Spherical Wavelet Descriptors for Content-based 3D Model Retrieval

Hamid Laga Hiroki Takahashi Masayuki Nakajima
Computer Science Department
Tokyo Institute of Technology, Japan
email:{hamid,rocky,nakajima}@img.cs.titech.ac.jp

Abstract

The description of 3D shapes with features that possess descriptive power and invariant under similarity transformations is one of the most challenging issues in content based 3D model retrieval. Spherical harmonics-based descriptors have been proposed for obtaining rotation invariant representations. However, spherical harmonic analysis is based on latitude-longitude parameterization of a sphere which has singularities at each pole. Consequently, features near the two poles are over represented while features at the equator are under-sampled, and variations of the north pole affects significantly the shape function. In this paper we discuss these issues and propose the usage of spherical wavelet transform as a tool for the analysis of 3D shapes represented by functions on the unit sphere. We introduce three new descriptors extracted from the wavelet coefficients, namely: (1) a subset of the spherical wavelet coefficients, (2) the L_1 and, (3) the L_2 energies of the spherical wavelet sub-bands. The advantage of this tool is three fold; First, it takes into account feature localization and local orientations. Second, the energies of the wavelet transform are rotation invariant. Third, shape features are uniformly represented which makes the descriptors more efficient. Spherical wavelet descriptors are natural extension of 3D Zernike moments and spherical harmonics. We evaluate, on the Princeton Shape Benchmark, the proposed descriptors regarding computational aspects and shape retrieval performance.

Keywords. *3D model retrieval, shape matching, spherical wavelets, rotation invariant representation.*

1 Introduction

The 21^{st} century is the era of digital media and a substantial progress has been achieved in acquisition, storage and transmission of different types of information. While text, images, sound and video have been the predominant form of digital media, 3D models emerge as a new form. They have applications in many fields including CAD, medicine, physical simulation, e-commerce and education. Consequently, a significant research effort is spent to developing effective techniques for content-based 3D model retrieval.

A challenging issue in content-based 3D model retrieval is the description of shapes with suitable numerical representations called *shape descriptors*. In general a shape descriptor should be discriminative, compact, easy to compute, and invariant under a group of transformations. While invariance to global transformations, such as translation, rotation, scale [15, 11, 23], and invariance under certain deformations [6, 20], have been extensively studied, existing shape descriptors have difficulties to handle local variations in scale and orientations. Moreover, they should capture only the key shape features.

In this paper we present a new 3D content-based retrieval method relying on spherical wavelet transform (SWT) of the shape function. Spherical Wavelets have been introduced by Schröder et. al. [18] and since, they have been used to solve many geometry processing problems including 3D model compression [7]. Similar to first generation wavelets [2], SWT is an effective tool to analyze shape functions defined on the sphere as they provide a natural partition of the function spectrum into multiscale and oriented sub-bands. SWT is a natural extension of spherical harmonics [5] and 3D Zernike moments [12, 13]. It offers better feature localization and takes all the advantages of wavelets over Fourier analysis.

1.1 Related work

Most of three-dimensional shape retrieval techniques proposed in the literature aim to extract from the 3D model meaningful descriptors based on the geometric and topological characteristics of the object. Survey papers to the related literature have been provided by Tangelder et. al [21] and Iyer et. al [8]. They fall into three broad categories; feature-based including global and local features, graph-based and view-based similarity.

View-based techniques compare 3D objects by comparing their two dimensional projections. The Lightfields [1] are reported to be the most effective descriptor [19]. View-based techniques are suitable for implementing query interfaces using sketches [8, 5].

Graph-based techniques reduce the problem of shape dissimilarity estimation to the problem of comparing graphs. They are suitable for retrieving articulated objects, and the Reeb graphes proposed by Hilaga *et al.*[6], and skeletons [20] are among the most popular. Other techniques include methods based on the distribution of features such as shape distributions [14], and local features such as spin images [9]. Shilane et. al. [19] provides a comparison of these techniques and reported that histogram based techniques are the less efficient in terms of discriminative power.

Feature-based methods aim to extract compact descriptors from the 3D object. A popular approach is to represent the shape using functions defined on the unit sphere. Funkhouser et al.[5] uses spherical harmonics (SH) to analyze the shape function. They demonstrated later that spherical harmonics can be used to achieve rotation invariance provided that the shape function is defined on the sphere [11]. Novotni et. al. [12] uses 3D Zernike moments (ZD) as a natural extension of SH. Representing 3D shapes as functions on concentric spheres has been extensively used. Our developed descriptors fall into this category and are a natural extension of SH and ZD.

The issue of extracting invariant shape features is always an important problem in content-based 3D model retrieval. While translation and scale invariance can be easily achieved [17, 5, 11], rotation invariance is still a challenging issue. Recently, much research has been focused on this issue and various methods have been proposed to cope with the problem. Some of them require pose normalization, where each shape is placed into a canonical coordinate frame. These methods are usually variant of PCA [10, 25], continuous PCA [25], and other variants for solving for axial ambiguity. However, PCA-based alignment is known to misbehave and therefore, it hampers significantly the retrieval performance [11].

Other methods, commonly referred as *invariant* methods, describe shapes in a transformation invariant manner by discarding alignment-dependent shape information. Spherical harmonics [24] and power spectrum-based [23, 11] are among the most popular methods. These approaches rely on the sampling of the shape function in the latitude and longitude directions. However, the problem that we observed in these approaches is that, the sampling is not rotation invariant, and therefore, the shape descriptor will vary by varying the sampling directions. This problem will be further discussed in Section 2.

1.2 Overview and contributions

This paper investigates for the first time the application of spherical wavelet analysis to content-based 3D model retrieval. To the best of our knowledge, SWT have not been applied to content-based retrieval of 3D models so far. We make use of them and propose three new descriptors, namely: spherical wavelet coefficients as feature vector (SWC_d), L1 energy of the spherical wavelet coefficients (SWEL1), and L2 energy of the spherical wavelet coefficients (SWEL2). This paper makes the following contributions:

1. We address for the first time the problem of rotation invariant sampling of the shape function. We found that the sensitivity of the latitude-longitude parameterization to rotations of the north pole affects the rotation invariance of the shape descriptors. This paper proposes a new parameterization method based on regular octahedron sampling.

