC2007 for our experimental comparison. Since in the VOC2007 one image may contain multiple objects, we classify each image to one specific class according to the largest object it contains. The dimensionality-reduced data are divided into two sets, one for training and the other for testing. Support Vector Machine is used for classification. The dimensionality of the low-dimensional embedding d is from 10 to 360. We compute the correct ratios of classification for every 10 values in this rage. With the

Table 4
The parameters used in the proposed method.

Name	Meaning
k_i	The number of nearest neighbors for labeled samples from the identical class
k_d	The number of nearest neighbors for labeled samples from the different class
k_u	The number of nearest neighbors for unlabeled samples

accuracy, we compare the proposed method with MSE [45] in our experiments.

5.2. Parameter optimization

In this subsection, we aim at optimizing the combination of parameters, which have been listed in Table 4. For the consistency of the weighting process, we should make the number of nearest neighbors for labeled samples equal to the number of nearest neighbors for unlabeled samples. In this way, we set the following prerequisite:

$$k_u = k_i + k_d \tag{16}$$

To clearly compare the accuracy, we show the average correct ratio of all dimensionality for each k_i and k_d. First, we fix the value of $k_d=5$. The average correct ratio is shown in Fig. 3. We could see that the best performance is achieved when $k_i=6$.

Second, we fix the value of $k_i=6$. The average correct ratio is shown in Fig. 4. We easily point out that the accuracy is the highest when $k_d=3$.

With the above experiment, we could set the best combination of MHL as Table 5.

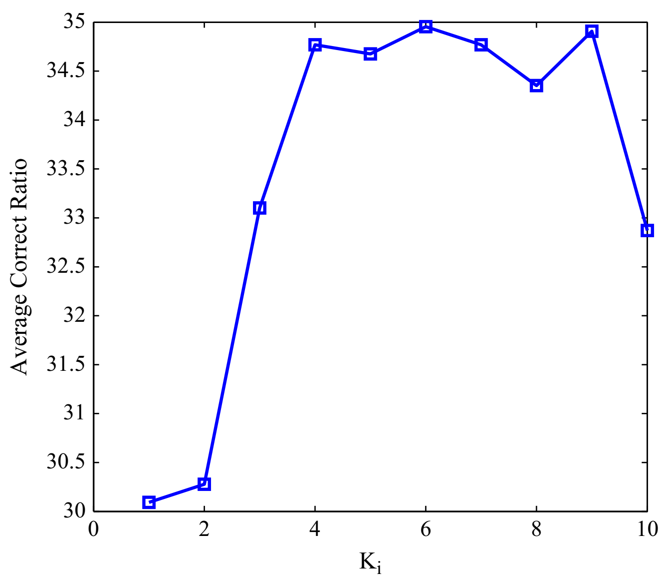

Fig. 3. The changes of average correct ratio with different k_i.

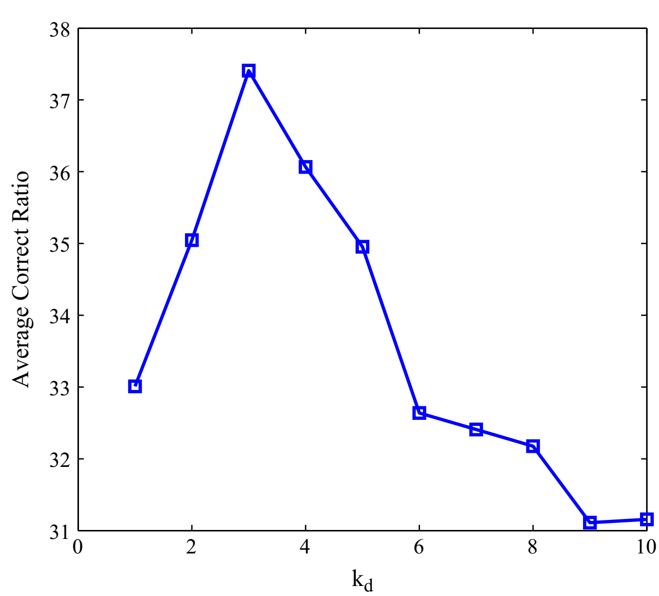

Fig. 4. The changes of average correct ratio with different k_d.

Table 5
The best combination of parameters used in the proposed method.

Name	Value
k_i	6
k_d	3
k_u	9

5.3. Comparison with previous methods

In this subsection, we compare MHL with MSE and PCA on classification accuracy. The settings of parameter for MSE is the same as the setting of MHL in Table 5. To make PCA work on multi-view data, we use two tricks. One is linking features from different views to form a long vector, the other one is computing the classification accuracy with different views and averaging the performance. The result is shown in Fig. 5. Although the correct ratios of previous methods are sometimes higher than the proposed MHL, MHL is the best of all in most of the cases. Especially when the reduced dimension is lower than 100, MHL is much better than all the existing methods. It indicates that MHL works well in low dimension. However, the performance of MHL significantly decreases when the dimension is set as 100. The average correct ratio of MHL is 37.4%, while the average correct ratio of MSE is 28.9%. The average correct ratio is %31.3 for PCA with long vectors and %28.5 for PCA with average performance. It indicates that the performance of MHL is higher than previous methods.

