Pattern Recognition Letters 80 (2016) 121–128

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Explicit discriminative representation for improved classification of manifold features[☆]

Arnold Wiliem[a,*], Raviteja Vemulapalli[b], Brian C. Lovell[a]

[a] School of Information Technology and Electrical Engineering, The University of Queensland, Australia
[b] Center for Automation Research, University of Maryland, United States

ARTICLE INFO

Article history:
Received 7 January 2016
Available online 22 June 2016

Keywords:
Riemannian manifolds
Grassmann manifold
Symmetric positive definite manifold
Kernel methods
Empirical kernel feature space

ABSTRACT

We tackle the problemof extracting explicit discriminative feature representation for manifold features. Manifold features have already been shown to have excellent performance in a number of image/video classification tasks. Nevertheless, as most manifold features lie in a non-Euclidean space, the existing machineries operating in Euclidean space are not applicable. The proposed explicit feature representation enables us to use the existing Euclidean machineries, significantly reducing the challenges of processing manifold features. To that end, we first embed the manifold features into a Reproducing Kernel Hilbert Space that can encode the manifold geometry. Then, we extract the explicit representation by using the empirical kernel feature space, an explicit lower dimensional space wherein the inner product is equivalent to the corresponding kernel similarity. The final feature representation is then derived from a linear combination of multiple explicit representations from various manifold kernels. We propose a max-margin approach to learn an effective linear combination that will improve the feature discriminative power. Evaluations in various image classification tasks show that the proposed approach consistently and significantly outperforms recent state-of-the-art methods.

1. Introduction

Representing image/video data using manifold features that lie in a non-Euclidean space often naturally arises in various applications such as pedestrian detection in Tuzel et. al. [32] and image set classification in Turaga et. al. [31]. Existing classification approaches such as Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) are not directly applicable to such features due to the non-Euclidean nature of the underlying space. This limits the applicability of existing machineries operating in Euclidean space.

This limitation gives a strong motivation to design a feature representation that is Euclidean but still captures the manifold geometry [11,26]. For example, Hong et. al. [11] proposed a second order statistic based region descriptor, named Sigma Set. The Sigma Set descriptor lies in Euclidean space and has similarity to covariance descriptors which have Riemannian structure. Unfortunately, in most cases, these works are specific to a particular manifold, and thus it is not trivial to extend to the other Riemannian manifolds.

Manifold features can also be accurately studied by first embedding them into a Reproducing Kernel Hilbert Space (RKHS) and then applying kernel-based approaches [10,13,34]. The embedding function, which normally is non-linear, can encode the Riemannian geometry by exploiting typical Riemannian metrics [10,23]. As Euclidean geometry applies in the Hilbert space, this effectively decouples machine learning from data representation [15]. Unfortunately, one may still need significant effort to develop existing methods for kernel space [5]. To that end, Vedaldi et. al. [33] proposed explicit kernel feature maps for the additive class of kernels such as the intersection, Hellinger's and χ^2 kernels. After mapping the input features into the explicit kernel feature space, one can apply any efficient learning method such as the linear SVM. Unfortunately, this approach cannot be applied directly for manifold features as the explicit feature maps of [33] are valid only for features in Euclidean space.

Inspired from the work of Vedaldi et. al. [33], we aim to learn an explicit discriminative feature representation that has Euclidean geometry but still encodes the manifold topological

[☆] This paper has been recommended for acceptance by Egon L. van den Broek

[*] Corresponding author. Tel.: +61 7 3365 1643; fax.: +61 7 3365 4999.
E-mail addresses: a.wiliem@uq.edu.au, arnold.wiliem@ieee.org (A. Wiliem), raviteja@umd.edu (R. Vemulapalli), lovell@itee.uq.edu.au (B.C. Lovell).

http://dx.doi.org/10.1016/j.patrec.2016.06.006
0167-8655/

structure. Having such a feature will significantly reduce the challenges present in processing manifold features. There are two issues that need to be addressed: (1) How to extract the Euclidean features that still encode the manifold topological structure and (2) once extracted, how to improve the feature discriminative power.

Using explicit kernel feature space for manifold features seems an ideal solution for our aim as it is possible to encode manifold geometry into RKHS and hence once mapped into the explicit feature space, the explicit representation still captures the manifold geometry information. Nevertheless, it is still not clear how to develop the explicit kernel feature map for manifold features. Therefore, we opt to use the empirical kernel feature space, first introduced in [27]. The Empirical kernel Feature Space (EFS) is a good approximate solution to the pre-image problem that specifically aims to find the existence of an explicit representation of data points defined in the kernel space [27]. The space can be considered as an explicit lower dimensional Euclidean space wherein the dot product between two points equals the corresponding kernel similarity function. As the dot product is preserved, then the distance is also preserved.

