Pattern Recognition ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Pattern Recognition

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

Efficient clustering on Riemannian manifolds: A kernelised random projection approach

Kun Zhao [a,*], Azadeh Alavi [b], Arnold Wiliem [a], Brian C. Lovell [a]

[a] School of ITEE, The University of Queensland, St Lucia, Brisbane, QLD 4072, Australia
[b] Center for Automation Research, University of Maryland, College Park, MD 20742-3275, United States

ARTICLE INFO

Article history:
Received 19 December 2014
Received in revised form
2 September 2015
Accepted 9 September 2015

Keywords:
Riemannian manifolds
Random projection
Clustering

ABSTRACT

Reformulating computer vision problems over Riemannian manifolds has demonstrated superior performance in various computer vision applications. This is because visual data often forms a special structure lying on a lower dimensional space embedded in a higher dimensional space. However, since these manifolds belong to non-Euclidean topological spaces, exploiting their structures is computationally expensive, especially when one considers the clustering analysis of massive amounts of data. To this end, we propose an efficient framework to address the clustering problem on Riemannian manifolds. This framework implements random projections for manifold points via kernel space, which can preserve the geometric structure of the original space, but is computationally efficient. Here, we introduce three methods that follow our framework. We then validate our framework on several computer vision applications by comparing against popular clustering methods on Riemannian manifolds. Experimental results demonstrate that our framework maintains the performance of the clustering whilst massively reducing computational complexity by over two orders of magnitude in some cases.

1. Introduction

Clustering analysis is an automated process that groups unlabelled data into subsets (here called clusters) that may express the underlying structure of the data. It is one of the most critical tools for understanding visual data [1,2]. For instance, significant amounts of visual data such as videos and pictures are uploaded every second [3]. Indeed, this is the case for YouTube where 100 h of video are uploaded every minute [4]. Although these videos have titles and some additional meta-information, it is often desirable to automatically group the videos in terms of specific criteria such as visual similarity or detected objects.

In recent years, modelling visual data in analytical manifolds such as Riemannian manifolds has enjoyed success in various computer vision application domains such as face recognition [5], action recognition [6] and pedestrian detection [7]. This is because visual features and models often possess special structures which Euclidean space fails to capture. Riemannian manifolds which form curved spaces are a more appropriate approach to model problems in various computer vision tasks.

Unfortunately, despite the fact that clustering methods have been studied since the 1950s [1,8], applying such methods directly on data represented on Riemannian manifolds is not trivial. Riemannian manifolds generally do not conform to Euclidean space [5,9]. To address this, one could use manifold tangent spaces which are locally homeomorphic to Euclidean space [9]. However, this brings another challenge to applying existing clustering algorithms as some general algebraic operations are not well defined [10]. For instance, *K*-means requires the computation of the mean within a cluster which cannot be computed directly. To this end, Pennec et al. [10] reformulated the computation of mean as a solution to an optimisation problem. Using this formulation, the mean point is considered as the point over the manifold minimising the geodesic distance (i.e. the true distance on the manifold between two points) from the mean point to all other points. The algorithm to solve this problem is called Karcher mean [10]. Thanks to the Karcher mean, Turaga et al. [5] extended the *K*-means algorithm into the Riemannian manifold, which is regarded as intrinsic *K*-means and has been applied to activity-based video clustering. Intrinsic *K*-means has further demonstrated better performance than Euclidean-based methods (for example, Protein Clustering [11]).

Generally, methods that completely honour the manifold topology lead to higher accuracy. We shall categorise these methods as intrinsic methods. Unfortunately, the computational cost of intrinsic methods is extremely high since these need to map all of the data to tangent spaces repeatedly.

* Corresponding author. Tel.: +61 7 33658302; fax: +61 7 3365 4999.
E-mail address: kun.zhao@uq.net.au (K. Zhao).

http://dx.doi.org/10.1016/j.patcog.2015.09.017
0031-3203/

Extrinsic methods, on the other hand, seek solutions that may not completely consider the manifold topology [12–17]. The most simplistic way, here called Log Euclidean methods, is to embed all of the points into a designated tangent space at the identity point [18]. Log Euclidean methods can be considered as flattening the manifold space. It has been used in various computer vision applications, such as human action recognition [13] and cell classification [14]. This addresses the computational cost issues suffered by the intrinsic methods, as the tangent space is homeomorphic to the Euclidean space and well-known Euclidean clustering approaches such as *K*-means can be directly applied. Unfortunately, as the flattening step distorts the pair-wise distances in regions far from the origin of the tangent space, accuracy is severely compromised. So much of the value of the manifold approach is lost.

Other approaches that fall in the extrinsic method category are kernel-based approaches [15–17], such as Kernel *K*-means. In essence, the data in manifold space are first embedded into the Reproducing Kernel Hilbert Space (RKHS) [19]. As the embedding function is defined implicitly, generally kernel-based approaches make use of the inner products in the RKHS in their formulation. These inner products are then arranged in a Gram matrix. It is often observed that the right choice of kernel could significantly improve the performance [15]. Furthermore, in general, kernel inner products with specified metrics have much less computational complexity than geodesic distances [20,21]. With these properties, kernel-based approaches could be suitable to address issues suffered in both the intrinsic approach and the Log Euclidean approach. Unfortunately, the kernel-based approaches cannot scale easily, as the Gram matrix computation is $O(n^2)$ where n is the number of data points. Also, it is often quite challenging to kernelise the existing algorithms that do not have known kernelised versions [22]. Furthermore, Nikhil et al. demonstrate that clustering data in the RKHS may lead to unexpected results since the clusters obtained in the RKHS may not exhibit the structure of the original data [23].

