Computers & Graphics 54 (2016) 8–17

Contents lists available at ScienceDirect

Computers & Graphics

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

Special Issue on CAD/Graphics 2015

Multiscale shape context and re-ranking for deformable shape retrieval

Zongmin Li [a], Zhenzhong Kuang [b,*], Yujie Liu [a], Jiayan Wang [b]

[a] College of Computer and Communication Engineering, China University of Petroleum, No. 66, Changjiang West Road, Huangdao District, Qingdao 266580, China
[b] School of Geosciences, China University of Petroleum, No. 66, Changjiang West Road, Huangdao District, Qingdao 266580, China

ARTICLE INFO

Article history:
Received 12 April 2015
Received in revised form
15 June 2015
Accepted 2 July 2015
Available online 15 July 2015

Keywords:
3D shape retrieval
Multiscale shape context
Local descriptor
BOW
Re-ranking

ABSTRACT

This paper proposes a distinct multiscale shape context (MSC) method for isometric 3D shape description and retrieval. For each feature point, a MSC descriptor is devised to capture the multiple spatial information on the basis of the intrinsic shape context, which is advantageous in solving the domain offset deficiency for intra-class shapes. Different from the traditional shape context method, the MSC descriptor is built based on the charts without angular bins and the shape distributions in local domains, which makes it not only simple but also efficient. To reduce the cost of shape representation, we detect a sparse set of feature points and design an improved bags-of-word model to encode the MSC descriptors. For retrieval improvement, an efficient while robust re-ranking algorithm by metric mapping is designed to alleviate the errors of the feature space. Finally, the experimental results have demonstrated significant performance gains on two public benchmarks.

1. Introduction

Over the past few decades, the analysis and recognition of non-rigid shapes has attracted wide attention from the researchers [1–5], where feature extraction plays a critical role. Initially, the global feature was widely used. Elad and Kimmel [6] proposed a bending invariant signature by the multi-dimensional scaling (MDS) procedure of the geodesic distance matrix. Smeets et al. [7,1] used the spectra of the geodesic distance matrix as shape feature (called SD-GDM). To perform partial shape matching and recognition, many of the researchers turned to the local shape descriptors, like the segmentation based method [8,9], the depth image based method (BF-DSIFT-E [1], PCA-based VLAT [10]). Local descriptors based on diffusion geometry also presented promising properties (e.g. isometric invariant) for shape analysis with promising results [2], such as heat kernel signature (HKS) [11] and scale invariant HKS (SIHKS) in [12]. To encode the local descriptors, the bags-of-word (BOW) model is widely used with many successful applications [13–15]. Besides, many other techniques would also be helpful for object recognition, such as the collaborative classification [16], the graph-based approach [10] and the covariance matrix [4].

The lack of contextual information is one of the critical deficiencies for the above descriptors. Although many spatial strategies were proposed [2], they suffer from additional computation and storage costs. Recently, Kokkinos et al. [17] proposed an intrinsic shape context (ISC) descriptor for deformable shapes, which is an extension of the original shape context (SC) descriptor [18] in planar field to 3D domain. The shape context chart of a starting node could grasp the local information naturally by structure, which is promising in providing context information for local point descriptors (e.g. HKS and SIHKS). However, the ISC descriptor [17] suffers from the problem of domain (or region) offset when constructing the chart-based descriptor for each keypoint, which would lead to the invariant description unreliable. This would further result in more errors and mismatchings for the correspondence-based recognition (which is usually cost expensive). Besides, the local descriptor SIHKS of ISC is problematic in describing the local domain because it is uneasy to find a visually intuitive way to perform scale selection for keypoints at different parts as well as different shape conditions. Another up-to-date work by Hedi et al. [4] employed the covariance matrices of the crafted descriptors rather than the descriptors themselves to capture the spatial information. But the approach suffers from three constrains: (1) the spatial information is insufficient due to the covariance matrix could only provide the feature correlation instead of the physical relationship; (2) the descriptor requires the local features to correlate with each other; (3) the method considers few about the elimination of the domain offset.

To obtain better retrieval results, the dataset contexts that can be utilized freely without supervision is discussed for offline optimization. Regardless of errors, the similarity between intra-class objects is small while it is large for inter-class objects, which would provide some evidence for the similarity results of a given

* Corresponding author.
E-mail addresses: lizongmin@upc.edu.cn (Z. Li), zzkuang@foxmail.com (Z. Kuang), liuyujie@upc.edu.cn (Y. Liu).

http://dx.doi.org/10.1016/j.cag.2015.07.002
0097-8493/

query. The idea has been proved to be feasible with many well designed solutions [19–21]. Although Donoser et al. [22] presented a systematic discussion of the existing methods, their precision and robustness need to be further enhanced.