6. Conclusions

In this paper, we propose a novel dimensionality reduction method for multi-view data, which is called Multi-view Hypergraph Learning (MHL). It is based on patch alignment framework. First, our method use hypergraph to represent the relationships among images

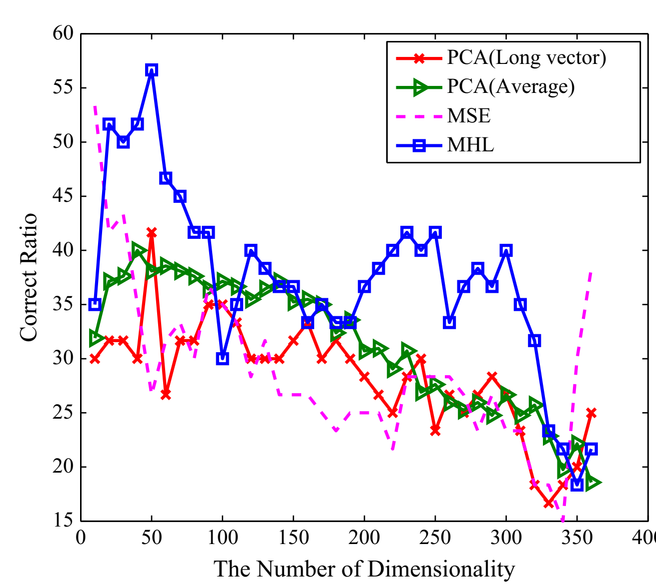

Fig. 5. The comparison of correct ratios.

in the data set. The hyperedges in the hypergraph are considered as patches in part optimization of patch alignment framework. Second, with the hypergraph and distances between neighboring pairs, we compute the weights of hyperedges. The hyperedges from different views could be integrated to construct Laplacian Matrix for standard eigen-decomposition. Experimental results on VOC2007 image data set show that the proposed method outperforms previous method on both retrieval performance.

Acknowledgments

This work is supported by the Natural Science Foundation of China (61202145, 61100104 and 61065007), the Program for New Century Excellent Talents in University (No. NCET-12-0323) and the Natural Science Foundation of Fujian Province of China (2012J01287).

