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