2. We propose new spherical wavelet-based shape descriptors. The SWC_d takes into account the localization and local orientations of the shape features, while the SWEL1 and SWEL2 are compact and rotation invariant.

3. We evaluate and compare the performance of the proposed descriptors using standard benchmarks.

In the next section we discuss the problem related to shape function sampling and motivate the use of spherical wavelet analysis. Section 3 reviews the general concepts of spherical wavelet transform of functions on the sphere, and how we use them for 3D shape analysis. Section 4 describes in detail the new shape signatures and the similarity estimation method. Section 5 presents some experimental results. Finally, we summarize in Section 6 the main findings of this paper and issues for future research.

2 Rotation invariant shape description

It is popular to represent a 3D shape with functions defined on the unit sphere, sampled on a regular grid of size $n \times n$ of angles of elevation $\theta, 0 \leq \theta \leq \pi$, and azimuth $\phi, 0 \leq \phi \leq 2\pi$. For simplicity, we consider the Spherical Extent Function (EXT) [17], but our analysis applies for any spherical function. Spherical Extent Function $f(\theta, \phi)$ represents the extension of the shape in the radial direction (θ, ϕ).

The steps commonly used to compare 3D shapes are:

1. **Normalization.** Transform the center of mass of the object to the origin, and scale the object to lie within a unit ball.

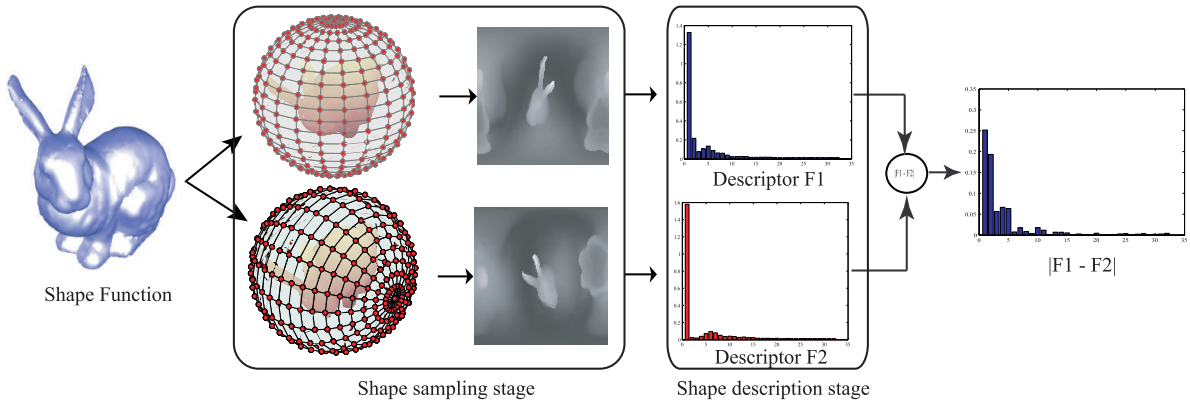


Figure 1: Problem illustration: latitude-longitude parameterization generates shape functions with singularities at each pole affecting the rotation invariance of the shape descriptor.

2. **Parameterization.** Compute the shape function at discrete locations $P = \{p_1, \ldots, p_k\}$ sampled from the shape. These points are obtained by casting rays in different radial directions (θ, ϕ).

3. **Spherical harmonic transform (SHT).** The shape function is expressed in terms of its frequency components.

4. **Shape descriptors.** Feature vectors are extracted and used as a mean for shape comparison.

We refer to step 2 as *sampling stage*, and steps 3 and 4 as *shape description stage*. Step 3 expresses the shape function in terms of its spherical harmonics:

$$f(\theta, \phi) = \sum_{l \geq 0} \sum_{|m| \leq l} f_{l,m} Y_l^m(\theta, \phi) \qquad (1)$$

The vector of spherical harmonics $Y_l^m, |m| \leq l$ forms a basis for the irreducible subspace V^l which is also invariant under the rotation group. Therefore, the norms of the harmonic coefficients:

$$f \rightarrow \{\|f_{l,m}\|\}_{|m| \leq l, l \geq 0} \qquad (2)$$

form a descriptor that is invariant to rotation about the north pole, and the power spectrum:

$$f \rightarrow \{\|f_l\|\}_{l \geq 0} = \left\{ \sqrt{\sum_{|m| \leq l} |f_{l,m}|^2} \right\}_{l \geq 0} \qquad (3)$$

forms a descriptor that is invariant to all rotations [11].

The key observation is that the rotation invariance concerns only the shape description stage, i.e, the shape descriptor is invariant to rotations applied to the input of the shape description stage. In this paper, we question the rotation invariance of the sampling stage.

2.1 Irregular sampling problem

At the sampling stage, the shape function f is sampled into an $n \times n$ regular grid along the elevation and azimuthal angles. Figure 1 shows the case of a unit sphere sampled into 32×32, the Bunny's power spectrum descriptors[1] computed using different poses (rotation of 90 degrees around the X axis), and the L_2 distance between frequency components of the two descriptors.

Notice that: (1) the shape function obtained with latitude-longitude sampling procedure has singularities at each pole where the shape is over-sampled. The areas near the equator are, however, under-sampled. Consequently, small variations of the shape near the two poles will affect significantly the descriptor. (2) the sampling is regular in the spherical coordinate frame, but not in the Euclidean space however. Increasing the sampling rate will alleviate the first problem but at the cost of higher computation time. The over and under sampling problems, however, still remain. (3) rotating the north pole around one of the other axis will result in a different sample of points, therefore a different discrete shape function.

