Vis Comput (2013) 29:513–524
DOI 10.1007/s00371-013-0815-3

ORIGINAL ARTICLE

A multiresolution descriptor for deformable 3D shape retrieval

Chunyuan Li · A. Ben Hamza

Published online: 26 April 2013

Abstract In this paper, we present a spectral graph wavelet framework for the analysis and design of efficient shape signatures for nonrigid 3D shape retrieval. Although this work focuses primarily on shape retrieval, our approach is, however, fairly general and can be used to address other 3D shape analysis problems. In a bid to capture the global and local geometry of 3D shapes, we propose a multiresolution signature via a cubic spline wavelet generating kernel. The parameters of the proposed signature can be easily determined as a trade-off between effectiveness and compactness. Experimental results on two standard 3D shape benchmarks demonstrate the much better performance of the proposed shape retrieval approach in comparison with three state-of-the-art methods. Additionally, our approach yields a higher retrieval accuracy when used in conjunction with the intrinsic spatial partition matching.

Keywords Spectral graph wavelet · Laplace–Beltrami · Shape retrieval · Multiresolution

1 Introduction

The content-based differentiation between 3D objects from different classes is being pursued in a number of established and emerging fields, including animation, molecular biology and medicine, computer-aided design, multimedia entertainment, and mobile game development. With the increasing use of 3D scanners and as a result of emerging multimedia computing technologies, large databases of 3D models are distributed freely or commercially on the World Wide Web. The availability and widespread usage of such databases, coupled with the need to explore 3D models in depth as well as in breadth, has sparked the need to organize and search these vast repositories, and efficiently retrieve the most relevant selections. The shape retrieval problem has been extensively investigated in the literature, from comprehensive surveys [1, 2] to comparable benchmarks [3–6].

Over the past few years, there has been a surge of interest in the spectral analysis of the Laplace–Beltrami (LB) operator, resulting in many applications to manifold learning [7], object recognition and deformable shape analysis [8–12]. It is worth pointing out that spherical harmonics [13] are nothing but the LB eigenfunctions on the sphere. The truncated sequence of the LB eigenvalues was proposed by Reuter et al. as an isometry-invariant global shape descriptor, dubbed shape-DNA [9]. Reuter also introduced a Morse-theoretic method for shape segmentation and registration using the topological features of the LB eigenfunctions [10]. These eigenfunctions are computed via a cubic finite element method on triangular meshes, and are arranged in increasing order of their associated eigenvalues. Rustamov [11] proposed a feature descriptor referred to as the global point signature (GPS), which is a vector whose components are scaled eigenfunctions of the LB operator evaluated at each surface point. GPS is invariant under isometric deformations of the shape, but it suffers from the problem of eigenfunctions' switching whenever the associated eigenvalues are close to each other. This problem was lately well handled by the heat kernel signature (HKS) [14], which is a temporal descriptor defined as an exponentially-weighted combination of the LB eigenfunctions. It is a local shape descriptor that has a number of desirable properties, including robustness to small perturbations of the shape, efficiency and invariance to isometric transformations. The idea of HKS

C. Li · A. Ben Hamza (✉)
Concordia Institute for Information Systems Engineering,
Concordia University, Montreal, QC, Canada
e-mail: hamza@ciise.concordia.ca

was also independently proposed by Gębal et al. [15] for 3D shape skeletonization and segmentation under the name of auto diffusion function. Using the Fourier transform's magnitude, Kokkinos et al. introduced the scale-invariant heat kernel signature (SIHKS) [16], which is constructed based on a logarithmically sampled scale-space. Observed from the graph Fourier perspective, HKS is highly dominated by information from low frequencies, which correspond to macroscopic properties of a shape. To give rise to substantially more accurate matching than HKS, the wave kernel signature (WKS) [17] was proposed as an alternative in an effort to allow access to high-frequency information. Despite being physically inspired, both WKS and HKS can be regarded as filters. On the other hand, in order to construct a good task-specific spectral descriptor, one has to be in the position of defining the spectral content of the geometric "signal" and the "noise." Bronstein [18] proposed to learn the signal and noise from examples in a way that resembles the construction of a Wiener filter that passes frequencies containing more signal than noise, while attenuating those where the noise covers the signal.

In this paper, we introduce a wavelet-based signature for nonrigid 3D shape retrieval. Wavelets are a class of a functions used to localize a given function in both space and scaling [19]. The wavelet analysis has some major advantages over Fourier transform, which makes it an interesting alternative for many applications. In particular, unlike the Fourier transform, wavelet analysis is able to perform local analysis and also makes it possible to perform a multiresolution analysis. Classical wavelets are constructed by translating and scaling a mother wavelet, which is used to generate a set of functions through the scaling and translation operations. The wavelet transform coefficients are then obtained by taking the inner product of the input function with the translated and scaled waveforms. The application of wavelets to graphs (or triangle meshes) is, however, problematic and not straightforward due in part to the fact that it is unclear how to apply the scaling operation on a signal (or function) defined on the mesh vertices. To tackle this problem, Coifman and Lafon [20] introduced the diffusion wavelets, which generalize the classical wavelets by allowing for multiscale analysis on graphs. The construction of diffusion wavelets interacts with the underlying graph through repeated applications of a diffusion operator, which induces a scaling process. Recently, Hammond et al. [21] showed that the wavelet transform can be performed in the graph Fourier domain, and proposed a spectral graph wavelet transform that is defined in terms of the eigensystem of the graph Laplacian matrix. More recently, Kim et al. [22] introduced a wavelet-based multiscale descriptor for the analysis of cortical surface signals (such as cortical thickness) using the spectral graph wavelet transform. While building on earlier efforts, we take a rather different approach in this paper by proposing a novel multiresolution shape signature that is not only isometric invariant, but also compact, easy to compute and combines the advantages of both band-pass and low-pass filters.

1.1 Contributions

Our main contributions in this paper may be summarized as follows:

(i) We present a general and flexible framework for the analysis and design of shape signatures from the spectral graph wavelet perspective.
(ii) We propose a multiresolution shape signature for deformable 3D shape retrieval using a cubic spline generating kernel. The main attractive properties of the proposed signature may be summarized as follows: It can capture both global and local geometry of shapes in a multiresolution fashion; its parameters can be automatically determined as a trade-off between effectiveness and compactness; it nicely fits the intrinsic spatial partition matching; and it yields the best retrieval accuracy on two standard 3D shape benchmarks compared to spectral signatures within the diffusion geometric framework.

The rest of the paper is organized as follows. In Sect. 2, we provide some background on the LB operator, its discretization and eigenanalysis. Then, we briefly describe the Fourier transform in the graph-theoretic framework. In Sect. 3, we propose a multiresolution shape signature in the spectral graph wavelet framework. Using a cubic spline wavelet generating kernel, we introduce an efficient approach for nonrigid 3D shape retrieval in Sect. 4. Experimental results on two standard 3D shape data sets are provided in Sect. 5. Finally, we conclude in Sect. 6.