Contributions: We summarise the advantages and shortcomings of the existing approaches in Table 1. Our goal is to develop an efficient clustering algorithm for Riemannian manifolds, which significantly reduces the computational complexity, but still maintains acceptable performance. The inspirations are drawn from the random projection for Euclidean spaces which has enjoyed success in various domains [24–26] due to its simplicity and theoretical guarantees [27]. We list our contributions as follows:

1. We propose a random projection framework for manifold features. In general, the term projection is not well defined in Riemannian manifolds. Therefore, we address this via the RKHS constructed from a small subset of data. Once projected, we choose to apply the *K*-means algorithm.
2. From our framework, it becomes clear that random hyperplane generation is essential. Thus, we describe three generation algorithms which are followed in our framework: (1) Kernelised Gaussian Random Projection (KGRP); (2) Kernelised Orthonormal Random Projection (KORP) and (3) Kernel Principal Component Analysis Random Projection (KPCA-RP).

Table 1
Summary of the existing works compared to our proposal.

Approach	Exploits manifold structure	Accuracy	Computational complexity
Intrinsic methods [5,11]	Yes	High	High
Log-Euclidean methods [12–14]	Minimal	Low	Low
Kernel methods [15–17]	Approximately	High	Moderate
Our proposal	Approximately	High	Low

We note that our method is different from manifold learning approaches for clustering analysis described in [28]. Manifold learning is the collection of non-linear dimensionality reduction (NLDR) techniques that seek for a low dimensional representation of a set of high-dimensional points lying on a non-linear manifold [29]. They assume that the structure of the underlying manifold was unknown. Contrary to this, in our paper, we are interested in Riemannian manifolds whose underlying geometry is known.

We continue the paper as follows. Section 2 provides a brief mathematical background of Riemannian manifolds. Section 3 details the proposed random projection framework for manifold points and develops three different random projection methods for clustering points on manifold spaces. The proposed methods are then contrasted with the state-of-the-art methods in Section 4. The conclusions and future directions are summarised in Section 5.

2. The geometry of Riemannian manifolds

A differentiable manifold $\mathcal{M}$ is a topological space that is locally similar to Euclidean space [30]. One can use the tangent space to model the neighbourhood structure on a differentiable manifold. The tangent space at a point $\boldsymbol{X}$ on the manifold, $\boldsymbol{T}_{\boldsymbol{X}}\mathcal{M}$, is a vector space that contains all possible directions tangentially passing through $\boldsymbol{X}$ [30].

A Riemannian manifold is a differentiable manifold, endowed with a Riemannian metric. The Riemannian metric is the family of inner products on all of the tangent spaces [31]. This metric enables us to define geometric concepts such as lengths, angles and distances. The geodesic distance between two points $\boldsymbol{X}, \boldsymbol{Y}$ is defined as the length of the shortest curve between $\boldsymbol{X}$ and $\boldsymbol{Y}$ [31].

In this section, we briefly introduce two well known Riemannian manifolds used in the computer vision community, namely Symmetric Positive Definite (SPD) manifold and Grassmannian manifold.

2.1. SPD manifolds

To compute a compact representation of an image, one method is to calculate the covariance matrix of a set of *d*-dimensional vector features extracted from the image [32]. Covariance matrices naturally arise in the form of SPD matrices, which can be considered as points on SPD manifolds [7]. The geodesic distance between points on SPD manifolds then can be calculated through an affine invariant Riemannian metric:

$$\text{dist } G^2(\boldsymbol{X}, \boldsymbol{Y}) = \|\log(\boldsymbol{X}^{-1/2}\boldsymbol{Y}\boldsymbol{X}^{-1/2})\|_F^2, \quad (1)$$

where $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{M}$ are two points over the SPD manifold. For further discussions on SPD manifolds, the readers are referred to [9].

To further improve clustering performance, SPD manifolds could be projected into RKHS by Mercer kernels. In this paper, we use one of the popular kernels for SPD manifolds, namely the Gaussian kernel, which is defined by

$$K(\boldsymbol{X}, \boldsymbol{Y}) = \exp(-\beta \cdot \text{dist}^2(\boldsymbol{X}, \boldsymbol{Y})), \quad (2)$$

where $\text{dist}(\boldsymbol{X}, \boldsymbol{Y})$ is the distance between point $\boldsymbol{X}$ and $\boldsymbol{Y}$. Note that, the Gaussian kernel can only be a positive definite kernel when dist(X,Y) is from a flat metric space [65]. Therefore, we use two popular distance functions:. In this paper, we use two popular approximate distance functions: Log Euclidean Distance (LED) [18] and Stein Divergence (SD) [33]. The Gaussian kernel with LED and

Please cite this article as: K. Zhao, et al., Efficient clustering on Riemannian manifolds: A kernelised random projection approach, Pattern Recognition (2015), http://dx.doi.org/10.1016/j.patcog.2015.09.017