Neurocomputing 156 (2015) 12–21

Contents lists available at ScienceDirect

Neurocomputing

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

Multi-view clustering via pairwise sparse subspace representation

Qiyue Yin, Shu Wu, Ran He, Liang Wang *

Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 26 September 2014
Received in revised form
4 January 2015
Accepted 4 January 2015
Communicated by Yi Zhe Song
Available online 19 January 2015

Keywords:
Multi-view clustering
Subspace clustering
Pairwise sparse representation

ABSTRACT

Multi-view clustering, which aims to cluster datasets with multiple sources of information, has a wide range of applications in the communities of data mining and pattern recognition. Generally, it makes use of the complementary information embedded in multiple views to improve clustering performance. Recent methods usually find a low-dimensional embedding of multi-view data, but often ignore some useful prior information that can be utilized to better discover the latent group structure of multi-view data. To alleviate this problem, a novel pairwise sparse subspace representation model for multi-view clustering is proposed in this paper. The objective function of our model mainly includes two parts. The first part aims to harness prior information to achieve a sparse representation of each high-dimensional data point with respect to other data points in the same view. The second part aims to maximize the correlation between the representations of different views. An alternating minimization method is provided as an efficient solution for the proposed multi-view clustering algorithm. A detailed theoretical analysis is also conducted to guarantee the convergence of the proposed method. Moreover, we show that the must-link and cannot-link constraints can be naturally integrated into the proposed model to obtain a link constrained multi-view clustering model. Extensive experiments on five real world datasets demonstrate that the proposed model performs better than several state-of-the-art multi-view clustering methods.

1. Introduction

Many kinds of real-world data appear in multiple views. For example, web pages contain both images and corresponding texts, and images can be encoded by different features such as color histogram and Fourier shape descriptors. Although learning tasks such as classification and clustering can be approached based on one single view, multiple views providing complementary information can improve the performance of learning tasks [1]. This leads to a surge of interest in multi-view learning, whose goal is to exploit multiple views to obtain better performance rather than relying on every single view. Till now, multi-view learning has been widely studied in different areas such as data mining, multimedia, computer vision and natural language processing [2–5].

As one of the basic tasks of multi-view learning, multi-view clustering has attracted more and more attention because it can handle large numbers of unlabeled datasets. The objective of multi-view clustering is to cluster multi-view datasets based on their latent groups. Generally, the main challenge lies in how to make use of the complementary characteristics embedded in the multiple sources of information. Plenty of multi-view clustering algorithms have been developed to solve this problem. Some methods aim to find a unified low-dimensional embedding to fuse the multi-view representations, and clustering is then performed when the unified representation is obtained [6,7]. These methods often map the original high-dimensional feature space to a latent low-dimensional space so as to well explore the feature correlation between different views. On the other hand, some methods perform multi-view clustering through merging the clustering results from different individual views [8,9]. These methods, called late fusion, obtain the final clustering results by voting or other fusion strategies. For more details about multi-view clustering, refer to Section 2.

Although various existing methods indeed improve the clustering performance for multi-view data, they often do not take some useful prior knowledge into consideration, such as collaborative [10], sparse [11] and low-rank [12] information, which has been shown to be helpful for clustering in some data mining applications. On the other hand, spectral-based subspace clustering methods [13] are recently developed, which can take advantage of such prior information and achieve promising results. These methods bring in different prior

* Corresponding author.
E-mail addresses: qyyin@nlpr.ia.ac.cn (Q. Yin), shu.wu@nlpr.ia.ac.cn (S. Wu), rhe@nlpr.ia.ac.cn (R. He), wangliang@nlpr.ia.ac.cn (L. Wang).

http://dx.doi.org/10.1016/j.neucom.2015.01.017
0925-2312/

knowledge to constrain the self-representation matrix of the dataset to be the ideal block diagonal matrix, and then use the spectral method to obtain the final clustering results. Moreover, algorithms in this class can well discover the relationship between data points and reflect the latent group structure of the dataset. However, these methods often focus on the single view data and could not been directly applied for multi-view datasets.

Inspired by the recent advances in subspace clustering, this paper proposes a novel multi-view clustering framework based on sparse subspace representation. The proposed model resorts to subspace clustering for efficiently using the prior knowledge compared with conventional multi-view clustering methods. Besides, pairwise co-regularization is developed to explore the complementary information embedded in the multi-view data. More specifically, the sparse representation of the dataset for each view is firstly constructed. At the same time, a pairwise co-regularization constraint is utilized to capture the interaction between the correlated view-specific sparse representations. Then, we develop an iterative algorithm to efficiently solve the proposed framework, and provide rigid theoretical analysis on the convergence of this algorithm. Moreover, we discuss the impacts of the proposed different co-regularization forms in exploring the correlation between views. In addition, we show that the link prior can be easily integrated into our proposed model and a link constrained multi-view clustering method is accordingly developed. Extensive experiments are conducted to demonstrate the effectiveness of the proposed methods.