2 Spectral geometric signatures

Spectral geometry is concerned with the eigenvalue spectrum of the LB operator on a compact Riemannian manifold, and aims at describing the relationships between such a spectrum and the geometric structure of the manifold.

2.1 Laplace–Beltrami operator

Let $\mathbb{M}$ be a smooth orientable 2-manifold (surface) embedded in $\mathbb{R}^3$. A global parametric representation (embedding) of $\mathbb{M}$ is a smooth vector-valued map $\boldsymbol{x}$ defined from a connected open set (parameterization domain) $U \subset \mathbb{R}^2$ to $\mathbb{M} \subset \mathbb{R}^3$ such that $\boldsymbol{x}(\boldsymbol{u}) = (x^1(\boldsymbol{u}), x^2(\boldsymbol{u}), x^3(\boldsymbol{u}))$, where $\boldsymbol{u} = (u^1, u^2) \in U$.

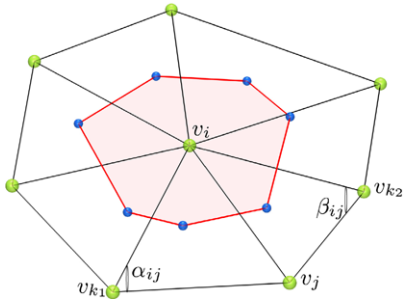


Fig. 1 Voronoi area and cotangent weighted scheme

Given a twice-differentiable function $f : \mathbb{M} \to \mathbb{R}$, the Laplace–Beltrami (LB) operator is defined as

$$\Delta_{\mathbb{M}} f = -\frac{1}{\sqrt{\det g}} \sum_{i,j=1}^{2} \frac{\partial}{\partial u^j}\left(\sqrt{\det g}\, g^{ij} \frac{\partial f}{\partial u^i}\right), \tag{1}$$

where the matrix $g = (g_{ij})$ is the Riemannian metric tensor on $\mathbb{M}$, and g^{ij} denote the elements of the inverse of g. The metric tensor g is an intrinsic quantity in the sense that it relates to measurements inside the surface [23, 24].

2.2 Discrete Laplace–Beltrami operator

In computer graphics, 3D objects are commonly represented as triangle meshes. A triangle mesh $\mathbb{M}$ is usually denoted by $\mathbb{M} = (\mathcal{V}, \mathcal{T})$, where $\mathcal{V} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is the set of vertices and $\mathcal{T} = \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m\}$ is the set of triangles. Two distinct vertices $\boldsymbol{v}_i, \boldsymbol{v}_j \in \mathcal{V}$ are adjacent (denoted by $\boldsymbol{v}_i \sim \boldsymbol{v}_j$ or simply $i \sim j$) if they are connected by an edge.

A signal (function) $f : \mathcal{V} \to \mathbb{R}^n$ defined on the vertices of the mesh may be represented as an n-dimensional vector $\boldsymbol{f} \in \mathbb{R}^n$, where the kth component of the vector $\boldsymbol{f}$ represents the signal value at the kth vertex in $\mathcal{V}$. Using a mixed finite element/finite volume method on triangle meshes [25], the value of $\Delta_{\mathbb{M}} f$ at a vertex $\boldsymbol{v}_i$ (or simply i) can be approximated using the cotangent weight scheme:

$$\Delta_{\mathbb{M}} f(i) \approx \frac{1}{a_i} \sum_{j \sim i} \frac{\cot \alpha_{ij} + \cot \beta_{ij}}{2} \big(f(j) - f(i)\big), \tag{2}$$

where α_{ij} and β_{ij} are the angles $\angle(\boldsymbol{v}_i \boldsymbol{v}_{k_1} \boldsymbol{v}_j)$ and $\angle(\boldsymbol{v}_i \boldsymbol{v}_{k_2} \boldsymbol{v}_j)$ of two faces $\boldsymbol{t}^{\alpha} = \{\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_{k_1}\}$ and $\boldsymbol{t}^{\beta} = \{\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_{k_2}\}$ that are adjacent to the edge $[i, j]$, and a_i is the area of the Voronoi cell (shaded polygon), as shown in Fig. 1. It is worth pointing out that the cotangent weight scheme is numerically consistent and preserves several important properties of the continuous LB operator, including symmetry and positive-definiteness [26].

2.3 Eigenanalysis

The eigenvalues λ_ℓ and the associated eigenfunctions $\boldsymbol{\chi}_\ell$ of the LB operator can be computed by solving the following generalized eigenvalue problem:

$$C\boldsymbol{\chi}_\ell = \lambda_\ell R \boldsymbol{\chi}_\ell, \quad \ell = 1, 2, \ldots, n, \tag{3}$$

where $\boldsymbol{\chi}_\ell$ is the unknown eigenfunction evaluated at n mesh vertices, $R = \mathrm{diag}(a_i)$ is a positive-definite diagonal matrix, and C is a sparse symmetric matrix given by

$$C = \begin{cases} \sum_{i=1}^{n} c_{ij} & \text{if } i = j \\ -c_{ij} & \text{if } i \sim j \\ 0 & \text{o.w.} \end{cases} \tag{4}$$

with

$$c_{ij} = \begin{cases} \frac{\cot \alpha_{ij} + \cot \beta_{ij}}{2} & \text{if } i \sim j \\ 0 & \text{o.w.} \end{cases} \tag{5}$$

Assuming that we have a connected triangle mesh, we may sort the eigenvalues in ascending order as $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n = \lambda_{\max}$ with associated orthonormal eigenfunctions as $\boldsymbol{\chi}_1, \boldsymbol{\chi}_2, \ldots, \boldsymbol{\chi}_n$. The eigensystem $\{\lambda_\ell, \boldsymbol{\chi}_\ell\}_{\ell=1,\ldots,n}$ of the LB operator enjoys nice properties, including isometry invariance.

2.4 Graph Fourier transform

The Fourier transform is a powerful mathematical tool for the analysis of functions, and defines a relationship between a function in the time domain and its representation in the frequency domain. The inverse Fourier transform then reconstructs the original function from its transformed frequency components. The eigensystem $\{\lambda_\ell, \boldsymbol{\chi}_\ell\}$ of the LB operator can be interpreted in the same vein as the Fourier transform basis: the eigenvalues λ_ℓ act as the frequencies, while the eigenfunctions $\boldsymbol{\chi}_\ell$ play the role of Fourier basis functions. For any function $f \in \mathbb{R}^n$ defined on the mesh vertex set $\mathcal{V}$, the forward and inverse graph Fourier transforms are defined, respectively, by

$$\hat{f}(\ell) = \langle \boldsymbol{\chi}_\ell, f \rangle = \sum_{i=1}^{n} \boldsymbol{\chi}_\ell^*(i) f(i), \quad \ell = 1, \ldots, n, \tag{6}$$

and

$$f(j) = \sum_{\ell=1}^{n} \hat{f}(\ell) \boldsymbol{\chi}_\ell(j), \quad j \in \mathcal{V}, \tag{7}$$

where $*$ denotes operation of complex conjugate. It is worth pointing out that since $\boldsymbol{\chi}_\ell$ is a real-valued function on the mesh vertices, it follows that $\boldsymbol{\chi}_\ell^* = \boldsymbol{\chi}_\ell$. However, for the sake of generality, we use the complex conjugate notation throughout the paper.