This paper is inspired most by the intrinsic shape context descriptor [17], multiscale HKS [11], shape distributions [23] and the diffusion geometry [22,24]. We concentrate on multiscale shape contexts for deformable shape retrieval. Our approach does not rely on any prior models of any object classes, and is fully automated and unsupervised. Besides, it has multiple advantages for shape description: (i) multi-scale, spatial sensitive and discriminative; (ii) it avoids the domain offset problem; (iii) it has superior clustering property for intra-class shapes, which favors re-ranking a lot. Different from the covariance method [4], our method simply encodes the spatial information in shape feature and proposes a offset invariant method for the BOW model, which provides a better characterization for shape representation. Fig. 1 shows a brief introduction of our method and the contributions are summarized as follows. (1) We propose a new multiscale shape context (MSC) descriptor for shape retrieval. (2) To alleviate the scale offset, we propose to regard each scale of MSC as an independent descriptor for BOW codebook learning. (3) We detect a sparse number of keypoints for each shape to reduce the cost for shape representation, where the dense scale for each keypoint is adopted to deal with the limited number of descriptors. (4) To achieve better retrieval results, we further discuss the robust metric mapping method for retrieval re-ranking. Finally, we demonstrate the performance of our method for deformable shape retrieval on two public benchmarks.

The rest of this paper is organized as follows. In Section 3, we introduce a new MSC descriptor. In Section 4, we present how to represent shapes by using MSC. In Section 5, we discuss the re-ranking method on the shape dataset. Finally, we show the experimental results in Section 6 and conclude the paper in Section 7.

2. Related work

The content-based differentiation between 3D objects from different classes is being pursued in a number of established and emerging fields, such as multimedia, entertainment and industry. Deformable 3D shape data currently available has posed interesting challenges for retrieval systems. To distinguish different shape classes, the common factor of existing methods is the use of descriptors that capture the major properties of 3D objects.

Shape distributions [23] is one of the most classical methods for rigid shapes and it usually updates with new distance metrics for some other applications. The method no longer works for deformable shape in the Euclidean space [25]. As an extension, the geodesic distance was employed instead of Euclidean distance to describe the intrinsic property of deformable shape [6,11], which resulted in significant performance lifts [7]. In the recent years, shape distributions based on the spectral distances were considered to realize robust representation [24,26].

For partial shape retrieval, the local descriptor together with the BOW model is quite popular for non-rigid shape representation [2,14]. Based on heat diffusion, the multiscale HKS [11] and SIHKS [12] were introduced to solve the redundancy of the heat kernel. To test the performance of HKS and SIHKS, the BOW model was used for shape retrieval based on the densely distributed descriptors [2]. Unfortunately, both HKS and SIHKS suffer from some drawbacks for shape retrieval: the multiscale property could not provide much spatial information; it is uneasy to find a visually intuitive way to perform scale selection; the dense descriptors are required to be evenly distributed. Toldo et al. [9] proposed to segment the shape into regions and represent each region by four types of descriptors. And then each shape was modeled as an occurrence histogram of sub-parts based on the learned codebooks. Although the method achieved good results, it depends highly on the segmentation and the topology. In [14], the authors proposed a parameter-free local Fourier descriptor and the BOW model was trained based on the uniform sampling of the feature points. Then, they introduced a further improvement version of the approach called Hybrid BOW [3]. Kokkinos et al. [17] designed an intrinsic shape context descriptor for deformable shape retrieval, but only the shape correspondence results were presented to verify its performance. In [27], an enhanced pose normalization method was studied for non-rigid shape retrieval with the help of the PANORAMA descriptor. Tabia et al. [4] proposed a covariance method for shape retrieval on the basis of covariance matrices and the BOW framework, which achieved state-of-the-art performance.

Recently, lots of attention have been paid to optimize the retrieval results by considering the latent structure of the dataset [21,19]. The basic idea of this technology is to discover the true similarities between dataset objects. The problem is that the similarity relationships between some objects are non-ideal due to inappropriate similarity metrics or features. The affinity learning by diffusion has been proved quite efficient [21], among which the iterative random walks method has been widely researched by using the K-nearest neighborhood (KNN) graph. The authors of [19] introduced a self-smoothing (SSO) operator to improve a given similarity result. Wang and Tu [20] extended the approach by self-diffusion (SD) for image segmentation and clustering. To obtain a good neighborhood for affinity diffusion, Yang et al. [21] utilized the tensor product graph to incorporate high order relations. An up-to-date work [22] summarized the prior works and presented a diffusion process for retrieval revisit. But, the sensitivity to topology change has constrained its real application.

3. The multiscale shape context

This section introduces a new multiscale shape context (MSC) descriptor for shape representation, including the local descriptor of each scale and the keypoint detection algorithm.

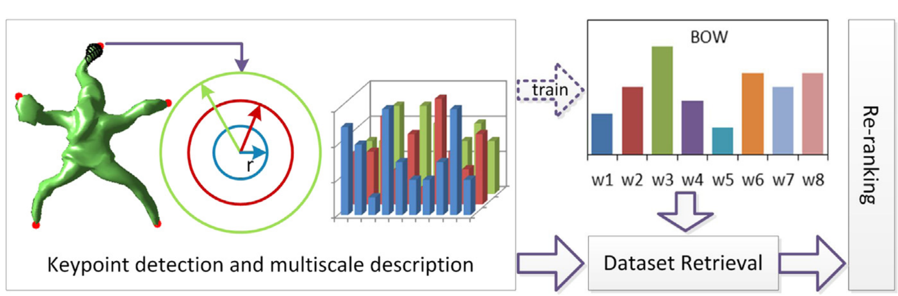

Fig. 1. Brief illustration of the proposed shape retrieval method (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.).