Pattern Recognition 71 (2017) 414–427

Contents lists available at ScienceDirect

Pattern Recognition

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

Scale space clustering evolution for salient region detection on 3D deformable shapes

Xupeng Wang[a,b,*], Ferdous Sohel[c], Mohammed Bennamoun[b], Yulan Guo[d,b], Hang Lei[a]

[a] School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China
[b] School of Computer Science and Software Engineering, University of Western Australia, Perth, WA, Australia
[c] School of Engineering and Information Technology, Murdoch University, Perth, WA, Australia
[d] College of Electronic Science and Engineering, National University of Defense Technology, Changsha, China

A R T I C L E I N F O

Article history:
Received 31 January 2017
Revised 5 May 2017
Accepted 20 May 2017
Available online 22 May 2017

Keywords:
Deformable shape segmentation
Salient region detection
Diffusion geometry
Clustering algorithm
Persistent homology

A B S T R A C T

Salient region detection without prior knowledge is a challenging task, especially for 3D deformable shapes. This paper presents a novel framework that relies on clustering of a data set derived from the scale space of the auto diffusion function. It consists of three major techniques: scalar field construction, shape segmentation initialization and salient region detection. We define the scalar field using the auto diffusion function at consecutive time scales to reveal shape features. Initial segmentation of a shape is obtained using persistence-based clustering, which is performed on the scalar field at a large time scale to capture the global shape structure. We propose two measures to assess the clustering both on a global and local level using persistent homology. From these measures, salient regions are detected during the evolution of the scalar field. Experimental results on three popular datasets demonstrate the superior performance of the proposed framework in region detection.

1. Introduction

Salient feature detection and description on three dimensional shapes is a fundamental problem in the field of computer vision and graphics. It has wide applications in surface matching, surface registration, shape recognition and shape retrieval [1,2]. Early research mainly focuse on rigid shapes [3–5]. The problem is much more challenging for non-rigid shapes due to the large degree of local deformations [6,7]. In the past decade, a large number of point based feature detectors and descriptors have been proposed, such as the MeshDOG and MeshHOG [8], Harris 3D [9], heat kernel signature [10], wave kernel signature [11], stable topological signature [12], local binary descriptor based on heat diffusion [13], heat propagation contours [14] and descriptors based on machine learning methods [15,16]. In recent years, an increasing interest has been given to the detection and description of the stable regions on 3D shapes [17–20]. These methods achieved a remarkable success in several applications [21–23], such as shape correspondence and retrieval, due to their higher robustness compared to local surface based descriptors.

Existing region detection approaches focus mainly on the property of stability under different deformations of a shape. As a result, some of their resulting detected shape components do not usually contain informative features [19]. Furthermore, not many works consider the relationship between region detection and human perception. According to cognitive theory, a *salient region* is a perceptually related subset of a shape determined by its relative size and degree of protrusion in comparison with the neighboring surfaces [24]. Salient region detection on three dimensional shapes is very useful, because it provides more insights and facilitates the interpretation of a shape in terms of both geometric and semantic information.

In this paper, we present a novel framework for the detection of salient regions on 3D deformable shapes, which combines two main approaches used for 3D shape analysis [25], i.e., diffusion geometry and persistent homology. This framework first defines a scalar field using the auto diffusion function [26] at consecutive time scales. Being a local surface descriptor derived from the Laplace-Beltrami decomposition, the auto diffusion function is robust to isometric transformations and is able to describe a surface at multiple scales. Therefore, during the evolution of the scalar field, all of the salient regions will appear in the scale space. Based on the constructed scalar field, we exploit the method of persis-

* Corresponding author at: School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China.
E-mail addresses: 201211220104@uestc.std.edu.cn, 1986wxp@gmail.com (X. Wang), F.Sohel@murdoch.edu.au (F. Sohel), mohammed.bennamoun@uwa.edu.au (M. Bennamoun), yulan.guo@nudt.edu.cn (Y. Guo), hlei@uestc.edu.cn (H. Lei).

http://dx.doi.org/10.1016/j.patcog.2017.05.018
0031-3203/

tent homology to extract salient regions. Thereby, we take advantage of the capabilities of the method in multivariate data analysis, i.e., persistence-based clustering [27] and clustering assessment [28]. Persistence-based clustering is performed to produce an initial segmentation of the shape. Persistent homology is then calculated to assess the clustering and discover the newly emerged salient regions during the process. Comparative experiments were performed on three popular datasets to demonstrate the superiority of the proposed method over the existing ones.

The rest of this paper is organized as follows. Section 2 provides a brief literature review of region detection algorithms on deformable shapes. Section 3 analyzes the limitations of the existing persistence-based shape segmentation method, which sets the stage for our approach. Section 4 describes the proposed salient region detection framework. Section 5 presents all evaluation results and analysis of our method. Section 6 concludes the paper.

2. Related work

Salient region detection is a non-trival problem, especially for shapes under isometric transformations. It has been well studied by the shape analysis community in the recent years.