3 Proposed approach

3.1 Spectral graph wavelet transform

Similarly to the Fourier transform which decomposes a signal into its constituent frequencies, the wavelet transform is a powerful multiresolution analysis tool that enables decomposition of a signal into a wavelet basis which allows simultaneous localization in space and frequency [19]. The idea of wavelets is based on the use of shifting and scaling operations on the signal. Using these two operations, a signal f can be represented as the sum of shifted and scaled versions of the mother wavelet function, ψ, and shifted versions of the scaling function, ϕ. The mother wavelet and scaling functions act as band-pass and low-pass functions, respectively. It is, however, not straightforward how to apply the scaling operation on a signal (or function) defined on the mesh vertices. In other words, for a function $f(i)$ defined on a mesh vertex $i \in \mathcal{V}$, it is meaningless to interpret $f(ti)$ on its own domain for a scaling constant t. To tackle this problem, Hammond et al. [21] introduced the spectral graph wavelet transform (SGWT). The idea is to cast the problem in the Fourier domain and then define the required scaling in that domain. The SGWT is determined by the choice of a spectral graph wavelet generating kernel $g : \mathbb{R}^+ \to \mathbb{R}^+$. To act as a band-pass filter, the kernel g should satisfy $g(0) = 0$ and $\lim_{x\to\infty} g(x) = 0$.

Wavelet function Let g be a given kernel function. The spectral graph wavelet coefficients of a given function f are defined as:

$$W_f(t, j) = \langle \psi_{t,j}, f \rangle = \sum_{\ell=1}^{n} g(t\lambda_\ell)\hat{f}(\ell)\boldsymbol{\chi}_\ell(j), \tag{8}$$

where $\psi_{t,j}$ is the spectral graph wavelet localized at vertex j and scale t, and it is given by

$$\psi_{t,j}(i) = \sum_{\ell=1}^{n} g(t\lambda_\ell)\boldsymbol{\chi}_\ell^*(j)\boldsymbol{\chi}_\ell(i). \tag{9}$$

It should be noted that $g(t\lambda_\ell)$ is able to modulate the spectral wavelets $\psi_{t,j}$ only for λ_ℓ within the domain of the spectrum of LB operator. Thus, an upper bound on the largest eigenvalue $\lambda_{\max}$ is required to provide knowledge on the spectrum in practical applications.

Scaling function Similarly to the low-pass scaling functions in the classical wavelet analysis, a second class of waveforms $h : \mathbb{R}^+ \to \mathbb{R}$ are used as low-pass filters to better encode the low-frequency content of a function f defined on the mesh vertices. To act as a low-pass filter, the function h should satisfy $h(0) > 0$ and $h(x) \to 0$ as $x \to \infty$. Similarly to the wavelet kernels, the scaling function coefficients of f are defined as:

$$S_f(j) = \langle \phi_j, f \rangle = \sum_{\ell=1}^{n} h(\lambda_\ell)\hat{f}(\ell)\boldsymbol{\chi}_\ell(j), \tag{10}$$

where ϕ_j is the scaling function at vertex j:

$$\phi_j(i) = \sum_{\ell=1}^{n} h(\lambda_\ell)\boldsymbol{\chi}_\ell^*(j)\boldsymbol{\chi}_\ell(i). \tag{11}$$

A major advantage of using the scaling function is to ensure that the original signal f can be stably recovered when sampling scale parameter t with a discrete number of values t_k. As demonstrated in [21], given a set of scales $\{t_k\}_{k=1}^K$, the set $F = \{\phi_j\}_{j=1}^n \cup \{\psi_{t_k,j}\}_{k=1}^K{}_{j=1}^n$ forms a spectral graph wavelet frame with bounds

$$A = \min_{\lambda\in[0,\lambda_{\max}]} G(\lambda) \quad \text{and} \quad B = \max_{\lambda\in[0,\lambda_{\max}]} G(\lambda), \tag{12}$$

where

$$G(\lambda) = h(\lambda)^2 + \sum_k g(t_k\lambda)^2. \tag{13}$$

The stable recovery of f is ensured when A and B are away from zero. Additionally, the crux of the scaling function is to smoothly represent the low-frequency content of the signal on the mesh. Thus, the design of the scaling function h is uncoupled from the choice of wavelet generating kernel g.

3.2 Proposed multiresolution shape signature

Wavelets are useful in describing functions at different resolution levels. Understanding and characterizing the differences between functions at different levels of resolution is what wavelets are all about. To characterize the localized context around a mesh vertex $j \in \mathcal{V}$, we assume that the signal on the mesh is a unit impulse function, that is $f(i) = \delta_j(i)$ at each mesh vertex $i \in \mathcal{V}$. Since the graph Fourier transform of δ_j is given by

$$\hat{\delta}_j(\ell) = \langle \boldsymbol{\chi}_\ell, \delta_j \rangle = \sum_{i=1}^{n} \boldsymbol{\chi}_\ell^*(i)\delta_j(i) = \boldsymbol{\chi}_\ell^*(j), \tag{14}$$

it follows that the spectral graph wavelet coefficients are

$$W_{\delta_j}(t, j) = \langle \psi_{t,j}, \delta_j \rangle = \sum_{\ell=1}^{n} g(t\lambda_\ell)\boldsymbol{\chi}_\ell^2(j) \tag{15}$$

and that the coefficients of the scaling function are

$$S_{\delta_j}(j) = \sum_{\ell=1}^{n} h(\lambda_\ell)\boldsymbol{\chi}_\ell^2(j). \tag{16}$$

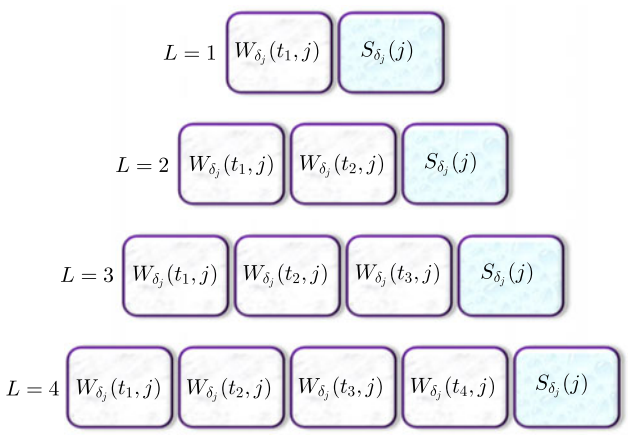


Fig. 2 Pyramid representation of the proposed spectral graph wavelet signature at various resolutions levels $L = 1, \ldots, R$, where $R = 4$