Consequently, while in the continuous case, the power spectrum based descriptors are rotation invariant, in the discrete case however, this property does not hold. In this paper, we propose an alternative solution using a uniform sampling of the unit sphere and spherical wavelet analysis to address these issues.

2.2 Rotation invariant sampling

The two points outlined in the previous section motivate our choice for rotation invariant and uniform shape sam-

[1] For illustration purpose we used 32×32 grids but the descriptors are computed using 128×128 grids. In the literature, grids of 64×64 are the most popular.

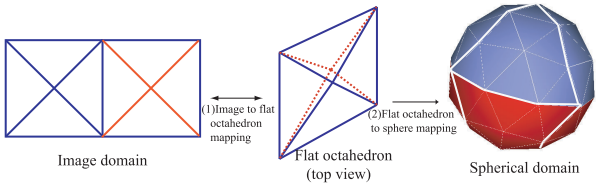


Figure 2: Flat octahedron parameterization procedure. The flat octahedron is isometrically unfolded on the image plane. The left half of the image plane is mapped to the top half of the flat octahedron, which is then mapped to the north hemisphere of the geodesic sphere.

pling stage. The key idea of our approach is that rotation invariant sampling can be achieved using an operator Φ that samples the shape uniformly (in the Euclidean distance sense) in all directions.

To achieve this in practice, we investigated two approaches originally proposed for spherical parameterization and geometry image compression [7, 16]:

1. **Geodesic sphere.** we sample the shape function by casting rays from the shape's center of mass to the vertices of a geodesic sphere. The vertices of the geodesic sphere are equidistant and its faces have equal area. These two conditions are sufficient to guarantee a uniform sampling. The coarsest (level-0) representation of the shape function is obtained using a basic octahedron of 20 vertices. Finer levels are obtained by recursive subdivisions.

2. **Flat octahedron parameterization.** Hoppe et. al [7, 16] maps the sphere onto a square domain using spherical parameterization of a *flattened octahedron* domain. The interesting property is that the flattened octahedron unfolds isometrically onto a square image. Therefore, image processing tools can be used with simple boundary extension rules.

The benefits of these two approaches are two fold: (1) the shape is sampled uniformly in all directions. This eliminates the singularities that appear at each pole in the latitude-longitude parameterization, and (2) the discrete shape function is invariant to rotations along the edges of the spherical triangles (by moving a vertex to its neighbor). We make use of these properties to build efficient shape descriptors.

2.3 Geometry image

In our implementation, we have chosen to use the flat octahedron parameterization (We will justify this choice in Section 3.2). Therefore, we represent each 3D model O in the database with a geometry image I of size $k = w \times h$. The parameterization process performs in three steps:

1. **Image - flat octahedron mapping.** Figure 2 shows how the flat octahedron is unfolded and mapped to different regions of the image I. We use barycentric coordinates mapping to map each pixel of I into the octahedron domain.

2. **Flat octahedron - sphere mapping.** we achieve this by simple spherical projection. This step generates a set of points $V = \{v_1, \ldots, v_k\}$ on the sphere. The mapping function associates each point v_i to a pixel location in the image I.

3. **Shape function.** we redefine the spherical extent function (EXT) f by: $f = \{f_i\}_{i=1}^{k}$, where f_i is the extent of the shape in direction v_i. f has the domain I as a regular support.

The set of points V, the result of step 1 and 2, can be calculated at once and used for all models. Now to describe the shape, a subset of the vector f can be used as shape descriptor [25], but it is well established that the l_p metric is not effective in the spatial domain. On the other hand, spherical harmonics can not be engaged as the sampling is not uniform in terms of azimuthal and elevation angles.

3 Spherical wavelets for 3D shape description

We now consider the problem of descriptor extraction from the spherical shape function. In this paper, we make use of wavelets [18, 7] to efficiently extract shape descriptors. In the following subsections we will review the general concepts, and then present how we use them to analyze the shape function.

3.1 Spherical wavelets

Wavelets are basis functions which represent a given signal at multiple levels of detail, called *resolutions*. They are suitable for sparse approximations of functions. In the Euclidean space, wavelets are defined by translating and dilating one function called *mother wavelet*. In S^2, however, the metric is no longer Euclidean. Schröder et. al. [18] introduced the second generation wavelets. The idea behind was to build wavelets with all desirable properties adapted to much more general settings than real lines and 2D images.

The general wavelet transform of a function λ is constructed as follows:

Analysis: (forward transform)

$$\lambda_{j,k} = \sum_{l \in K(j)} \tilde{h}_{j,k,l} \lambda_{j+1,l}$$
$$\gamma_{j,k} = \sum_{l \in M(j)} \tilde{g}_{j,m,l} \lambda_{j+1,l}$$

Synthesis: (backward transform)

$$\lambda_{j+1,l} = \sum_{k \in K(j)} h_{j,k,l} \lambda_{j,k} + \sum_{m \in M(j)} g_{j,m,l} \gamma_{j,m}$$

where $\lambda_{j,\bullet}$ and $\gamma_{j,\bullet}$ are respectively the approximation and the wavelet coefficients of the function at resolution j. The decomposition filters $\tilde{h}, \tilde{g}$, and the synthesis filters h, g denote spherical wavelet basis functions. The forward transform is performed recursively starting from the shape function $\lambda = \lambda_{n,\bullet}$ at the finest resolution n, to get $\lambda_{j,\bullet}$ and $\gamma_{j,\bullet}$ at level $j, j = n-1, \ldots, 0$. The coarsest approximation $\lambda_{n-i,\bullet}$ is obtained after i iterations ($0 < i \leq n$). The sets $M(j)$ and $K(j)$ are now index sets on the sphere such that $K(j) \cup M(j) = K(j+1)$, and $K(n) = K$ is the index set at the finest resolution.

3.2 Analysis of the spherical shape function

To analyze a 3D model, we first apply spherical wavelet transform (SWT) to the spherical shape function and collect the coefficients to construct discriminative descriptors. The properties and behavior of the shape descriptors are therefore determined by the spherical wavelet basis functions used for transformation.