In order to further improve the feature discrimination, we derive our features by combining multiple EFS from various kernels defined over the manifold space. Here, our idea is somewhat similar to the Multiple Kernel Learning (MKL) problem [4,6,18]. The difference is, MKL aims to find discriminative kernel from a convex combination of the given kernels. In our case, we learn the explicit discriminative feature representation from multiple EFS. The optimal combination is jointly learned with a max-margin classifier. We draw our inspiration from the recent method proposed in [36] that describes Multiple Kernel Learning in the empirical kernel feature space using a Modification of the Ho-Kashyap algorithm (MultiK-MHKS). In a similar manner to the canonical correlation analysis, MultiK-MHKS aims at finding a transformation function which maximizes the correlation of explicit feature representations of different kernels. On the other hand, we directly model the learned features as a linear combination of EFS. As will be shown in the experiments section, our approach is more effective than MultiK-MHKS.

It is noteworthy to mention that our work primarily focuses on manifold features whose underlying geometry is known. In contrast to our work, the manifold learning techniques described in Lin et. al. [20] and Roweis and Saul [25] perform non-linear dimensionality reduction wherein the low dimensional projection space lie on a non-linear manifold. In other words, these techniques assume the underlying manifold was unknown.

Contributions: Our main contribution comes from the proposed max-margin learning approaches to extract the explicit discriminative features representation. We show empirically that the extracted features are more discriminative than the recent state-of-the-art approaches. To the best of our knowledge, this is one of the first studies to extract explicit kernel feature space for manifold features. We note that, although our proposed approach could be applied to any other Euclidean features, the benefit of using our approach is more significant when it is applied to manifold features due to the non-linear topological structure presenting significant challenges in processing manifold features. Fig. 1 illustrates our proposal.

Organization: We first discuss related works in Section 2. In Section 3, we describe the procedure to embed data from an RKHS into the corresponding EFS. We describe the proposed approach in Section 4. Section 5 provides a brief overview of two Riemannian manifold features commonly used for computer vision tasks. Section 6 presents our experimental results comparing the proposed approach with various state-of-the-art approaches on four benchmark datasets and Section 7 concludes the paper.

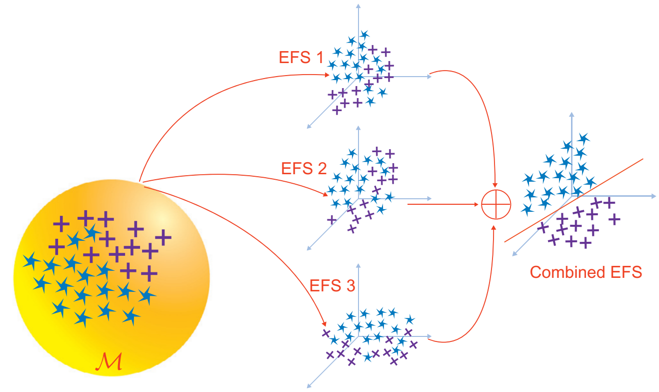

Fig. 1. Illustration on how the proposed explicit discriminative feature representation is extracted. First, each manifold point is mapped into several **E**mpirical kernel **F**eature **S**pace (EFS). Then, the final representation is derived from a linear combination of EFS representations.

2. Related works

Euclidean features for Riemannian manifolds: There are various works focusing on extracting Euclidean features that encode manifold topology [11,26]. Similar to the Sigma Set, San Biagio et. al. proposed descriptors that capture non-linear relationships which are not captured by covariance descriptor [26]. Their descriptors are also Euclidean which makes it convenient to process. Unfortunately, most of these works primarily aim at a particular manifold such as the Symmetric Positive Definite (SPD) manifolds. Another line of work is to flatten the manifold space by embedding the manifold points into a designated tangent space [1]. Nevertheless, the tangent space embedding may introduce significant distortions that could affect the classification.

Fusing multiple views: The idea of fusing information from multiple sources to achieve better performance has been extensively studied both algorithmically and theoretically [29]. For instance, when each source of information can be represented using a kernel (also referred to as base kernel), one can use Multiple Kernel Learning (MKL) which selects a kernel from a family of convex combination of M based kernels and learns a predictor based on the selected kernel [4,6,18]. These two tasks can be performed either individually in stages [6] or simultaneously [18].

Our aim is similar to MKL framework as we aim to combine information extracted from multiple kernels and learn a predictor from the combined space. On the other hand, our approach differs from MKL and multi-view learning in two aspects: (1) the combination is done in the empirical feature spaces of the associated kernels, and (2) unlike the multi-view learning methods, our proposal only learns one predictor.

Multiple Kernel Learning in manifold space: The idea of combining multiple kernels has also been studied for Riemannian manifolds with known geometry [14,34]. For instance, Vemulapalli et. al. [34] used the manifold structure as a regularizer. More precisely, they found a linear combination of basis kernels such that the distance induced by the combined kernel is close to the manifold geodesic distance. Unfortunately, the regularization may prevent the method from picking useful kernels to achieve better performance in a given task. In [14], Jayasumana et. al. focused on selecting the optimal parameters for radial basis kernels on Riemannian manifolds. They specifically showed that radial kernels can be expressed in a simple parametric form. Then a linear SVM framework could be used to automatically obtain optimal parameters for the kernels. Different from their work, we mainly focus on extracting explicit Euclidean features that still encode manifold geometry and improve the discrimination by combining multiple