References

[1] Y. Yang, J. Song, Z. Huang, Z. Ma, N. Sebe, A.G. Hauptmann, Multi-feature fusion via hierarchical regression for multimedia analysis, in press.
[2] M. Wang, H. Li, D. Tao, K. Lu, X. Wu, Multimodal graph-based reranking for web image search, IEEE Trans. Image Process. 21 (2012) 4649–4661.
[3] J. Yu, D. Tao, Y. Rui, J. Cheng, Pairwise constraints based multiview features fusion for scene classification, Pattern Recognition 46 (2013) 483–496.
[4] J. Yu, M. Wang, D. Tao, Semisupervised multiview distance metric learning for cartoon synthesis, IEEE Trans. Image Process. 21 (2012) 4636–4648.
[5] H. Hotelling, Analysis of a complex of statistical variables into principal components, J. Educ. Psychol. 24 (1933) 417–441.
[6] R. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugen. 7 (1936) 179–188.
[7] S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (2000) 2323–2326.
[8] V.S.J. Tenenbaum, J. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (2000) 2319–2323.
[9] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, Adv. Neural Inf. Process. Syst. 14 (2002) 585–591.
[10] D. Donoho, C. Grimes, Hessian eigenmaps: new locally linear embedding techniques for high-dimensional data, Proc. Natl. Acad. Sci. 100 (2003) 5591–5596.
[11] Z. Zhang, H. Zha, Principal manifolds and nonlinear dimension reduction via local tangent space alignment, SIAM J. Sci. Comput. 26 (2005) 313–338.
[12] X. Zhu, Z. Ghahramani, J. Lafferty, Semi-supervised learning using Gaussian fields and harmonic functions, in: International Conference on Machine Learning, vol. 2, AAAI Press, 2003, pp. 912–919.
[13] D. Zhou, O. Bousquet, T.N. Lal, J. Weston, B. Scholkopf, Learning with local and global consistency, in: Advances in Neural Information Processing Systems, vol. 16, MIT Press, 2003, pp. 321–328.
[14] Y. Zhuang, Y. Yang, F. Wu, Mining semantic correlation of heterogeneous multimedia data for cross-media retrieval, IEEE Trans. Multimedia 10 (2008) 221–229.
[15] M. Wu, B. Scholkopf, Transductive classification via local learning regularization, in: International Conference on Artificial Intelligence and Statistics, Microtome, 2007, pp. 628–635.
[16] M. Szummer, T. Jaakkola, Partially labeled classification with Markov random walks, in: Advances in Neural Information Processing Systems, vol. 14, MIT Press, 2002, pp. 945–952.
[17] Z. Ma, Y. Yang, F. Nie, J. Uijlings, N. Sebe, Exploiting the entire feature space with sparsity for automatic image annotation, in: The 19th ACM International Conference on Multimedia, ACM, 2011, pp. 283–292.
[18] Y. Yang, F. Nie, D. Xu, J. Luo, Y. Zhuang, Y. Pan, A multimedia retrieval framework based on semi-supervised ranking and relevance feedback, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2012) 723–742.
[19] Y. Yang, Y.-T. Zhuang, F. Wu, Y. he Pan, Harmonizing hierarchical manifolds for multimedia document semantics understanding and cross-media retrieval, IEEE Trans. Multimedia 10 (2008) 437–446.
[20] G. Li, M. Wang, Z. Lu, R. Hong, T.-S. Chua, In-video product annotation with web information mining, ACM Trans. Multimedia Comput. Commun. Appl. 8 (2012) 55:1–55:19.
[21] J. Yu, D. Liu, D. Tao, H.S. Seah, On combining multiple features for cartoon character retrieval and clip synthesis, IEEE Trans. Syst. Man Cybern. Part B 42 (2012) 1413–1427.
[22] S. Yan, H. Wang, Semi-supervised learning by sparse representation, in: The SIAM International Conference on Data Mining, SIAM, 2009, pp. 792–801.
[23] J. Yu, D. Tao, M. Wang, Adaptive hypergraph learning and its application in image classification, IEEE Trans. Image Process. 21 (2012) 3262–3272.
[24] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 888–905.
[25] Y. Huang, Q. Liu, D. Metaxas, Video object segmentation by hypergraph cut, in: IEEE International Conference on Computer Vision and Pattern Recognition, IEEE Press, 2009, pp. 1738–1745.
[26] D. Zhou, J. Huang, B. Scholkopf, Learning with hypergraphs: clustering, classification, and embedding, in: Advances in Neural Information Processing Systems, vol. 19, MIT Press, 2006, pp. 1601–1608.
[27] Y. Huang, Q. Liu, F. Lv, Y. Gong, D.N. Metaxas, Unsupervised image categorization by hypergraph partition, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 1266–1273.
[28] D. Tao, X. Li, X. Wu, S.J. Maybank, General tensor discriminant analysis and Gabor features for gait recognition, IEEE Trans. Pattern Anal. Mach. Intell. 29 (2007) 1700–1715.
[29] M. Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, Y. Song, Unified video annotation via multigraph learning, IEEE Trans. Circuits Syst. Video Technol. 19 (2009) 733–746.
[30] D. Tao, X. Li, X. Wu, W. Hu, S.J. Maybank, Supervised tensor learning, Knowl. Inf. Syst. 13 (2007) 1–42.
[31] D. Tao, X. Li, X. Wu, S.J. Maybank, Geometric mean for subspace selection, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2009) 260–274.
[32] D. Tao, X. Tang, X. Li, X. Wu, Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval, IEEE Trans. Pattern Anal. Mach. Intell. 28 (2006) 1088–1099.
[33] C. Berge, Hypergraphs, North-Holland, Amsterdam, 1989.
[34] Y. Wen, Y. Gao, S. Liu, Q. Cheng, R. Ji, Hyperspectral image classification with hypergraph modelling, in: The 4th International Conference on Internet Multimedia Computing and Service, 2012, pp. 34–37.
[35] Y. Gao, T.-S. Chua, Hyperspectral image classification by using the pixel spatial correlation, in: International Conference on Multimedia Modeling, 2013.
[36] J. Yu, D. Liu, D. Tao, H.S. Seah, Complex object correspondence construction in two-dimensional animation, IEEE Trans. Image Process. 20 (2011) 3257-3269.
[37] Y. Gao, M. Wang, Z.-J. Zha, J. Shen, X. Li, X. Wu, Visual-textual joint relevance learning for tag-based social image search, IEEE Trans. Image Process. 22 (2013) 363–376.
[38] Y. Gao, M. Wang, H. Luan, J. Shen, S. Yan, D. Tao, Tag-based social image search with visual-text joint hypergraph learning, in: The 19th ACM International Conference on Multimedia, ACM, 2011, pp. 1517–1520.
[39] M. Wang, B. Ni, X.-S. Hua, T.-S. Chua, Assistive tagging: a survey of multimedia tagging with human–computer joint exploration, ACM Comput. Surv. 4 (2012).
[40] M. Wang, R. Hong, G. Li, Z.-J. Zha, S. Yan, T.-S. Chua, Event driven web video summarization by tag localization and key-shot identification, IEEE Trans. Multimedia 14 (2012) 975–985.
[41] C. Hong, J. Zhu, Hypergraph-based multi-example ranking with sparse representation for transductive learning image retrieval, Neurocomputing 101 (2012) 94–103.
[42] Y. Gao, M. Wang, D. Tao, R. Ji, Q. Dai, 3-d object retrieval and recognition with hypergraph analysis, IEEE Trans. Image Process. 21 (2012) 4290–4303.
[43] T. Zhang, D. Tao, X. Li, J. Yang, Patch alignment for dimensionality reduction, IEEE Trans. Knowl. Data Eng. 21 (2009) 1299–1313.
[44] T. Zhang, D. Tao, J. Yang, Discriminative locality alignment, in: The 10th European Conference on Computer Vision: Part I, Springer-Verlag, 2008, pp. 725–738.
[45] T. Xia, D. Tao, T. Mei, Y. Zhang, Multiview spectral embedding, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 40 (2010) 1438–1446.
[46] M. Everingham, L. Van Gool, C.K. I. Williams, J. Winn, A. Zisserman, The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results, 〈http://www.pascal-network.org/challenges/VOC/voc2007/workshop/index.html〉, 2007.