Similar to 3D Zernike moments [12] and spherical harmonics [11, 23], the desired properties of a descriptor are: (1) Invariance to a group of transformations, (2) Orthonormality of the decomposition, and (3) Completeness of the representation. The orthonormality ensures that the set of features will not contain redundant information. The completeness property implies that we are able to reconstruct approximations of the signal from the decomposition.

The SW basis function should reflect these properties. In our work we have experimented with the second generation wavelets [18] including the linear and butterfly spherical wavelets with lifting scheme, and image wavelets with spherical boundary extension rules [7]. In our experiments on the Princeton Shape Benchmark, we found that the performance of both the linear and butterfly spherical wavelets is very low (comparable to shape distribution based descriptors). Therefore, we decided to use the image based wavelet with spherical boundary extension rules to build our shape descriptors.

The image wavelet transform uses separable filters, so at each step it produces an approximation image A and three detail images HL, LH, and HH. The forward transformation algorithm, illustrated in Figure 3, performs as follows:

1. **Initialization:**
   (a) Generate the geometry image I (therefore the function f) of size $w \times h = 2^{n+1} \times 2^n$ as explained in Section 2.3.
   (b) $A^{(n)} \leftarrow f, l \leftarrow n$.
2. **Forward transform:** repeat the following steps until $l = 0$:
   (a) Apply the forward spherical wavelet transform on $A^{(l)}$, we get the approximation $A^{(l-1)}$, and the detail coefficients $C^{(l-1)} = \{LH^{l-1}, HL^{l-1}, HH^{l-1}\}$ of size $2^l \times 2^{l-1}$.
   (b) $l \leftarrow l - 1$.
3. **Collect the coefficients:** the approximation $A^{(0)}$ and the coefficients $C^{(0)}, \ldots, C^{(n-1)}$ are collected into a vector F.

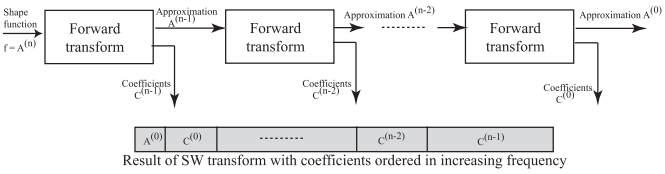


Figure 3: Spherical wavelet-based shape descriptors computation.

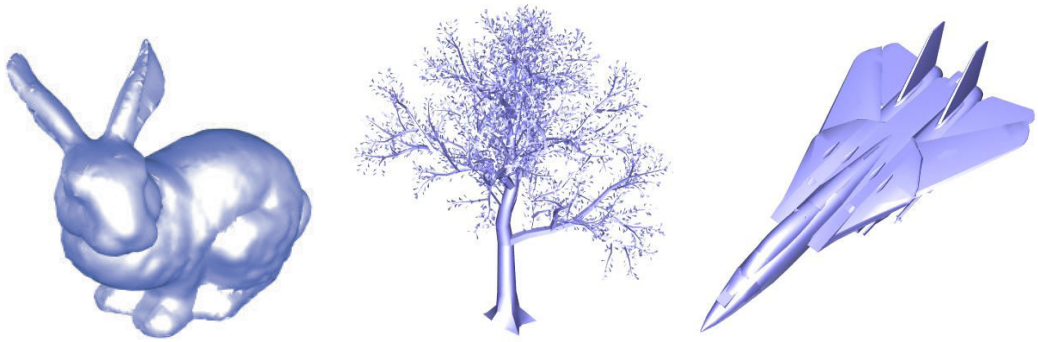
In this paper, we experimented with the Haar wavelets, where the scaling function is designed to take the rolling average of the data, and the wavelet function is designed to take the difference between every two samples in the signal. Other wavelet basis can also be used but requires a further investigation.

4 Spherical wavelet-based descriptors

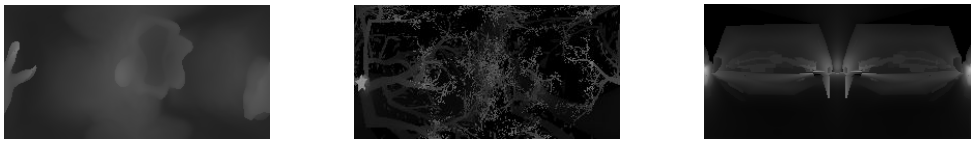
We now consider the computation of shape descriptors. We propose three methods to compare 3D shapes using their spherical wavelet transform: (1) Wavelet coefficients as a shape descriptor (SWC_d) where the shape signature is built by considering directly the spherical wavelet coefficients, and (2) spherical wavelet energies: SWEL1 that uses the $L1$ energy, and (3) SWEL2 using the $L2$ energy of the wavelet sub-bands. Figure 4 shows three models and their SW descriptors. The following sections detail each method.

4.1 Wavelet coefficients as shape descriptor

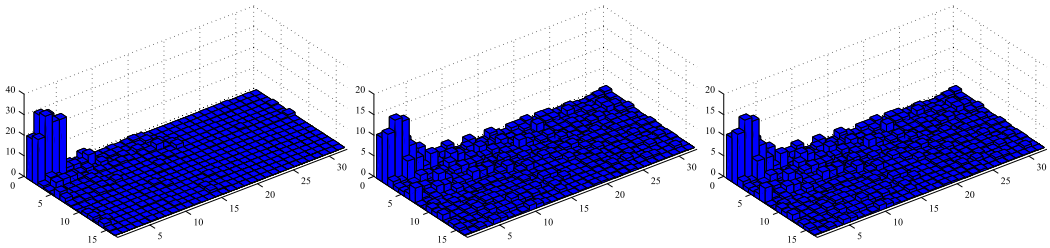
Once the spherical wavelet transform is performed, one may use the wavelet coefficients as shape descriptor. Using the entire coefficients is computationally expensive. Instead, we have chosen to keep the coefficients up to level d. We call the obtained shape descriptor SWC_d, where $d = 0, \ldots, n-1$. In our implementation we used $d = 3$,