Following the multiresolution analysis, the spectral graph wavelet and scaling function coefficients are collected to form the following multiresolution shape signature:

$$\mathcal{S}_R(j) = \left\{ \boldsymbol{s}_L(j) \mid L = 1, \ldots, R \right\}, \tag{17}$$

which we refer to as the *spectral graph wavelet signature* (SGWS), where R is the resolution parameter, and $\boldsymbol{s}_L(j)$ is the shape signature at resolution level L:

$$\boldsymbol{s}_L(j) = \left\{ W_{\delta_j}(t_k, j) \mid k = 1, \ldots, L \right\} \cup \left\{ S_{\delta_j}(j) \right\}. \tag{18}$$

The wavelet scales t_k ($t_k > t_{k+1}$) are selected to be logarithmically equispaced between maximum and minimum scales t_1 and t_L, respectively. Thus, the resolution level L determines the resolution of scales to modulate the spectrum. The proposed shape signature can be represented as a pyramid, as depicted in Fig. 2, which shows that at resolution level $L = 1$ the signature $\boldsymbol{s}_L(j)$ consists of two elements: one element, $W_{\delta_j}(t_1, j)$, of spectral graph wavelet function coefficients and another element, $S_{\delta_j}(j)$, of scaling function coefficients. And at level $L = 4$, the signature $\boldsymbol{s}_L(j)$ consists of five elements (four elements of spectral graph wavelet function coefficients and one element of scaling function coefficients). Hence, if the resolution is set to $R = 4$, then the multiresolution signature $\mathcal{S}_R(j)$ is composed of a total of 14 elements, as illustrated in Fig. 2. It should be noted that these elements change from one resolution level L to another, albeit with a slight abuse of notation we used the same symbols to denote the spectral graph wavelet function coefficients as well as the scaling function coefficients at each level.

The spectral graph wavelet coefficients, given by Eq. (15), yield a general form for spectral signatures, which includes both HKS and WKS as particular cases, indicating a close relationship between these two signatures and our proposed

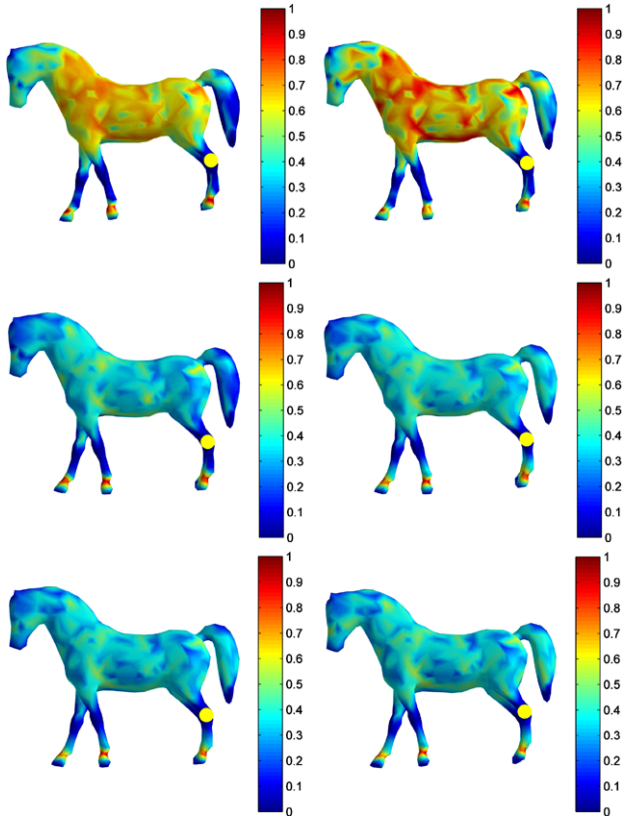


Fig. 3 Normalized χ^2-distance between a reference point (*yellow colored point* on the horse's back left leg) and other surface points using SGWS at different resolution levels. From top to bottom and left to right, the resolution levels are $L = 1, 2, 3, 4, 5$ and 6

SGWS. Moreover, our approach provides a general and flexible framework for the analysis and design of shape signatures from the wavelet viewpoint. Unlike HKS and WKS, our proposed signature allows a multiresolution representation of shapes. The multiresolution analysis using spectral graph wavelets model the shape content in different levels. These levels are used to capture different details inherently found in different structures, as shown in Fig. 3. Given a mesh vertex j, we calculate the dissimilarity between j and a reference vertex j_{ref} (the yellow point on the horse's back left leg in Fig. 3) at resolution level L, using the χ^2-distance given by

$$d(j, j_{\text{ref}}) = \sum_{k=1}^{L+1} \frac{(s_k(j) - s_k(j_{\text{ref}}))^2}{s_k(j) + s_k(j_{\text{ref}})}, \tag{19}$$

where s_k denotes the elements of the signature vector $\boldsymbol{s}_L$ of length $L + 1$. For visualization purposes, we normalize all the χ^2-distances to $[0, 1]$ via division by $\max d(j, j_{\text{ref}})$. As can be seen in Fig. 3, the discriminative power from the reference point to other points on the horse surface is dominated by the resolution level $L = 2$. From $L = 3$ to $L = 6$, more detailed geometry starts to dominate the representa-

tion and it always maintains a large dissimilarity between the reference and feet, resulting in relatively small dissimilarities for other points.

4 Cubic spline wavelet for deformable shape retrieval

In this section, we discuss what a reasonably good descriptor is for shape retrieval [18], and how to design a signature that satisfies the following three properties:

- *Invariance*. The descriptor should be invariant or at least insensitive to a certain class of transformations that the shape may undergo. In this paper, we consider shapes with only isometric or near-isometric transformations. Since the eigensystem of LB operator is intrinsic, the spectral signature is naturally isometry invariant.
- *Efficiency*. The descriptor should capture as much information as possible within as little number of dimensions as possible. Additionally, a wavelet provides very compact support because of its band-pass nature, and thus less redundant information is contained among different supports. To guarantee the localization ability of the spectral graph wavelet in the limit of fine scale, the kernel function g should behave as a monic power of a mesh vertex near the origin [21].
- *Discriminative power*. The descriptor should be able to distinguish between shapes belonging to different classes. First, shapes from different classes usually have different micro structures. We use the multiresolution strategy to capture the micro structures at different resolution levels in a principled fashion. Second, low-frequency information represents the macro structures which are critical to shape comparison. Consequently, the scaling function h is an integral part of our proposed shape signature. The guidance to design a proper h is not only to make up for the information loss by the wavelet kernel g in low frequency, but also not to overlap with g.

As suggested in [21], we choose the cubic spline wavelet and scaling function kernels given by

$$g(x) = \begin{cases} x^2 & \text{if } x < 1 \\ -5 + 11x - 6x^2 + x^3 & \text{if } 1 \le x \le 2 \\ 4x^{-2} & \text{if } x > 2, \end{cases} \tag{20}$$

and

$$h(x) = \gamma \exp\left(-\left(\frac{x}{0.6\lambda_{\min}}\right)^4\right), \tag{21}$$

respectively, where $\lambda_{\min} = \lambda_{\max}/20$, γ is set such that $h(0)$ has the same value as the maximum value of g. The maximum and minimum scales are set to $t_1 = 2/\lambda_{\min}$ and $t_L = 2/\lambda_{\max}$.