Main contributions in this paper are summarized as follows:

(1) A novel pairwise co-regularization model is proposed for the multi-view clustering problem. It harnesses the prior information to obtain the view specific sparse representation and meanwhile utilizes the correlation between different views. Besides, different co-regularization forms are discussed as special examples in our framework.
(2) A novel link constrained multi-view clustering algorithm is developed to naturally integrate the partially observed supervisory information (e.g., must-link and cannot-link). To the best of our knowledge, this is rarely studied in the literature of multi-view clustering.
(3) We verify the effectiveness of the proposed multi-view clustering algorithms with extensive experiments on five real world datasets, achieving state-of-the-art results in terms of accuracy and normalized mutual information.

The rest of this paper is organized as follows. In Section 2, we briefly review multi-view clustering and subspace clustering algorithms. Then our multi-view sparse subspace clustering method is introduced in Section 3. Section 4 gives the extensions of our multi-view clustering model. Extensive experimental results and analysis are given in Section 5. Finally, Section 6 concludes the paper.

2. Related work

In this section, we briefly introduce the background of our proposed model, which consists of multi-view clustering and subspace clustering.

2.1. Multi-view clustering

Multi-view clustering, which aims to cluster the dataset with multiple views, can be roughly classified into three categories based on the usage of multiple sources of information in the clustering process [1,14]. Algorithms in the first category find a unified low-dimensional embedding of multi-view data, and then cluster the dataset using this representation like the single view clustering methods [15,6,2,16–19]. These methods, also called subspace learning-based methods, are widely studied. Kumar et al. [6] proposed a co-regularization framework to regularize the difference between view-specific Laplacian embeddings. Liu et al. [7] developed a multi-view non-negative matrix factorization framework to gain a consensus low-dimensional feature matrix from the original high-dimensional data, and He et al. [20] further improved the idea of multi-view non-negative matrix factorization based clustering algorithms. Recently, Wang et al. [3] proposed a regression-like clustering method, which directly obtains the final consensus label matrix.

The second category directly integrates the information of different views in the clustering process. Popular examples are the co-EM clustering algorithm [21] and the co-training framework [22–24]. Kumar and Daume [23] resorted to the co-training framework, which is widely used in the semi-supervised learning, to design the first co-training based multi-view spectral clustering algorithm. Zhao et al. [22] combined LDA, K-means with the co-training framework and developed a subspace co-training framework for the multi-view clustering task. In contrast, the third category is late fusion (or called ensemble clustering). That is, the final clustering result is derived from integrating each individual clustering result [25,9,26]. Long et al. [8] proposed to use mapping functions to make clusters from different views comparable and learn the best clusters from these multiple views. Greene and Cunningham [9] developed a matrix factorization based method to group the clusters obtained from each view.

Overall, these multi-view clustering methods indeed improve clustering performance for multi-view datasets. However, they rarely consider some useful prior knowledge, such as sparse or low rank information of the latent group structure, which has been shown to be helpful for clustering in some data mining applications.

2.2. Subspace clustering

Subspace clustering aims to cluster the high-dimensional data into multiple subspaces as well as find the subspaces fitting each group of data points. Generally it can be divided into four categories based on different techniques [13], and our discussion mainly focuses on the recently developed spectral-based subspace clustering methods [27,28,10,29]. The key idea of these approaches is to obtain a self-representation matrix by taking different prior information into consideration. Usually, these prior information is utilized as different constraints to achieve different self-representation matrices. Then the above matrix is applied to construct the affinity matrix, which is used for the final spectral clustering. Examples lie in this category are briefly introduced as follows.

Sparse Subspace Clustering (SSC) [11,28] is based on the fact that each point in a union of subspaces can be written as a linear (affine) combination of points belonging to the same subspace. Thus, the representation coefficients for a data point should be sparse, and this prior information is brought into the model by using an l_1-norm to constrain the representation coefficients. Different from SSC, Low Rank Representation (LRR) based subspace segmentation algorithms [12,27] seek the lowest rank representation for all points. The prior information lies in the low rank characteristic of the optimal representation matrix. Whereas, the Multi-Subspace Representation (MSR) based subspace segmentation methods [30,31] regularize the representation matrix to be both low rank and sparse. By a careful parameter configuration, the subspace structure can be well revealed. Least Squares Regression (LSR) [10,32] applies ridge regression to