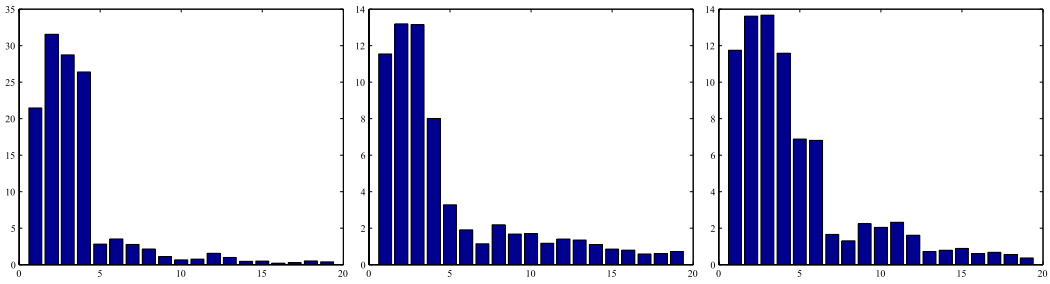
(a) 3D shapes



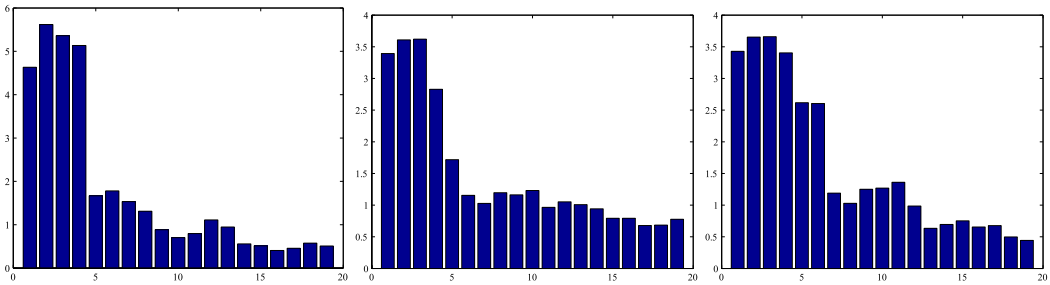
(b) Their associated geometry images



(c) Spherical wavelet coefficients as descriptor (SWC_3).



(d) L_2 energy descriptor (SWEL2).



(e) L_1 energy descriptor (SWEL1).

Figure 4: Example of different models with their spherical wavelet-based descriptors.

therefore we obtain two dimensional feature vectors F of size $N = 2^{d+2} \times 2^{d+1} = 32 \times 16$.

Comparing directly wavelet coefficients requires efficient alignment of the 3D model prior to wavelet transform. A popular method for finding the reference coordinate frame is pose normalization using Principal Component Analysis (PCA) [15], and continuous PCA [25]. We perform the pose normalization in three steps;

1. First we translate the shape's center of mass to the origin $(0, 0, 0)$.
2. Then we align the shape to its principal axis using continuous PCA [25]. We use the maximum area technique to resolve for the positive and negative directions of the principal axis.
3. Finally we scale the shape such that the average distance between the center of mass to any point in the surface is equal to $1/2$.

Figure 4c shows the SWC_3 descriptor extracted on three different models. Notice that, the vector F provides an embedded multi-resolution representation for 3D shape features. This approach performs as a filtering of the 3D shape by removing outliers. A major difference with spherical harmonics is that SWT preserves the localization and orientation of local features. However, a feature space of dimension 512 is still computationally expensive.

4.2 Spherical wavelet energy signatures

The wavelet energy signatures have been proven to be very powerful for texture characterization [22]. Commonly the $L2$ and $L1$ norms are used as measures [3, 4]:

$$F_l^{(2)} = \left(\frac{1}{k_l} \sum_{j=1}^{k_l} x_{l,j}^2 \right)^{\frac{1}{2}} \quad (4)$$

$$F_l^{(1)} = \frac{1}{k_l} \sum_{j=1}^{k_l} \|x_{l,j}\| \quad (5)$$

where $x_{l,j}, j = 1 \dots k_l$ are the wavelet coefficients of the l^{th} wavelet sub-band. Using the observation that rotating a spherical function does not change its energy, we propose to use it to build general rotation invariant shape descriptors. For this purpose we perform $n-1$ decompositions, then we compute the energy of the approximation $A^{(1)}$ and the energy of each detail sub-band $HV^{(l)}, VH^{(l)}$ and $HH^{(l)}$ yielding into a one-dimensional shape descriptor $F = \{F_l\}, l = 0 \dots 3 \times (n-1)$ of size $N = 3 \times (n-1)+1$. In our case we use $n = 7$, therefore $N = 19$. We refer to $L1$ energy and $L2$ energy-based descriptors by **SWEL1** and **SWEL2** respectively.

The main benefits of this descriptor are its compactness, and it is rotation invariant. Therefore, the storage and computation time required for comparison are reduced. Since the sampling stage is also rotation invariant, we obtain shape descriptors that are invariant to general rotations. However, similar to the power spectrum [11], information such as feature localization are lost in the energy spectrum.