Chaoqun Hong is currently a lecturer in the department of Computer Science, Xiamen University of Technology, PR China. He received the Ph.D. degree in 2011 from Zhejiang University,PR China. His research interests include video codec, image processing, computer vision and patter recognition.

Jun Yu is an associate professor in Xiamen University, China. He received the B.E. degree and Ph.D. degree from the Department of Computer Science, Zhejiang University, Hangzhou, China, respectively. His current research interests include computer graphics, computer vision, machine learning and multimedia. He has authored and coauthored more than 30 journal and conference papers in these areas.

Jonathan Li holds the Ph.D. degree in geomatics engineering from the University of Cape Town, South Africa. He has been with the Key Laboratory for Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry of Education, Xiamen, China, since 2011 and also holds Professorship with the Department of Geography and Environmental Management, University of Waterloo, Canada. He has co-authored more than 200 publications, over 70 of which were published in refereed journals. He has co-edited six books and six theme issues. His current research interests are in the areas of remote sensing and spatial informatics ranging from remote sensing of inland and coastal waters, mobile laser scanning for 3D critical infrastructure modeling, and geomatics solutions to disaster management. Dr. Li was the recipient of the 2011 Talbert Abrams Award for Best Paper in Photogrammetry and Remote Sensing, the 2008 ESRI Award for Best Paper in GIScience, and the 2006 MDA Award for Best Paper in Photogrammetry and Hydrography. He is Chair of ISPRS ICWG I/Va on Mobile Scanning and Imaging Systems (2012–C2016), Vice Chair of ICA Commission on Mapping from Remote Sensor Imagery (2011–C2015), and Vice Chair of FIG Commission on Hydrography (2011–C2014). He is IEEE Senior Member and PEng licenced with PEO. junli@xmu.edu.cnJonathan Li holds the Ph.D. degree in geomatics engineering from the University of Cape Town, South Africa. He has been with the Key Laboratory for Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry of Education, Xiamen, China, since 2011 and also holds Professorship with the Department of Geography and Environmental Management, University of Waterloo, Canada. He has co-authored more than 200 publications, over 70 of which were published in refereed journals. He has co-edited six books and six theme issues. His current research interests are in the areas of remote sensing and spatial informatics ranging from remote sensing of inland and coastal waters, mobile laser scanning for 3D critical infrastructure modeling, and geomatics solutions to disaster management. Dr. Li was the recipient of the 2011 Talbert Abrams Award for Best Paper in Photogrammetry and Remote Sensing, the 2008 ESRI Award for Best Paper in GIScience, and the 2006 MDA Award for Best Paper in Photogrammetry and Hydrography. He is Chair of ISPRS ICWG I/Va on Mobile Scanning and Imaging Systems (2012–C2016), Vice Chair of ICA Commission on Mapping from Remote Sensor Imagery (2011–C2015), and Vice Chair of FIG Commission on Hydrography (2011–C2014). He is IEEE Senior Member and PEng licenced with PEO. junli@xmu.edu.cn

Xuhui Chen is currently the deputy secretary of the department of Computer Science, Xiamen University of Technology, PR China. He received the Ph.D. degree in 2004 from Xi'an Jiaotong University. He was a postdoc in Arizona State University of U.S.A. His research interests include wireless sensor network and business intelligence.