Neurocomputing 208 (2016) 174–182

Contents lists available at ScienceDirect

Neurocomputing

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

A Laplacian structured representation model in subspace clustering for enhanced motion segmentation

Wanjun Chen [a], Erhu Zhang [a,*], Zhuomin Zhang [b]

[a] Department of Information Science, Xian University of Technology, Xian 710048, China
[b] Department of Computer Science and Technology, Nanjing University, Nanjing 210046, China

ARTICLE INFO

Article history:
Received 25 September 2015
Received in revised form
5 December 2015
Accepted 5 December 2015
Available online 2 June 2016

Keywords:
Motion segmentation
Subspace clustering
Local feature similarity
Sparse subspace clustering (SSC)
Low-rank representation (LRR)

ABSTRACT

Segmenting a moving object from its background is a fundamental step in many computer vision applications ranging from visual surveillance to multimedia image analysis. Although subspace clustering based motion segmentation methods (such as Sparse Subspace Clustering (SSC) and Low-Rank Representation (LRR)) have achieved state-of-the-art performance, they have suffered from the loss of local structure problem, i.e., local similar features may be encoded as totally different codes due to the overcomplete codebooks. Such instability may harm the connectivity of the similarity graph and affect the performance of clustering algorithms finally. To remedy this issue, we propose a Laplacian structured representation model to enhance the representation-based clustering methods by importing local feature similarity prior information to guide the encoding process, and then develop an efficient Alternating Direction Method of Multipliers (ADMM) algorithm for optimization. Two improved subspace clustering methods, the Enhanced Sparse Subspace Clustering (E-SSC) and Enhanced Low Rank Representation (E-LRR), are devised in this work. Experiments on Hopkins 155 motion segmentation dataset and airport dataset demonstrate the advantage of our proposed model over state-of-the-art methods, and achieve 0.77% and 0.85% segmentation error rate, respectively.

1. Introduction

Motion segmentation aims at labeling a set of tracked feature points from several moving objects into different groups based on their motions, as illustrated in Fig. 1. It is an essential building block for robotics, face recognition [1–6], video classification [7], action recognition [8–12], event detection [13] and many other applications [14,15]. Under the affine camera model, motion segmentation from tracked feature points can be formulated as a subspace clustering problem [16], where each subspace corresponds to a different motion. In subspace clustering, given the data from a union of subspaces, the objective is to find the number of subspaces, their dimensions, the segmentation of the data points and a basis for each subspace [17]. Subspace clustering has numerous applications, such as image segmentation [18], face clustering [19], and motion segmentation [20]. In this paper, we will focus on subspace clustering for motion segmentation.

Most existing subspace clustering methods perform subspace clustering by two steps: firstly learning an affinity matrix (i.e., an undirected similarity graph) from the given data, and then obtaining the segmentation results by using spectral segmentation algorithms such as the Normalized Cuts (NCut) [21]. The major difference among various methods is the approach for learning the affinity matrix that encodes the subspace memberships among data points.

Generally, there are two kinds of metrics to build an affinity matrix, i.e., pairwise distance and representation coefficients, also called reconstruction coefficients. The former measures the similarity by computing the distance between two data points, e.g., Euclidean distance. Pairwise distance could capture the local structure of data set, whose value only depends on the distance between two data points. As a result, it is sensitive to noise and outliers. Alternatively, representation coefficients-based methods assume that each data point can be denoted as a linear combination of other data points, and thus the representation coefficient can be regarded as a kind of measurement. The metric is robust against noise and outliers since the value of coefficient not only depends on the two connected data points, but also depends on the other data points. In other words, representation coefficients are data-adaptive. Several recent works have shown that representation coefficient is superior to pairwise distance based similarity in subspace clustering. For example, ℓ_1-norm based Sparse Subspace Clustering (SSC) [20,22].

* Corresponding author. Tel.: +86 02982312435.
E-mail addresses: wjchen@xaut.edu.cn (W. Chen), eh-zhang@xaut.edu.cn (E. Zhang).

http://dx.doi.org/10.1016/j.neucom.2015.12.123
0925-2312/

Fig. 1. Motion segmentation: segmentation the feature points according to the multiple moving objects. Some sample frames taken from the car3 sequence in the Hopkins155 database.