4.3 Similarity metric

Since 3D shapes are now represented using N-dimensional vectors with real-valued components (SWC_d is two-dimensional while SWEL1 and SWEL2 are one dimensional feature vectors), a natural way to compute distances in the feature space is to use a vector norm, called also L_p norm. In our implementation we experimented with the L_2 distance. If F_1 and F_2 are the feature vectors of a database object O_1 and a query object O_2 respectively, of dimension N, then the dissimilarity between O_1 and O_2 is the L_2 distance between their descriptors:

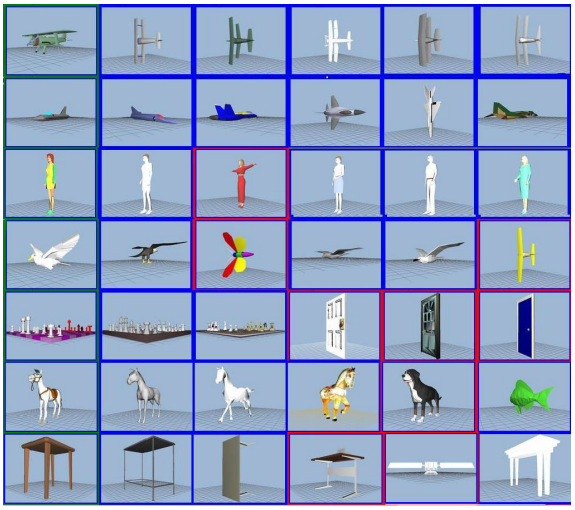
$$D(F_1, F_2) = \left(\sum_{i=1}^{N} \left(F_1(i) - F_2(i) \right)^2 \right)^{1/2} \quad (6)$$

Notice that the proposed spherical wavelet analysis framework supports retrieval at different acuity levels. In some situations, only the main structures of the shapes are required for comparison, while in others, fine details are essential. In the former case, shape matching can be performed by considering only the wavelet coefficients at large scales, while in the later, coefficients at small scales are used. Hence the flexibility of the developed method benefits different retrieval requirements.

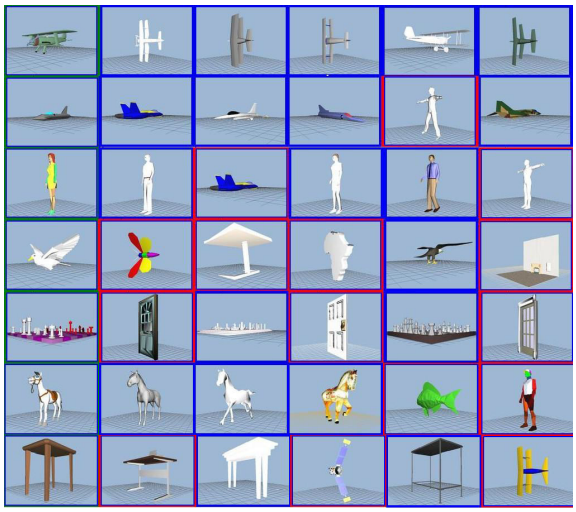
Finally, Table 1 summarizes the length of the proposed descriptors. The SWEL1 and SWEL2 are more efficient in terms of storage requirement and comparison time. They are also rotation invariant. Their discrimination efficiency will be discussed in the following section.

5 Experimental results

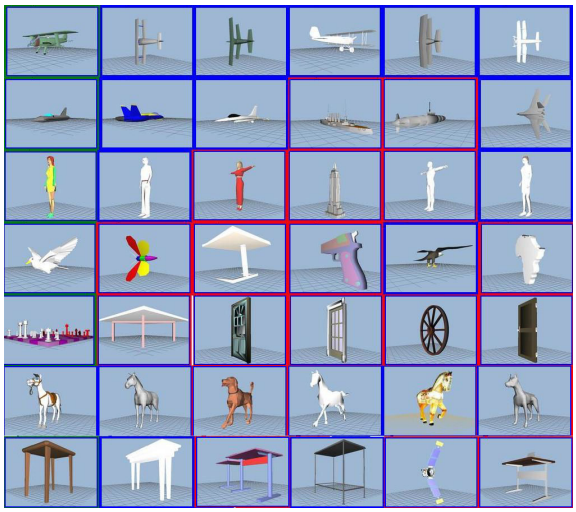
We have implemented the algorithms described in this paper and evaluated their performance on the Princeton Shape Benchmark (PSB)[19]. We represent each model using Spherical Extent Function (EXT). At the early stage of this research, we have experimented with linear and butterfly spherical wavelets using six decomposition levels ($n = 6$). We found, however, that the performance of the descriptors is very low. Instead, we used image wavelets with boundary extension rules. SWC_d requires pose normalization while SWEL1 and SWEL2 are rotation invariant. For the SWC_d descriptor, we use $d = 3$, therefore, we keep the first 512 coefficients.



(a) Retrieval results using spherical wavelet coefficients as shape descriptor (SWC$_d$).



(b) Retrieval results using spherical wavelet L1-energy as shape descriptor (SWEL1).



(c) Retrieval results using spherical wavelet L2-energy as shape descriptor (SWEL2).

Figure 5: Retrieval results using spherical wavelet-based shape descriptors. The query model is plotted in green, the correctly retrieved shapes are highlighted in blue, and the others in red.

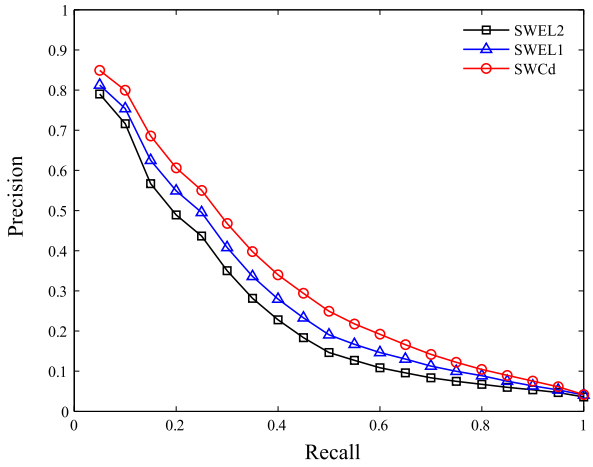


Figure 6: Precision-recall curves for SW-based descriptors.

5.1 Retrieval results

First we executed series of shape matching experiments on the base test classification of the PSB. We select randomly a 3D polygon soup model, and then compare it to the objects in the database. We show in Figure 5 the results of several queries for each of our three descriptors SWC$_d$, SWEL1 and SWEL2. Five most similar objects are displayed. A retrieved model is considered relevant if it belongs to the same class as the query model.

