Pattern Recognition Letters 000 (2016) 1–8

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Linear discrimination dictionary learning for shape descriptors☆

Meng Wang [a,b,c], Jin Xie [a,b], Fan Zhu [a,b], Yi Fang [a,b,d,*]

[a] NYU Multimedia and Visual Computing Lab, USA
[b] Department of Electrical and Computer Engineering, NYU Abu Dhabi, UAE
[c] Department of Computer Science and Engineering, NYU Tandon School of Engineering, USA
[d] Department of Electrical and Computer Engineering, NYU Tandon School of Engineering, USA

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Linear discriminant analysis
Dictionary learning
Shape retrieval

ABSTRACT

The complexity and variation of 3D models have posed a lot of challenges in 3D shape retrieval area, for example, the invariant representation and retrieval of nonrigid and noisy 3D shapes. This paper proposed a supervised dictionary learning scheme called Linear Discrimination Dictionary Learning (LDDL) which can learn shape representations that are insensitive to 3D shape deformations in the same category and different for shapes from different categories in the meantime. Besides, it can extract the subtle differences between 3D shapes for fine-grained shapes. To be specific, in this paper, category-specific dictionaries are learnt to encode subtle visual differences of shapes among different categories, a shared dictionary is learnt to encode common patterns of shapes among all the categories; with the Linear Discriminant Analysis (LDA) constraint on the learnt descriptors, the new descriptors can have small within-class scatter and big between-class scatter. Our method is efficient in training and can obtain promising shape retrieval performance on representative shape benchmark datasets.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Background

The advancement of 3D shape modeling, digitizing and visualizing techniques has given rise to an explosion of 3D models, which leads to a growing demand of efficient retrieval of desired shapes [4,5,8,10,28,37]. 3D shape retrieval, which aims to find similar shapes provided a shape as the query, is an important task in computer vision with broad applications in many fields, such as engineering, entertainment, and medical imaging [36].

In 3D shape retrieval, shape descriptors, compact and informative representations of shapes are often used for the retrieval of a given query shape by comparing the similarities of corresponding shape descriptors. Traditionally, the shapes are represented with vertices and triangular faces of the shape surfaces. However, the traditional representations are only intended for shape visualization, which are inappropriate to use in 3D shape retrieval. To address this issue, an informative shape descriptor is usually taken as a better representation of 3D shape, which can effectively represent the essential features of the shapes. 3D shape retrieval is particularly challenging when the shapes are undergoing large non-rigid deformations, noise and incompleteness [19,36]. In this case, it is very useful to find a shape descriptor which is insensitive and invariant to transformations. What's more, it is even better if the descriptors are insensitive to topological and numerical noises. It is also challenging when the shape descriptors have much in common and the effect of subtle differences between different categories of shapes are weakened. In this case, a descriptor is useful if it models the subtle differences between different categories of shapes.

1.2. Related work

There have been a lot of research works on 3D shape descriptors to address those challenges. In general, two kinds of approaches exist for 3D shape retrieval. One is to develop better hand-crafted shape descriptors; the other is learning-based shape descriptor.

1.2.1. Hand-crafted 3D shape descriptor

Hand-crafted shape descriptors are compact representations of 3D shape which extract geometric essence from 3D shapes [15]. In the past decades, several hand-crafted shape descriptors have been presented, for example, D2 shape distribution [28], model statistical moments [32], Fourier descriptor [13], Eigenvalue descriptor [20], and histogram of local orientation(HLO) [27]. An efficient

☆ This paper has been recommended for acceptance by Xiang Bai.

* Corresponding author at: Department of Electrical and Computer Engineering, NYU Abu Dhabi, UAE. Tel.: +971 3075090310.
E-mail address: yfang@nyu.edu (Y. Fang).

http://dx.doi.org/10.1016/j.patrec.2016.05.028
0167-8655/© 2016 Elsevier B.V. All rights reserved.

re-ranking method Sparse Contextual Activation (SCA) [3] is also presented to shape retrieval. Although these descriptors can extract geometric features from 3D shapes, they are sensitive to non-rigid or topological transformations. To address this problem, some local geometric based descriptors are proposed. Local geometric descriptors such as spin image [2], shape context [7], and mesh HOG [42] extract local geometric features from 3D shapes to achieve isometric transformation-invariance. However, these methods tend to be sensitive to local geometric noise or topological changes.

Apart from these local geometric descriptors, the diffusion-based descriptors are very popular nowadays. These diffusion-based methods are mostly based on an important property of Laplace–Beltrami (LB) operator on 3D surface: Eigenfuctions of LB operator are insensitive to isometric deformations of shapes. Based on this property, Global Point Signature (GPS) [31] is proposed to characterize each vertex with a high-dimensional vector of scaled eigenfunctions of the LB operator calculated at that vertex. However, it suffers the eigenfunction switching problem when eigenvalues are too close [24]. The heat kernel signature (HKS) [35] uses the heat kernel to represent the shape descriptor and it is defined in terms of eigenvalues and eigenfunctions of LB operator over shapes to obtain isometric-invariance. To remove the influence of shape scales, a scale invariant heat kernel signature [11] is proposed, removing the dependence of area of voronoi cell in the LB operator over shapes.