All representation-based subspace clustering methods have a common characteristic, i.e. they encode a feature using an overcomplete codebook, which means the base number is much larger than the feature dimensions. Although these methods have achieved state-of-the-art performance in data clustering, they have suffered from the loss of local structure problem, i.e. local similar features may be encoded as totally different codes (large difference in representation coefficients, or different codebook items, or both) under such overcomplete codebooks. That is to say, these methods suffer from the instability of the coding, which may destroy the connectivity of the similarity graph. As an illustration, Fig. 2(e) shows the sparse coding representation coefficients of two local similar features in Fig. 2(a) or (c), but the sparse codes of these two features have a big difference (The location of dash line in Fig. 2 (e) indicates the difference). Such instability may harm the connectivity of the similarity graph and misclassify features into incorrect subspaces finally. To address this issue, we propose a Laplacian structured representation model, which is a uniform formulation to the representation-based subspace clustering methods, and apply the model to two most popular approaches, SSC [20,22] and LRR [19,23]. By adding an additional locality preserving term to the formulation of representation-based approaches, our model can learn more discriminative encoding coefficients and preserve the local feature similarity in the process of encoding, and therefore the robustness of the encoding is enhanced, seeing Fig. 2(b), (d) and (f). Comparing Fig. 2(e) with 2(f), we can clearly find that: by importing local feature similarity prior information into representation-based subspace clustering methods to guide the encoding process, the difference of the codebook items and coefficients become less, and so the connectivity of the similarity graph can be further enhanced.

The main contributions of our work can be summarized as follows: 1) The proposed method is a uniform formulation which makes representation-based subspace clustering algorithms be capable of preserving the local feature similarity while encoding the feature points. This is beneficial to the connectivity of similarity graph constructed by the features which lie in the same subspace. 2) Following our proposed model, we present extensions of SSC and LRR, i.e., Enhanced Sparse Subspace Clustering (E-SSC) and Enhanced Low Rank Representation (E-LRR). Both E-SSC and E-LRR obtain the local feature similar representation coefficients without losing their discriminability. Extensive experiments show that both E-SSC and E-LRR significantly outperform SSC and LRR, and achieve better performance compared with the state-of-the-art methods for motion segmentation in Hopkins 155 dataset and airport dataset.

The remaining of this paper is organized as follows: Section 2 provides a brief review of existing results on sparse representation and rank minimization for subspace clustering, including the other work related to representation-based methods. In Section 3, we present our proposed model and apply it to two popular representation-based spectral clustering algorithms: SSC and LRR, and formulate E-SSC and E-LRR as their extension, respectively. The application of E-SSC and E-LRR in motion segmentation is described in Section 4. In Section 5, we conclude our work and propose future work.

2. Related works

Many subspace clustering methods for motion segmentation have been proposed in the past two decades. In principle, these algorithms can be roughly divided into four categories: algebraic methods [24], iterative methods [25], statistical methods [26] and spectral clustering-based methods [19,20,22,23,27–31]. A review of these methods can be found in [17]. Most recently, spectral clustering-based methods play a dominant role in subspace clustering problem. In the following, we will focus on some major spectral clustering-based approaches.

Sparse Subspace Clustering (SSC) [20,22] expresses each feature data as a linear combination of all other features, where the combination coefficients are required to be sparse and achieves state-of-the-art performance in subspace clustering. However, SSC treats each sample individually in the sparse representation computation, and no global prior on the affinity matrix is considered. Though the data are from the same subspace and are highly correlated, the SSC will generally ignore this correlation information and most probably select other uncorrelated feature points, even in different subspace, to act as encoding items to represent each sample. This leads to a sparse solution but misses data correlation information, i.e., the local feature similarity. Thus SSC may result in a sparse affinity matrix but lead to unsatisfactory performance. In order to capture the global prior that the union of the underlying subspaces is still low-dimensional, Liu et al. [19,23] propose the Low Rank Representation (LRR) for subspace segmentation, which enforces the constructed affinity matrix to be low-rank. The grouping effect ensures that the highly corrected data which are usually from the same subspace can be grouped together. Unfortunately, however, minimizing the rank of a matrix is known to be NP-hard and a very challenging problem. LRR adopted nuclear norm as a convex relaxation of the rank function. In contrast to LRR, most recently, Kang et al. [32] propose to use a particular log-determinant function to approximate the rank function, which is a tighter rank approximation function than the nuclear norm, and the derived method is called CLAR (Clustering with Log-determinant Approximation to Rank). Lu et al. [27] propose a Least Square Regression (LSR) based method for the affinity matrix construction. It is claimed that grouping effect brought by least square regression for the samples from the same subspace is able to improve the performance of subspace segmentation. Like the LSR, [28] proposes the Correlation Adaptive Subspace Segmentation (CASS) method by using trace Lasso, which simultaneously performs automatic data selection and groups correlated data together. Another prior is introduced by Zhang et al. in [29]. They propose to use correntropy as a robust regularization term to enforce a block-diagonal structure for the affinity matrix. On the other hand, directly pursuing the block-diagonal structure by a graph Laplacian constraint is adopted by Feng et al. in [30]. They propose a rank constraint on the graph Laplacian matrix to effectively generate