By visually inspecting these results, we noticed that SWC$_d$ descriptor performs better than the others. The $L1$ energy of the spherical wavelet coefficients comes in the second rank.

5.2 Performance evaluation

The precision-recall curves on the base test classifications of the PSB, and using the three shape signatures are shown in Figure 6. We refer the reader to the Princeton Shape Benchmark paper [19] for comparison with other descriptors on the precision-recall measure.

We evaluated the performance of our descriptors using the nearest neighbor, first and second-tier, E-measure and Discount Cumulative Gain measures [19]. The results are summarized in Table 1. We made all the experiments on the base test classification of the PSB. This comes to confirm the visual evaluation, that is, spherical wavelet coefficients perform better, while the L1 and L2 energy comes in the second and third rank respectively. Note that the SWC$_d$ requires more storage and comparison time.

Shilane et. al. [19] summarized the performance on the PSB of several shape descriptors and we use their results to compare with our descriptors. In this paper, we show the performance of four descriptors, but we refer the reader

Table 1: Performance of SW descriptors on the PSB base test classification, values of the length are in bytes, other are in (%). The length refers to the dimension of the feature space.

	length	NN	$1^{st}-$ tier	$2^{nd}-$ tier	E-measure	DCG
SWC_d.	512	46.9	31.4	39.7	20.5	65.4
SWEL1	19	37.3	27.6	35.9	18.6	62.6
SWEL2	19	30.3	24.9	31.5	16.1	59.4

Table 2: Performance of the LFD, EXT, and D2 on the PSB base classification [19]. The length refers to the dimension of the feature space. Other values are in (%)

	length	NN	$1^{st}-$ tier	$2^{nd}-$ tier	E-measure	DCG
LFD	4500	65.7	38.0	48.7	28.0	64.3
EXT	153	54.9	28.6	37.9	21.9	56.2
H-EXT	33	28.1	24.5	31.3	16.3	58.6
D2	64	31.1	15.8	23.5	13.9	43.4

to the original paper for a complete evaluation. More precisely, we consider the:

1. **Lightfields descriptors (LFD) [1]**: the features representing a 3D model are extracted from 2D images, which are rendered from cameras positioned on the vertices of a regular dodecahedron. Each image is encoded with 35 coefficients of Zernike moments, and 10 coefficients to represent Fourier descriptors. The dimension of the feature space is then 4500.

2. **Spherical Extent Function (EXT) [17]:** this is a measure of the extent of the shape in the radial direction. It was computed on 64×64 spherical grid (latitude-longitude parameterization) and then represented by its harmonic coefficients up to order 16. We obtain feature vectors of 153 floating point numbers.

3. **Harmonics of the Spherical Extent Function (H-EXT) [11]:** a rotation invariant representation of the EXT obtained by computing the norm of each harmonic frequency. In our implementation, we consider the harmonic coefficients up to order 32 (similar to [11]) obtaining feature vectors of 33 floating point numbers. We used geometry images of size 128×128.

4. **Osada's $D2$ shape distribution (D2) [14]**: a one dimensional histogram that measures the distribution of the pairwise distance between pairs of random points on the shape surface. Similar to [19], we used histograms of 64 bins.

Table 3: Evaluating retrieval performance for the SWC_d descriptor on different classes using the PSB coarse2 test classification (6 classes).

	NN	1^{st}-tier	2^{nd} -tier	E-measure	DCG
Animal	74.2%	41.4%	63.0%	19.9%	82.7%
Vehicle	72.2%	36.5%	65.8%	11.0%	82.2%
Household	63.2%	23.5%	38.2%	11.4%	74.8%
Furniture	53.2%	8.6%	14.3%	7.7%	62.3%
Plant	25.0%	8.1%	15.3%	5.9%	55.6%
Buildings	10.6%	11.1%	19.0%	10.3%	56.2%

In the literature, the LFD is considered as the best descriptor. Table 2 shows the results according to the quantitative measures computed on these descriptors (the results of LFD, EXT and D2 are the one reported in the original paper [19], while the results of H-EXT are from our implementation).

These results indicate that, spherical wavelet descriptors perform better than the LFD, shape distributions and spherical harmonic descriptors on DCG measure. An interesting observation is that the lightfield descriptor, which is considered a very good signature [1], performs better than spherical wavelet descriptors for the $k-$nearest neighbors related measures (nearest neighbor, first and second tier), while the spherical wavelet descriptors perform better than the lightfields descriptor for the precision/recall measures (DCG), which are considered more indicative.

Spherical wavelet descriptors have several benefits over lightfields, shape distributions and spherical harmonic descriptors in terms of storage and computational costs. Table 1 and 2 summarize the length of each shape descriptor. An interesting result is that the performance on the DCG measure of the SWEL1, a very compact descriptor, is almost similar to the LFD. A comparison with the performance of the EXT and H-EXT descriptors shows that energy-based wavelet descriptors (SWEL1 and SWEL2) have several benefits: (1) compactness, (2) rotation invariant without pose normalization, and (3) easy to compute.

Performance on different shape classes

Finally, we evaluate the performance of the SWC_d and SWEL1 descriptors on different shape classes. Table 3 and Table 4 summarize the average performance of the two descriptors with respect to the quantitative measures. Six classes of the coarse2 test classification of the PSB are used. The results show that spherical wavelet coefficients perform better on animal, vehicle, household and furniture classes, while the L1 energy is more efficient on plants and building classes.

Table 4: Evaluating retrieval performance for the SWEL1 descriptor on different classes using the PSB coarse2 test classification (6 classes).