One of the first notable works was proposed by Litman et al. [17], which formulated the problem as seeking the maximally stable components on a shape. This approach exploited the diffusion geometry to derive weighting functions in order to achieve invariance to isometric transformations, and proposed both vertex and edge weighted graph representations of the mesh. The edge weighted graph representation is more general than the vertex weighted representation and shows superior performance. This framework was extended to handle volumetric shapes in [29]. Sipiran et al. [18] considered the salient regions as the “Key-components” on a shape, and assumed that they contain rich discriminative local features. This approach was inspired by the cognitive theory on saliency of visual parts. According to this theory, these salient parts correspond to the regions with a high protrusion, and are detected by a clustering process in the geodesic space. However, this method produces an incomplete decomposition of a shape [19].

Diffusion geometry based methods have achieved a remarkable success as they exploit the intrinsic properties of a shape [10]. Reuter et al. [30] used eigenvectors of the Laplace–Beltrami operator, due to their invariance to isometric transformations, and exploited the persistence diagram to perform a hierarchical segmentation of a shape. In [31], a global point signature is calculated for each point, and the shape is mapped into an intrinsic space. A clustering algorithm is then applied in this space to segment the shape. However, these methods use eigenfunctions of the Laplace–Beltrami operator for segmentation. The use of eigenfunctions suffers from the problems of sign flipping and eigenvectors switch, especially when the difference between corresponding eigenvalues is small [32]. Rodola et al. [19] introduced the idea of consensus clustering into this area to achieve a stable segmentation. They generated a heterogeneous ensemble of segmentations by applying multiple clusterings in the global point signature space. Their work assumes that a robust segmentation can be obtained by gathering statistical information from these segmentations. This approach produces the state-of-the-art results on a wide range of transformations.

More stable variants based on the heat kernel from the theory of diffusion geometry have also been introduced. The persistence-based shape segmentation approach was proposed in [32,33]. This framework first computes the prominence of the basins of attractions that are associated with the local maxima, which are then grouped in the form of a component tree. A persistence threshold is then selected to merge the components and to produce a stable segmentation of the shape. This framework is stable under isometric transformations. However, it uses a vertex weighted function and is relatively less robust to noise compared to the edge weighting scheme [24]. In addition, as discussed in this paper (Section 3.3), the method depends heavily on the selection of the merging parameter, which is hard to derive. Heat walk [24] was proposed to derive a salient and stable segmentation by making full use of the information contained in the heat kernel. It achieves a robust performance and can be considered as an edge weighted method [19].

As noted in [19], the aforementioned approaches focus on the robustness property of a region detector under different transformations of a shape. They are unable to automatically determine the optimal number of components to be extracted. As a result, some of the regions produced by these methods are lacking in features, which limits their applications, e.g., for dense correspondence and shape retrieval. In this paper, we propose a *region detector*, which is stable and at the same time captures the salient regions on a shape. We demonstrate that the regions detected by our method are distinctive and closely related to human perception.

The main contributions of this paper include:

First, we analyze the limitations of the existing persistence-based segmentation methods Section 3), and propose a method to detect the salient regions on deformable shapes. *Second*, we take advantage of the recent progress in persistence homology to achieve a salient shape segmentation (Section 4). *Third*, we provide a comprehensive evaluation of the proposed method to demonstrate its saliency and robustness (Section 5).

3. Deformable shape segmentation using persistence-based clustering

In this section, we give a brief overview of the basic concepts behind our method.

3.1. Persistent homology

Persistent homology has originated from computational topology. It describes a data set using topological features with various dimensions. The topological features consist of *connected components* (dimension 0), *holes* (dimension 1) and *voids* (dimension 2). In this paper, we focus on the 0-dimensional persistent homology because of its significant effectiveness for data clustering [27], which can be directly applied to shape segmentation. The fundamental idea of persistent homology is to provide a framework for data set characterization, which incorporates both topological information and the geometrical properties measured by a function.

Given a manifold $\mathcal{M}$ and a scalar function defined on it $f : \mathcal{M} \to \mathcal{R}$, the function f is assumed to have a finite number of *critical points*, i.e., points where the gradient of the function values vanishes. Fig. 1 gives an example of $\mathcal{M}$ and its associated height function f. Since $\mathcal{M}$ is a 2-dimensional manifold, the critical points include both local extrema and saddles.

Sub-level sets of the scalar function $\mathcal{L}_\alpha(f) = f^{-1}([\alpha, +\infty))$ induces a filtration of the manifold $\mathcal{M}$, i.e., a family of subspaces nested by inclusion. Persistent homology encodes the evolution of connectivity information when the value of function f changes from large to small. The topology of the space created by $\mathcal{L}_\alpha(f)$ changes only at the critical points. At a local maximum, a new connected component emerges in the space. At a local minimum or a saddle, two connected components are merged. As α decreases from $+\infty$ to $-\infty$, a *hierarchy of components* of the manifold is generated. To be consistent with the construction of the component tree, the component with a smaller maximum is merged into the larger one.