The geometry captured at each resolution level L of the SGWS can be viewed as the area under the curve G (see Fig. 4). For a given resolution level L, we can understand the information from a specific range of the spectrum as its associated area under the curve G. As the resolution level L increases, the partition of spectrum becomes tighter, and thus a larger portion of the spectrum is highly weighted.

Relation to HKS The kernel function $g_{\text{HKS}} = \exp(-t\lambda)$ yields HKS. Since $g_{\text{HKS}} \neq 0$, it follows that g_{HKS} does not satisfy the admissibility condition. As can be seen in Fig. 4(a), there is a rich redundance along the spectrum between different scales. However, multiresolution analysis is not possible for HKS. Since g_{HKS} acts as a low-pass filter, HKS will fail to capture micro structures. Consequently, HKS is not able to conduct high-precision feature localization. As can be observed in the top row of Fig. 5, the χ^2-distance changes slowly along the surface. The same distance value remains with the isometric transformation. When a human body (shape from a different class) compares with the reference, large distances occupy most of the area.

Relation to WKS The WKS is obtained using the kernel function $g_{\text{WKS}} = C_t \exp(\frac{-(\log t - \log\lambda)^2}{\sigma^2})$, which is a log-normal distribution function that forms a wavelet generating kernel, where C_{t_k} is a normalization constant and σ^2 is the variance of the distribution [17]. Unlike our proposed shape signature, there is, however, no scaling function for WKS. Therefore, less low-frequency information is used in WKS, as illustrated in Fig. 4(b), resulting in a substantial loss of global geometry of shapes. In the middle row of Fig. 5, small χ^2-distances appear on a large percentage of the surface even when comparing the reference to a body shape. The advantage of WKS is also vividly depicted in the sense that micro structures that are different with the local geometry of the reference are detected.

Proposed signature revisited Observed from the reconstruction perspective, the cubic spline kernel is devised obeying the rules of the wavelet generating kernel function. Therefore, the signal defined on the surface can be stably recovered, whereas neither HKS nor WKS allow stable recovery in the sense that the signal is only recovered approximately. In Figs. 4(c) to (h) display the cubic wavelet spline kernels and their squared sum function G for different values of the resolution parameter R, ranging from 1 to 6. Each value of the function G indicates the energy contribution from each frequency, and it also acts as a good tool to visualize the resolution oscillation for each resolution R. The last row of Fig. 5 shows our proposed signature with $R = 2$. In the next section, our experiments show that

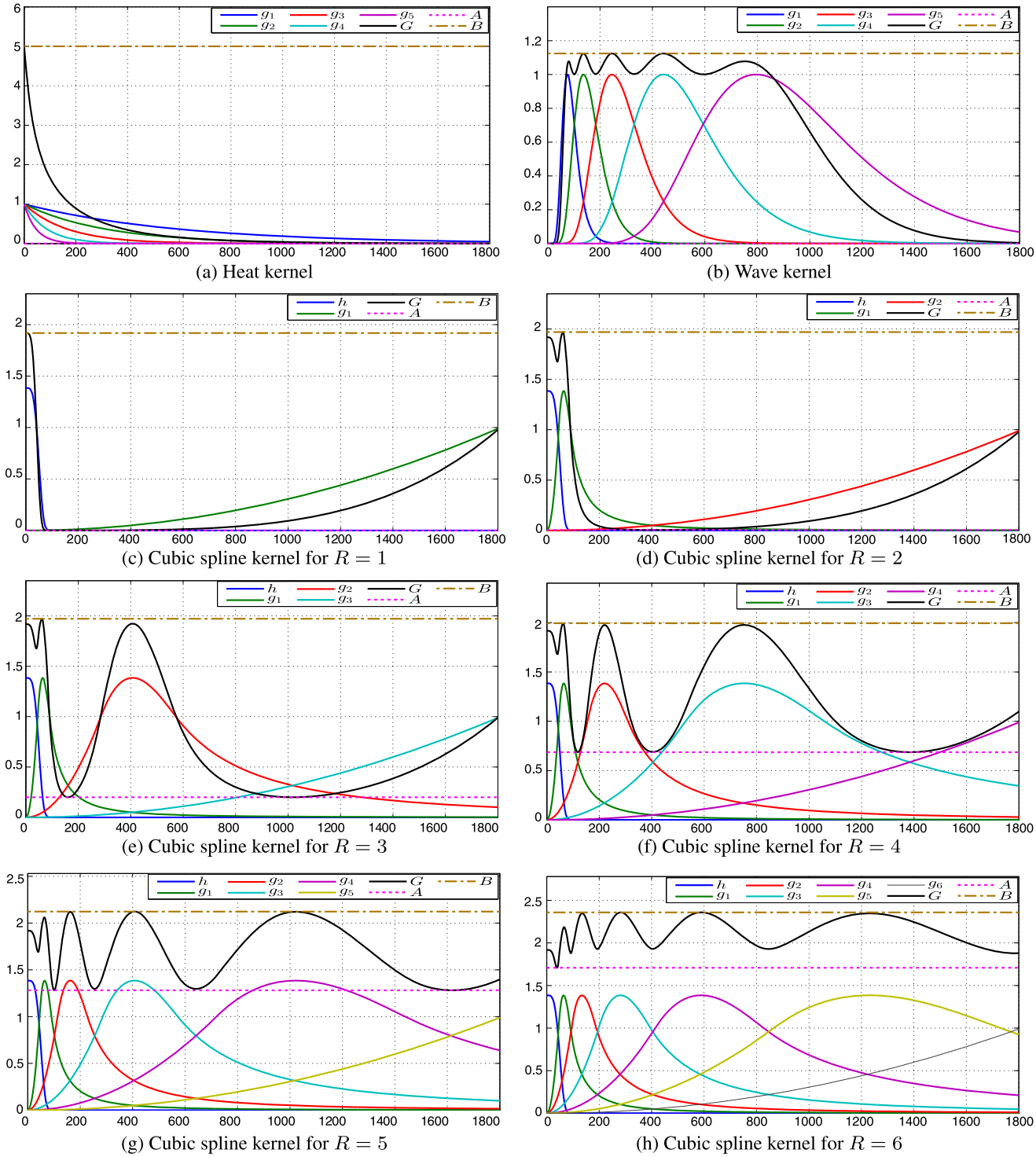


Fig. 4 Spectrum modulation using different kernel functions at various resolutions. The *dark line* is the squared sum function G, while the *dash-dotted* and the *dotted lines* are upper and lower bounds (B and A) of G, respectively

$R = 2$ gives the best results. As can be seen, the SGWS integrates the advantages of both HKS (global geometry for discriminative power) and WKS (local geometry for localization). The former is demonstrated by the body surface colored in red. The χ^2-distances are larger than the horse model. The latter is evidenced by sharply detected features close to the end of the four legs of the horse. As a whole, that is why we observe that the SGWS brings steeper color change within a neighborhood and deeper red on a different shape.

