RESEARCH

CrossMark

A novel framework for 3D shape retrieval

Ye Jin[*], ZhiXun Li, YingTao Zhang and XiangLong Tang

*Correspondence:
jinye@hit.edu.cn
Harbin Institute
of Technology, Harbin, China

Abstract

The ability to accurately and effectively search for 3D shape is crucial for many applications. In this study, we proposed a novel framework for 3D shape retrieval. We compensate the loss of high frequencies of heat kernel signature from two aspects. One is to introduce the weight for each point to highlight the details of the salient points. The other is to directly capture microgeometry structure through wave kernel's access to high frequencies. Thus, our method can capture geometric features at different frequencies of a shape, which satisfy the property of an ideal descriptor. We conduct shape retrieval experiments on a standard benchmark and compared with another heat kernel-based method. Experimental results demonstrate that the proposed method is effective and accurate.

Keywords: Shape retrieval, Laplace–Beltrami operator, Spectral geometry, Shape descriptor

Background

With advancement in 3D imaging and processing, collections of 3D models have become an increasingly prominent in many fields, such as engineering, entertainment, and medical imaging (Liu et al. 2006). In order to find relevant objects from a given query, measuring the similarity between different objects has become acute. Since many shapes manifest rich variability, retrieval is required to be invariant to different kinds of deformations. One of the most challenging tasks is to deal with non-rigid shapes, in which transformation category is very broad.

The key component of 3D shape analysis is to define a descriptor, for distinguishing different parts of the shapes. Elad (2001) extended a series of rigid descriptors for non-rigid shapes by replacing Euclidean metric space with geodesic distance, which is invariant to inelastic deformations. However, it suffers from a huge flaw because of its strong sensitivity to topological noise, which limits the usefulness of such descriptors (Li and Hamza 2013).

Recently, research efforts have shown growing popularity of spectral analysis of the Laplace–Beltrami (LB) operator. An isometry-invariant global descriptor was proposed by Reuter (2006) which used eigenvalues of LB operator as Shape-DNA for a 3D manifold. Despite the good performance for non-rigid shapes, it cannot be used for local shape analysis. Rustamov (2007) proposed to construct global point signature (GPS) at each point based on diffusion geometry. A major drawback of such a descriptor is the problem of eigenfunctions' switching whenever the associated eigenvalues are close to

Jin *et al. Appl Inform* (2017) 4:3

Page 2 of 7

each other. A remedy was proposed lately by Sun (2009) through constructing heat kernel signature, which was based on the fundamental solutions of the heat equation. It is a point-based signature which has a number of desirable properties, including invariance to isometric transformations, robustness to small perturbations (Abdelrahman et al. 2012), and a multi-scale interpretation (Sun et al. 2009) of the shapes. As of today, this descriptor achieves state-of-art performance in shape retrieval and other applications.

HKS successes in wide applications, however, it highly depends on the information derived from low frequencies corresponding to the global structure of the shape (Aubry et al. 2011). The information of low frequency is effective to discriminate distinct shapes, which usually differ greatly at coarse scales; however, the loss of high-frequency information damages the ability to conduct feature localization precisely (Aflalo et al. 2012).

To solve the problem that high frequencies are avoided, in this study, we proposed a novel framework for shape retrieval by integrating the advantages of both global geometry for discriminative power and microgeometry for localization. In proposed method, not only the macroshape information, but also the microshape information is considered. In the experiment, some cases from the standard benchmark (SHREC 2010) are employed for testing and validating the proposed approach; and a well-studied approach (Ovsjanikov et al. 2009) based on heat kernel signature is utilized for comparison. Experimental results demonstrate that the proposed method is effective and accurate for 3D shape retrieval.

The rest of this study is organized as follows. "The proposed method" section describes our framework in details. "Object representation and matching" section presents our retrieval procedure using Bag of words. "Results and discussion" section shows some experimental results on a 3D shape benchmark. Finally, we conclude in "Conclusions" section.

The proposed method

We proposed a new framework to handle the loss information of heat kernel signature.

Highlight details

The heat kernel $h_t(x, y)$ is the fundamental solution of the heat equation (Sun et al. 2009), which is closely associated with Laplace–Beltrami operator by:

$$\frac{\partial h_t(x, y)}{\partial t} + \Delta h_t(x, y) = 0. \tag{1}$$

It can be further defined in terms of the eigenvalues and eigenfunctions of Δ as follows:

$$h_t(x, y) = \sum_i e^{-\lambda_i t} \phi_i(x) \phi_i(y). \tag{2}$$

Intuitively, $h_t(x, x)$ denotes the amount of heat remaining at point x after time t. Therefore, HKS at point x is represented in the discrete temporal domain by a n-dimensional feature vector:

$$\text{HKS}(x) = [h_{t1}(x, x), h_{t2}(x, x), \dots, h_{tn}(x, x)], \tag{3}$$

where t is the time scale.