1.2.2. Learnt shape descriptor

The hand-crafted shape descriptors are sensitive to some common 3D shape variations, such as noise and structural changes. Learning discriminative descriptors from a large library of 3D models is a popular way to construct robust shape descriptors [9,25]. There have been a lot of works on learning shape descriptors. Some neural network learning based methods like convolutional neural network [22], deep belief network [18], auto-encoder [39,44] based methods have been used to learn shape descriptors. For example, DeepPano [33] is a robust representation of 3D shapes learnt using deep convolutional neural networks; a multi-view Convolutional Neural Networks [34] is presented for 3D shape recognition. However, these methods suffer the problems of computational complexity and uncertainty of parameter selections. And view-based methods normally have the requirement that the models are consistently upright oriented, which could not be met sometimes. Another kind of learning method is dictionary learning. The purpose of dictionary learning is to learn a new space from the training samples where the codings will better represent the original descriptors [26]. A lot of dictionary learning methods have been applied to image processing and computer vision [1,21,43]. In some methods, training samples are set as the dictionary, however, it may not work effectively for noisy data and outliers in the original training set. In addition, using all the training samples increases the coding complexity and computing cost, and it cannot fully take advantage of the discriminative information hidden in the samples. Another method is learning over-complete dictionary from the training set [1]. KSVD is a decomposition-based method to learn an over-complete dictionary to sparsely represent the original descriptors [1]. However, the learnt dictionary does not have any discriminative abilities because it only pursue the fidelity and low-reconstruction error of the original descriptors. DKSVD [43] is proposed to acquire discriminative abilities for different classes. However, by using a common dictionary shared by all classes, it fails to separate different categories in the learnt dictionary. LCKSVD [21] is proposed to learn a category-specific and label-consistent dictionary to distinguish different classes. Yang et al. [40] proposes an approach using Fisher discrimination criterion to learn a category-specific and label-consistent dictionary to make the dictionary discriminative. Currently in the processing of 3D shapes, bag of features [9] is used to learn global shape descriptors based on local vertices of 3D shapes. The dictionary is learnt by K-means clustering and quantization of local descriptors. A geometric-aware bag of feature [12] method is proposed to build features without losing the geometric information. Bai [5] proposes a new way of representing spatial arrangement of views for 3D models when learning the codewords. To enhance distinctive power to the learnt dictionary, Litman et al. [25] replace the K-means clustering with sparse coding based dictionary learning to get better discriminative performance.

1.3. Our descriptor

We propose to learn category-specific dictionaries to encode subtle visual differences among different categories, and a shared dictionary to encode common patterns among all the categories to gain both representative power and discriminative power. With the LDA constraint on the coefficients of dictionary learning, the new learnt coefficients have small within-class scatter and big between-class scatter. The main contributions of this paper are summarized as follows: (i) With the sparse coding scheme, we construct a structured dictionary to represent 3D shapes, including the shared common sub-dictionary and class-specified sub-dictionaries. (ii) Based on the constructed dictionary, we impose the LDA constraint to representation coefficients to learn a novel 3D shape descriptor that is insensitive to different deformations (iii) Our method is efficient in training and obtain promising shape retrieval performance which is demonstrated on the representative benchmark shape dataset.

The differences between our proposed method and the method in [40] are two-fold. First, in [40], by imposing the LDA constraint on representation coefficients, the authors employed reconstruction error to construct a discriminative sparse coding based classifier for face classification. However, in our proposed method, we impose the LDA constraint to representation coefficients to learn discriminative representation coefficients. Based on the learned representation coefficients, a high-level shape descriptor is formed to represent 3D shape for retrieval. Second, different from the method in [40], we consider the similarities shared by all classes of 3D shapes and their differences to construct a structured dictionary to represent 3D shapes. Based on the constructed dictionary, by imposing the LDA constraint on representation coefficients, we form a novel objective function to learn the discriminative representation coefficients. And the major differences with [6] are: (1) At the low-level, [6] describes shapes using 2D projections, which are sensitive to deformations; on the other hand, the proposed approach uses local shape descriptors to capture invariance of the isometric and scale deformations of 3D shapes. (2) [6] adopts a strategy of pre-division of the codewords. However, this training process requires the prior knowledge of positions of visual words to train the codewords. In 3D shape, it may require that the shape to be segmented or matched first, while the only knowledge we need in our method is the label of the model.

In addition, our method is different from [14,39] in three aspects: (1) The learning strategy used in [14,39] is based on neural network, which builds non-linear mappings between inputs and outputs, while the learning strategy in the proposed method is based on dictionary learning, which builds linear mapping between inputs and outputs. (2) The LDA constraint used in the proposed method is more capable of simultaneously minimizing intra-class distances and maximizing inter-class distances based on the learned 3D shape representations while the learning strategies in [14,39] cannot guarantee max-margin between shape categories. (3) The structured dictionary learning strategy in this work can capture both similarities and dissimilarities across shape categories

Please cite this article as: M. Wang et al., Linear discrimination dictionary learning for shape descriptors, Pattern Recognition Letters (2016), http://dx.doi.org/10.1016/j.patrec.2016.05.028

Download English Version:

https://daneshyari.com/en/article/4970216

Download Persian Version:

https://daneshyari.com/article/4970216

Daneshyari.com