Fig. 5 Normalized χ^2-distance between a reference point (*yellow colored* on the horse's back left leg) and other surface points using different signatures. *Top row*: heat kernel. *Middle row*: wave kernel. *Bottom row*: cubic spline wavelet kernel with $R = 2$

5 Experimental results

To assess the performance of our proposed approach on 3D shape retrieval, we conducted experiments on two standard benchmarks: SHREC 2010 [27] and SHREC 2011 [4]. Instead of adopting the Gaussian kernel for the codeword ambiguity modeling [12], we use the traditional L_0-norm codebook because it can evaluate signatures without the effect of the kernel size. Shape comparison is then performed via intrinsic spatial pyramid matching (ISPM). The key idea of ISPM is to adopt the level sets of the second eigenfunction of LB operator as cuts to perform surface partition. ISPM is largely motivated by spatial pyramid matching [28], which partitions an image into increasingly fine subregions and then computes histograms of local features found inside each subregion.

5.1 Settings

Evaluation measure We evaluated the shape retrieval performance using the Discounted Cumulative Gain (DCG) [29]. The DCG score reflects the performance of the algorithm when correct results that are retrieved earlier are weighted higher than those retrieved later. All normalized DCG calculations are relative values in the interval [0, 1], and higher values imply better retrieval performance.

Comparing signatures We compared the proposed method with classical spectral signatures, including SIHKS, HKS and WKS. The SIHKS is chosen because of its excellent performance in the Shape Google algorithm [12]. The first 150 eigenvalues and eigenvectors of the LB operator on each shape are used. We experimentally select the best parameters for each signature as well as on each SHREC data set as follows. For HKS, we formulate the diffusion time as $t = t_0\alpha^\tau$, where τ is sampled from 0 to a given scale T with a resolution $1/4$. In our case, we set $T = 5$, $t_0 = 0.01$ and $\alpha = 4$. In order to construct the SIHKS, we use $t = \alpha^\tau$, where τ ranges from 1 to a given scale with finer increments of $1/16$, and we set the values $T = 25$ and $\alpha = 2$. After applying the logarithm, derivative and Fourier transform, all the frequencies are used to obtain the best result. For WKS, we seek the best increment by dividing the spectrum interval by a scalar M. The variance σ is also inferred as a certain percentage of the interval. The best parameters are $M = 50$ and percentage equal to 0.2.

Complexity We implemented our algorithms in MATLAB 7.14 (R2012a), and we performed the experiments on a desktop computer with an Intel Core i3-2100 running at 3.1 GHz and 4 GB RAM. We computed offline in advance the vocabulary, which depends on the number of the descriptors (number of mesh vertices), the dimension of the descriptor, and the vocabulary size (number of clusters). To confirm getting optimal results, the clustering is repeated 3 times, and each time by a new set of initial cluster centroid positions. The solution with the lowest value for the sum of distances is returned. We also simplify our mesh to 2000 faces for each shape. For a set of approximately 2×10^5 descriptors, the runtime (in seconds) for a vocabulary with size 32 ranges from 77 s for $R = 1$ to 501 s for $R = 6$. On the other hand, the runtimes for SIHKS, HKS and WKS with their best parameters are 5702, 340 and 725 s, respectively.

5.2 SHREC 2010

The first data set we consider is SHREC 2010 [27], which is a standard data set of nonrigid 3D models used in the Shape Retrieval Contest, organized by National Institute of the Standards and Technology (NIST). The data set consists of 200 shapes spread over 10 categories with 20 shapes each, and ranges from human bodies to man-made tools like glasses. In Fig. 6, we display four models of each class in this data set.

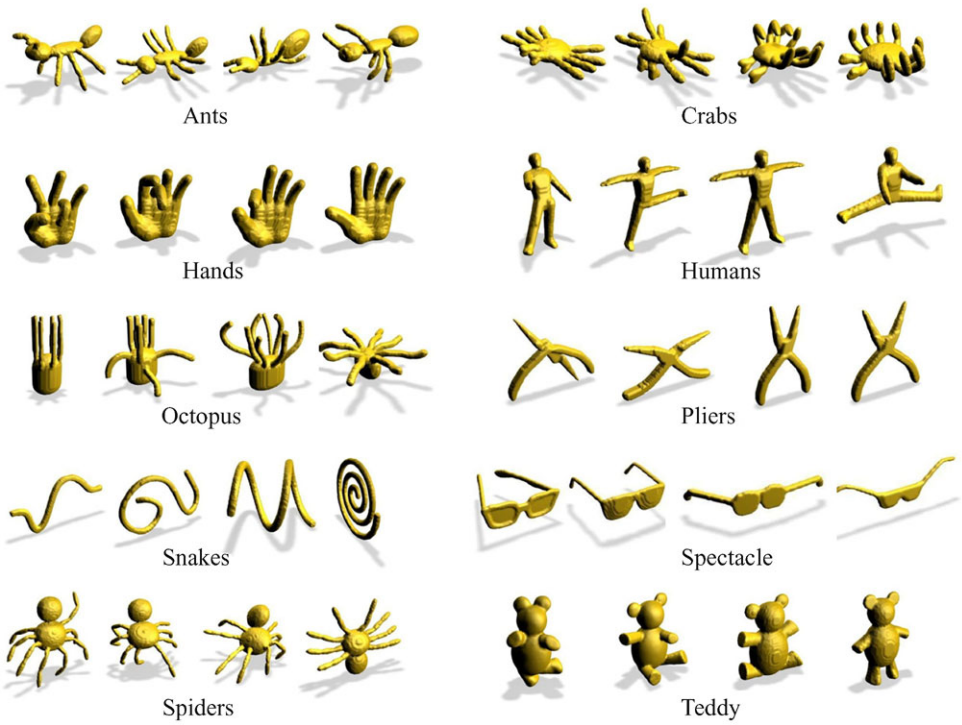


Fig. 6 Sample shapes from SHREC 2010

Table 1 Performance (DCG) of different spectral descriptors on SHREC 2010 with intrinsic spatial pyramid matching

Level l (Partitions)	SIHKS	HKS	WKS	Spectral graph wavelet signature (R)					
				1	2	3	4	5	6
0 (1)	0.8719	0.8448	0.7280	0.8124	**0.8635**	0.8541	0.8454	0.8547	0.8497
1 (2)	0.8734	0.8420	0.7290	0.8172	0.8593	0.8603	0.8498	0.8554	0.8527
2 (4)	0.8793	0.8509	0.7509	0.8789	0.8792	0.8732	0.8681	0.8665	0.8634
3 (8)	0.8817	0.8531	0.7612	0.8781	0.8737	0.8682	0.8651	0.8604	0.8615
4 (16)	0.8841	0.8520	0.7634	0.8687	0.8721	0.8636	0.8640	0.8563	0.8590
Improvement	0.0122	0.0083	0.0354	0.0665	0.0157	0.0191	0.0227	0.0118	0.0137