	NN	1^{st}-tier	2^{nd} -tier	E-measure	DCG
Animal	54.8%	35.6%	58.3%	16.1%	79.6%
Vehicle	65.3%	35.4%	64.8%	10.3%	81.4%
Household	47.6%	22.3%	37.0%	10.2%	73.5%
Furniture	47.9%	15.0%	22.9%	11.9%	65.8%
Plant	41.7%	16.6%	25.3%	15.3%	62.7%
Buildings	34.0%	14.5%	24.2%	13.2%	60.1%

6 Conclusions and future work

We proposed in this paper a spherical wavelet-based framework for the analysis of 3D shapes represented by functions on the sphere. We developed and tested using the Spherical Extent Function three new shape descriptors. Our results on the Princeton Shape Benchmark show that the new framework outperforms, in terms of Discount Cumulative Gain measure, the spherical harmonic based descriptor, while the spherical harmonic descriptors perform better on nearest neighbor measures. We found that our sampling procedure is more efficient since it is rotation invariant and samples uniformly all the shape features. An interesting property is that the SWEL1 descriptor, which is very compact, outperforms the LightField descriptor on the DCG measure.

Our best results have been achieved using SWC_3 after efficient pose normalization. We argue this improvement in the performance by the fact that spherical wavelet transform filters small details that affect negatively the performance, while it takes into account the spatial localization of the salient features. The SWEL1 and SWEL2 are equivalent to the power spectrum of the spherical harmonic analysis. They have many desirable properties. They are compact and faster to compute, and invariant under similarity transformations.

This work suggests a number of challenges that we would like to consider in the future. First we found from our experiments that the developed descriptors behave poorly on stick like shapes. We believe that this is the drawback of the sampling procedure. We plan in the future to elaborate more on this issue. Second, the proposed descriptors have been tested using only the Spherical Extent function. We will experiment with other representations. Third, we plan to test the proposed descriptors on other 3D model databases. Another issue is to experiment with different spherical wavelet basis and compare their performance on different classes of shapes. Finally, none of the developed descriptors perform equally in all situations and on all classes of shapes. A challenging issue is to investigate on how to combine and select features in order to achieve best performance.

Acknowledgement

We would like to thank Gabriel Peyre for providing us with an implementation of the second generation wavelets (GeoWave: Geometric Wavelets on Surfaces). We thank also the Princeton Shape Retrieval and Analysis Group which provided us with the Princeton Shape Benchmark (PSB). All models that appear in this paper are from the PSB.

References

[1] D.-Y. Chen, X.-P. Tian, Y.-T. Shen, and M. Ouhyoung. On visual similarity based 3d model retrieval. *Computer Graphics Forum*, 22(3):223–232, 2003.

[2] I. Daubechies. *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.

[3] M. N. Do and M. Vetterli. Texture similarity measurement using kullback-leibler distance on wavelet subbands. In *International Conference on Image Processing ICIP2000*, pages 730–733, 2000.

[4] M. N. Do and M. Vetterli. Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. *IEEE Transactions on Image Processing*, 11(2):146–158, February 2002.

[5] T. A. Funkhouser, P. Min, M. M. Kazhdan, J. Chen, J. A. Halderman, D. P. Dobkin, and D. P. Jacobs. A search engine for 3d models. *ACM Transactions on Graphics*, 22(1):83–105, 2003.

[6] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3d shapes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 203–212. ACM Press, 2001.

[7] H. Hoppe and E. Praun. Shape compression using spherical geometry images. In *MINGLE 2003 Workshop. In Advances in Multiresolution for Geometric Modelling, N. Dodgson, M. Floater, M. Sabin (eds.), Springer-Verlag*, number 2, pages 27–46, 2003.

[8] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, and K. Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, 37(5):509–530, 2005.

[9] A. Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.

[10] I. T. Jolliffe. *Principal Component Analysis*. Springer, 2nd edition edition, 2002.

[11] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. In *SGP '03: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing*,

pages 156–164, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[12] M. Novotni and R. Klein. 3d zernike descriptors for content based shape retrieval. In *SM '03: Proceedings of the eighth ACM symposium on Solid modeling and applications*, pages 216–225, New York, NY, USA, 2003. ACM Press.

[13] M. Novotni and R. Klein. Shape retrieval using 3d zernike descriptors. *Computer Aided Design*, 36(11):1047–1062, 2004.

[14] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Matching 3d models with shape distributions. In *Shape Modeling International*, pages 154–166, Genova, Italy, May 2001.

[15] E. Paquet, M. Rioux, A.Murching, T.Naveen, and A.Tabatabai. Description of shape information for 2-d and 3-d objects. *Signal Processing: Image Communication*, 16(1-2):103–122, 2000.

[16] E. Praun and H. Hoppe. Spherical parametrization and remeshing. *ACM Transactions on Graphics*, 22(3):340–349, 2003.

[17] D. Saupe and D. V. Vranic. 3d model retrieval with spherical harmonics and moments. In B. Radig and S. Florczyk, editors, *DAGM-Symposium*, volume 2191 of *Lecture Notes in Computer Science*, pages 392–397. Springer, 2001.

[18] P. Schroder and W. Sweldens. Spherical wavelets: efficiently representing functions on the sphere. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 161–172. ACM Press, 1995.

[19] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. In *SMI '04: Proceedings of the Shape Modeling International 2004 (SMI'04)*, pages 167–178. IEEE Computer Society, june 2004.

[20] H. Sundar, D. Silver, N. Gagvani, and S. J. Dickinson. Skeleton based shape matching and retrieval. In *Shape Modeling International*, pages 130–142, 2003.

[21] J. W. Tangelder and R. C. Veltkamp. A survey of content based 3d shape retrieval. In *Shape Modeling International 2004, Genova, Italy*, pages 145–156, June 2004.

[22] G. van de Wouwer, P. Scheunders, and D. van Dyck. Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing*, 8(4):592–598, April 1999.

[23] D. V. Vranic. An improvement of rotation invariant 3d-shape based on functions on concentric spheres. In *ICIP2003*, pages 757–760, 2003.

[24] D. V. Vranic and D. Saupe. Description of 3d-shape using a complex function on the sphere. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2002)*, pages 177–180, August 2002.

[25] D. V.Vranic. *3D Model Retrieval*. Phd dissertation, Universitat Leipzig, Institut Fur Informatik, 2003.