Jin *et al. Appl Inform* (2017) 4:3

Page 3 of 7

The function $\exp(-\lambda t)$ is mainly dominated by low frequencies, which correspond to the macrostructure. In this study, we use heat mean signature (HMS) (Fang et al. 2011) to evaluate the weight of a point. The bigger is the value of HMS, the more influential of a point is. Using weight, we can further enhance the importance of salient points so that highlighting their corresponding details.

$$\text{weight}(x) = \text{HMS}_t(x) = \frac{1}{m} \sum_{y \neq x} h_t(x, y) \tag{4}$$

We empirically choose a smaller parameter t to compute the weight of a point. A new descriptor, enhanced heat kernel signature (EHSK) is defined:

$$\text{EHKS}(x) = \text{weight}(x) \cdot \text{HKS}(x) \tag{5}$$

$$= \text{weight}(x) \cdot [h_{t1}(x, x), \ldots, h_{tn}(x, x)]. \tag{6}$$

Microgeometry structure

Besides highlighting details by introducing the weight, we also directly capture microgeometry of a point. In our method, we use the wave kernel. The wave kernel is from Schrodinger's equation (Aubry et al. 2011)

$$\frac{\partial \varphi(x, t)}{\partial t} = i\Delta\varphi(x, t). \tag{7}$$

The wave function $\varphi(x, t)$, which governs the evolution of a quantum particle on the surface, is the solution of Schrodinger's equation and can be further expressed as following:

$$\varphi_e(x, t) = \sum_k e^{i\lambda_k t} \varphi_k(x) f_e(\lambda_k), \tag{8}$$

where e is the energy of the particle at $t = 0$. f_e is the initial distribution. The probability to measure the particle at point x is then $|\varphi_e(x, t)|^2$. Thus, the average probability is obtained by integrating over time as following:

$$P_e(x) = \lim_{T \to \infty} \frac{1}{T} \int_0^T |\varphi_E(x, t)|^2 = \sum_k \varphi_k(x)^2 f_e(\lambda_k)^2, \tag{9}$$

where we use a log-normal energy distribution.

$$f_e(\lambda) \propto \exp\left(-\frac{(\log e - \log \lambda)^2}{2\sigma^2}\right). \tag{10}$$

Finally, the wave kernel signature at point x is defined as:

$$\text{WKS}(x) = (P_{e1}(x), P_{e2}(x), \ldots, P_{en}(x)), \tag{11}$$

where e_i is the logarithmic energy scale. The function $\exp\left(-\frac{(\log e - \log \lambda)^2}{\sigma^2}\right)$ yielding WKS can be considered as band-pass filters. Thus, it provides an access to high frequencies,

Jin *et al. Appl Inform* (2017) 4:3

Page 4 of 7

which corresponds to the microgeometry of a point. Thus, we use wave kernel signature to remedy HKS's poor feature location capability.

Analysis of the proposed method

HKS can capture global geometry of shapes well, but it suppressed the local structure. In order to compensate this, we proposed an approach from two aspects. First, we consider the weights of the points. In shape representation, salient points make significantly greater contribution. After applying the weights, we define the EHKS which will not only to further enhance details of salient points but also maintain HKS's global geometry property. Second, we use WKS as band-pass filters to obtain high-frequency information and thus to capture the local structures of shapes. Thus, we can obtain full information of the shape both local and global, which is critical for shape analysis. In order to avoid overlapping of the two types of descriptors, we create two codebooks based on EHKS and WKS, respectively. Using Bag of words representation, we can finally obtain the global geometry distribution and local geometry distribution based on different codebooks which will be explained in "Object representation and matching" section in detail.

Object representation and matching

Codebook construction

Given a set of point-wise signatures, we use a codebook to represent the distribution of a shape. In our framework, each point has two types of signatures based on EHKS and WKS, respectively. To avoid the influence of different frequencies, we create two codebooks based on different type of descriptors.

To create a codebook $D = \{c_1, c_2, \ldots c_w\}$, we just employ simple k-means clustering on the set of descriptors and use the center of the cluster as a visual word.