First, we examine the performance of the SGWS with varying resolutions. In Table 1, the best DCG value of the SGWS matched by bag-of-feature (BoF) is displayed in the row named "Level 0," and it was obtained using a resolution $R = 2$. Interestingly, for $R = 2$ the proposed shape signature has only 5 elements, but it achieves a comparably good result. The element represented by the scaling function captures the global features, whereas the element represented by the spectral graph wavelet function captures more detailed local features and gives evidence that macro structures dominate the deformable shape retrieval. This argument is demonstrated again using $R = 2$, where the best result 0.8635 improves when only the sub-low-frequency band is added. In addition, it can be seen in Table 1 that the accuracy is consistently better than HKS and WKS from $R = 2$ all the way up to $R = 6$. We stop at $R = 6$ for two main reasons: (i) The higher is the resolution, the more expensive is the computational cost. However, the performance remains relatively stable; (ii) In Fig. 4, each value of G indicates the energy contribution of each frequency. The first local minimum close to 0 becomes the global minimum since $R = 6$, indicating the end of low-frequency domination.

Next, we examine the behavior of the SGWS with intrinsic spatial pyramid matching using different pyramid levels l and different surface partitions. The rows named "Level 1" to "Level 4," in Table 1, list the DCG performance achieved at different pyramid levels. The improvement of ISPM over BoF is displayed in the last row of the table, where it can be seen that WKS has improved most significantly by gaining 0.0354, while SIHKS and HKS are improved slightly by 0.0122 and 0.0083, respectively. This can be understood as the potential of signatures, independently occupying the frequency band, is well improved by

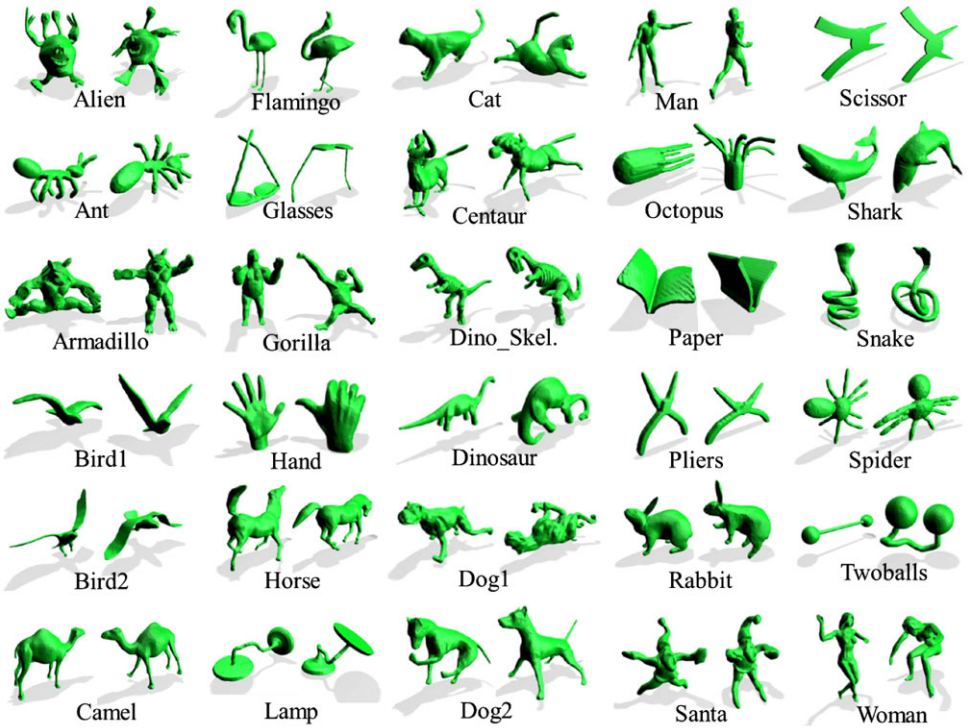


Fig. 7 Sample shapes from SHREC 2011

Table 2 Performance (DCG) of different spectral descriptors on SHREC 2011 with intrinsic spatial pyramid matching

Level l (Partitions)	SIHKS	HKS	WKS	Spectral graph wavelet signature (R)					
				1	2	3	4	5	6
0 (1)	0.8262	0.8114	0.6801	0.8043	**0.8948**	0.8536	0.8731	0.8613	0.8633
1 (2)	0.8436	0.8277	0.7097	0.8419	0.9208	0.8858	0.9002	0.8873	0.8911
4 (16)	0.8671	0.8721	0.7933	0.9203	0.9508	0.9443	**0.9526**	0.9471	0.9517
7 (128)	0.8771	0.8878	0.8042	0.9132	0.9427	0.9366	0.9383	0.9344	0.9359
9 (512)	0.8793	0.8902	0.8029	0.8982	0.9344	0.9241	0.9319	0.9277	0.9272
Improvement	0.0531	0.0788	0.1241	0.1160	0.0560	0.0907	0.0795	0.0858	0.0884

ISPM. However, WKS still gives the worst result because of its low baseline. The problem is well-balanced by the SGWS since the critical low-frequency is also incorporated independently. As expected, the improvement as well as the accuracy of the SGSW are always higher than of the HKS. On the other hand, all the signatures are improved considerably as we go from a pyramid level $l = 1$ to a multi-level setup. The proposed SGWS achieves the best result at $l = 2$, while the other signatures at $l = 3$ or higher.

5.3 SHREC 2011

Our second set of experiments is performed on a larger data set, SHREC 2011 [4], which contains 600 watertight triangle meshes that are equally classified into 30 categories. SHREC 2011 is the most diverse nonrigid 3D shape database available today in terms of object classes and deformations. In Fig. 7 we show two models of each class in the data set.

Table 2 gives a breakdown of shape retrieval accuracy for different resolutions R and at different pyramid levels l. For $l = 0$, the best DCG result of our approach is 0.8948, which is achieved with resolution $R = 2$. This significantly exceeds 0.8262 obtained by SIHKS. The proposed method with other resolutions is very convincing as well, as shown in Table 2. The behavior with ISPM on this database is kind of surprising. The best accuracy of the SGWS is obtained at $l = 4$, and the DCG performance increases to 0.9526, which is much higher than all the signatures in the literature within the framework of diffusion geometry. It can also be observed that the improvement gains of proposed shape signature via ISPM are also higher than SIHKS and HKS. This is in fact consistent with the results on SHREC 2010.

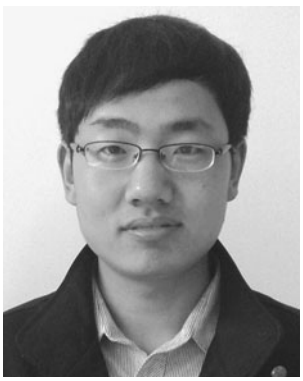
6 Conclusions

In this paper, we introduced a graph wavelet framework for generalized spectral shape signatures. Our proposed approach provides a general and flexible framework to design shape descriptors for specific applications. By concentrating on finding informative spectrum for shape retrieval, we devised a representation that is multiresolution, compact, highly discriminative and parameter-insensitive. We also demonstrated through extensive experiments the effectiveness of our shape signature by achieving state-of-the-art results on two standard benchmarks of 3D shapes.

Acknowledgements This work was supported in part by NSERC Discovery Grant.

References

1. Yang, Y., Lin, H., Zhang, Y.: Content-based 3-D model retrieval: a survey. IEEE Trans. Syst. Man Cybern., Part C, Appl. Rev. **37**(6), 1081–1098 (2007)
2. Tangelder, J., Veltkamp, R.: A survey of content based 3D shape retrieval methods. Multimed. Tools Appl. **39**(3), 441–471 (2008)
3. Shilane, P., Min, P., Kazhdan, M., Funkhouser, T.: The Princeton shape benchmark. In: Proc. Shape Modeling International, pp. 167–178 (2004)
4. Lian, Z., Godil, A., Bustos, B., Daoudi, M., Hermans, J., Kawamura, S., Kurita, Y., Lavoue, G., Nguyen, H., Ohbuchi, R., Ohkita, Y., Ohishi, Y., Porikli, F., Reuter, M., Sipiran, I., Smeets, D., Suetens, P., Tabia, H., Vandermeulen, D.: SHREC'11 track: shape retrieval on non-rigid 3D watertight meshes. In: Proc. Eurographics Symp. 3D Object Retrieval, pp. 79–88 (2011)
5. Li, B., Godil, A., Aono, M., Bai, X., Furuya, T., Li, L., López-Sastre, R., Johan, H., Ohbuchi, R., Redondo-Cabrera, C., Tatsuma, A., Yanagimachi, T., Zhang, S.: SHREC'12 track: generic 3D shape retrieval. In: Proc. Eurographics Conf. 3D Object Retrieval, pp. 119–126 (2012)
6. Lian, Z., Godil, A., Bustos, B., Daoudi, M., Hermans, J., Kawamura, S., Kurita, Y., Lavoué, G., Nguyen, H.V., Ohbuchi, R., Ohkita, Y., Ohishi, Y., Porikli, F., Reuter, M., Sipiran, I., Smeets, D., Suetens, P., Tabia, H., Vandermeulen, D.: A comparison of methods for non-rigid 3D shape retrieval. Pattern Recognit. **46**(1), 449–461 (2013)
7. Belkin, M., Niyogi, P., Sindhwani, V.: Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. J. Mach. Learn. Res. **7**, 2399–2434 (2006)
8. Lévy, B.: Laplace–Beltrami eigenfunctions: towards an algorithm that "understands" geometry. In: Proc. IEEE Int. Conf. Shape Modeling and Applications, p. 13 (2006)
9. Reuter, M., Wolter, F., Peinecke, N.: Laplace–Beltrami spectra as 'Shape-DNA' of surfaces and solids. Comput. Aided Des. **38**(4), 342–366 (2006)
10. Reuter, M.: Hierarchical shape segmentation and registration via topological features of Laplace–Beltrami eigenfunctions. Int. J. Comput. Vis. **89**(2), 287–308 (2010)
11. Rustamov, R.: Laplace–Beltrami eigenfunctions for deformation invariant shape representation. In: Proc. Symposium on Geometry Processing, pp. 225–233 (2007)
12. Bronstein, A., Bronstein, M., Guibas, L., Ovsjanikov, M.: Shape Google: geometric words and expressions for invariant shape retrieval. ACM Trans. Graph. **30**(1), 1–20 (2011)
13. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Proc. Eurographics Sympo. Geometry Processing, pp. 156–164 (2003)
14. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. Comput. Graph. Forum **28**(5), 1383–1392 (2009)
15. Gçbal, K., Bærentzen, J.A., Aanæs, H., Larsen, R.: Shape analysis using the auto diffusion function. Comput. Graph. Forum **28**(5), 1405–1513 (2009)
16. Kokkinos, I., Bronstein, M., Litman, R., Bronstein, A.: Intrinsic shape context descriptors for deformable shapes. In: Proc. Computer Vision and Pattern Recognition, pp. 159–166 (2012)
17. Aubry, M., Schlickewei, U., Cremers, D.: The wave kernel signature: a quantum mechanical approach to shape analysis. In: Proc. Computational Methods for the Innovative Design of Electrical Devices, pp. 1626–1633 (2011)
18. Bronstein, A.: Spectral descriptors for deformable shapes. CoRR abs/1110.5015 (2011)
19. Mallat, S.: A Wavelet Tour of Signal Processing: The Sparse Way. Academic Press, San Diego (2008)
20. Coifman, R., Lafon, S.: Diffusion maps. Appl. Comput. Harmon. Anal. **21**(1), 5–30 (2006)
21. Hammond, D., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Appl. Comput. Harmon. Anal. **30**(2), 129–150 (2011)
22. Kim, W., Pachauri, D., Hatt, C., Chung, M., Johnson, S., Singh, V.: Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination. In: Proc. NIPS, pp. 1135–1143 (2012)
23. Rosenberg, S.: The Laplacian on a Riemannian Manifold. Cambridge University Press, Cambridge (1997)
24. Bronstein, A., Bronstein, M., Kimmel, R.: Numerical Geometry of Non-Rigid Shapes. Springer, Berlin (2008)
25. Meyer, M., Desbrun, M., Schröder, P., Barr, A.: Discrete differential-geometry operators for triangulated 2-manifolds. Vis. Math. III **3**(7), 35–57 (2003)
26. Wardetzky, M., Mathur, S., Kälberer, F., Grinspun, E.: Discrete Laplace operators: no free lunch. In: Proc. Eurographics Sympo. Geometry Processing, pp. 33–37 (2007)
27. Lian, Z., Godil, A., Fabry, T., Furuya, T., Hermans, J., Ohbuchi, R., Shu, C., Smeets, D., Suetens, P., Vandermeulen, D., Wuhrer, S.: SHREC'10 track: non-rigid 3D shape retrieval. In: Proc. Eurographics Sympo. 3D Object Retrieval, pp. 101–108 (2010)
28. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: Proc. CVPR, vol. 2, pp. 2169–2178 (2006)
29. Järvelin, K., Kekäläinen, J.: IR evaluation methods for retrieving highly relevant documents. In: Proc. SIGIR, pp. 41–48 (2000)



Chunyuan Li is a Master's student in Concordia University, Montreal, Canada. He received his B.Sc. degree from the Department of Electronics and Information Engineering at Huazhong University of Science and Technology (HUST), Wuhan, China, in 2011. His research interests include computer vision, computer graphics and pattern recognition.