Bag of words representation and matching

For a given model M with a set of descriptors $Q = \{q_i, i = 1, 2,\ldots,n\}$, where n is the number of points. The codebook $D = \{c_1, c_2, \ldots c_w\}$ is obtained as discussed above. In this study, we use visual word uncertainty to describe M as the distribution of visual words. Hence, we assign a descriptor q_i over all visual words instead of its nearest visual word. Relevancy between q_i and word c_j is evaluated by a Gaussian kernel $K_\sigma(q_i, c_j)$. The distribution on word c_j is obtained as follow:

$$\mathrm{rel}(c_j) = \sum_{i=1}^{n} K_\sigma(q_i, c_j) \tag{12}$$

According EHKS and codebook D_1—we can obtain distribution histogram h_1—on macro structures using above word uncertainty method. Similarly, we can obtain distribution histogram h_2—on local structures. So, for a given shape X, the fully representation is $h_X = [h_1, h_2]$. To compare two shapes X and Y, we define their distance as follow:

$$d(X, Y) = ||h_X - h_Y||_1. \tag{13}$$

The summary of the proposed approach is given in Fig. 1.

Jin *et al. Appl Inform* (2017) 4:3

Page 5 of 7

Input: 3D database   D and a query shape Q
Output: first r retrieved shapes for Q
1. Compute heat kernel signature, HKS, at each point x , $x \in M$, $M \in D$
2. Compute weight of each point x
3. Compute global descriptor EHKS at each point x
4.   Compute micro descriptor WKS at each point x ,

5. Create codebooks V_1 and V_2 based on EHKS and WKS, respectively

6. Compute distribution histograms h_1 and h_2 based on two codebooks

7.   Compute full representation $h_M = [h_1, h_2]$

8.   Compute $S = \{h_M, M \in D\}$

9.   Compute query Q's full representation h_Q

10. Compute distance $d(Q, M) = |h_Q - h_M|_1$, $\forall h_M \in S$, return the first r shapes
   with minimum distances

Fig. 1 Retrieval procedure of the proposed approach

Results and discussion

Dataset and measures

To test the proposed approach, we conducted experiments on a benchmark: SHREC 2010. The dataset contains three collections: TOSCA, Sumner and Princeton (Bronstein et al. 2011). TOSCA contained seven shape classes, and Sumner contained six shape classes. After applying different transformations on 13 shape classes, the total set size has 596 shapes used as positives, while Princeton contained 347 shapes (exclude the shapes as the positives) used as negatives. Retrieval quality is assessed by following measures. Mean Average Precision (mAP) is defined as:

$$\text{mAP} = \sum_r P(r) \cdot \text{rel}(r), \tag{14}$$

where $P(r)$ is the percentage of relevant shapes in the first r top-ranked retrieved shapes and rel(r) is the relevance of a given rank. False positive rate (FPR) is the percentage of dissimilar shapes wrongfully identified as similar. False negative rate (FNR) is the percentage of similar shapes wrongfully identified as dissimilar. Equal error rate (EER) is the value of FPR at which it equals FNR.

Results

We compared results with the state-of-art method proposed by Ovsjanikov (2009). In this study, we use the largest 200 eigenvalues and eigenfunctions of LB operator. For EHKS, we choose six time scales with $\alpha = 1.32$ and $t_i = 1024 \cdot \alpha^{i-1}$ ($i = 1, 2, ..., 6$). For

Jin *et al. Appl Inform* (2017) 4:3

Page 6 of 7

WKS, we select energy scale $N = 20$ and variance $\sigma = 1$ with the best results through repeating experiments.

Figure 2a and b shows the results of a query "man." Ovsjanikov's method considers only the global geometry, so it cannot discriminate man and woman. Our method is more accurate in capturing both global and local structures. Figure 2c and d shows results of a query "dog." In this query experiment, a large number of noise samples were added. Ovsjanikov's method retrieved 3 wrong shapes, while our method retrived only 1 shape. This demonstrates the robust of our method.

To evaluate accuracy of each transformation, we conducted another experiment. Ovsjanikov's results and our method are shown in Tables 1 and 2, respectively. The higher mAP of our method indicates that the retrieved related shapes have a top ranking, while the low EER denotes better identification capability. The overall superior performance validates the effectiveness and accuracy of the proposed approach.

Conclusions

In this study, we proposed a novel retrieval framework. Althrough enhancing the details and capturing microstructures, this approach can handle the substantial loss of high frequencies of the state-of-art heat kernel signature; therefore, it can obtain both the global geometry and local geometry of the shapes. The experimental results also demonstrate
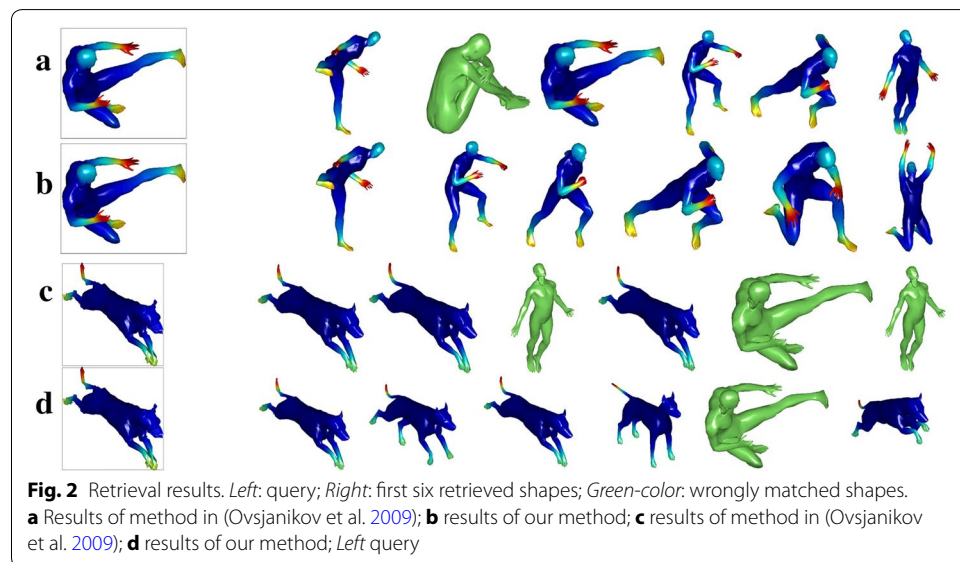

Fig. 2 Retrieval results. *Left*: query; *Right*: first six retrieved shapes; *Green-color*: wrongly matched shapes. **a** Results of method in (Ovsjanikov et al. 2009); **b** results of our method; **c** results of method in (Ovsjanikov et al. 2009); **d** results of our method; *Left* query

Table 1 Results of our method with vocabulary of size 48

Transformation	mAP (%)	EER (%)	FPR@FNR = 1%
Isometry	96.93	0.62	0.42
Isometry + topology	96.55	0.90	0.78
Noise	96.89	1.70	4.56
Partiality	90.34	2.49	4.08
Topology	95.85	0.50	0.22
Triangulation	94.73	1.87	2.54

Jin *et al. Appl Inform* (2017) 4:3

Page 7 of 7

Table 2 Results of Ovsjanikov's method based on heat kernel signature

Transformation	mAP (%)	EER (%)	FPR@FNR = 1%
Isometry	95.66	1.35	1.65
Isometry + topology	94.43	1.95	3.02
Noise	93.09	3.58	12.77
Partiality	88.12	3.95	5.46
Topology	93.99	1.75	2.83
Triangulation	92.68	2.62	4.42

that the proposed method is more accurate and robust compared with state-of-the-art methods. In future work, we would develop a more general and flexible way to obtain the shape's geometric features at different frequencies.

References
Abdelrahman M, El-Melegy et al. (2012) Heat kernels for non-rigid shape retrieval: sparse representation and efficient classification. In: Ninth IEEE conference on computer and robot vision, pp 153–160
Aflalo Y, Bronstein AM et al. (2012) Deformable shape retrieval by learning diffusion kernel. In: International conference on scale space and variational methods in computer vision. Springer, Berlin, pp 689–700
Aubry M, Schlickewei U et al. (2011) The wave kernel signature: a quantum mechanical approach to shape analysis. In: IEEE workshop on computer vision, pp 1626–1633
Bronstein AM, Bronstein MM et al (2011) Shape google: geometric words and expressions for invariant shape retrieval. ACM Trans Graph 30(1):623–636
Elad A, Kimmel R (2001) Bending invariant representations for surfaces. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition, 2001—CVPR 2001, vol 1, pp 1–168
Fang Y, Sun M et al. (2011) Heat-mapping: a robust approach toward perceptually consistent mesh segmentation. In: Proceedings of the 2011 IEEE conference on computer vision and pattern recognition, 2011—CVPR, pp 2145–2152
Li C, Hamza AB (2013) Spatially aggregating spectral descriptors for nonrigid 3D shape retrieval: a comparative survey. Multimedia Syst 20(3):253–281
Liu Y, Zha H, Qin H (2006) Shape topics: A compact representation and new algorithms for 3d partial shape retrieval. CVPR 2025–2032
Ovsjanikov M, Bronstein AM et al. (2009) Shape google: a computer vision approach to isometry invariant shape retrieval. In: IEEE 12th international conference on computer vision workshops, pp 320–327
Reuter M, Wolter FE, Peinecke N (2006) Laplace–Beltrami spectra as 'shape-DNA'of surfaces and solids. Comput Aided Des 38(4):342–366
Rustamov (2007) Laplace–Beltrami eigenfunctions for deformation invariant shape representation. In: Symposium on geometry processing, pp 225–233
Sun J, Ovsjanikov M et al (2009) A concise and provably informative multi-scale signature based on heat diffusion. Comput Graph Forum 28(5):1383–1392