

DISCRIMINATIVE SPHERICAL WAVELET FEATURES FOR CONTENT-BASED 3D MODEL RETRIEVAL

HAMID LAGA

*Global Edge Institute, Tokyo Institute of Technology,
S6-401B-C 2-12-1 Ookayama Meguro-ku Tokyo 152-8552, Japan.
hamid@img.cs.titech.ac.jp*

MASAYUKI NAKAJIMA

*Computer Science Department, Tokyo Institute of Technology,
W8-64, 2-12-1 Ookayama Meguro-ku Tokyo 152-8552, Japan.
nakajima@img.cs.titech.ac.jp*

KUNIHIRO CHIHARA

*Graduate School of Information Science,
Nara Institute of Science and Technology,
8916-5 Takayama-cho Ikoma Nara 630-0101, Japan.
chihara@is.naist.jp*

The description of 3D shapes using features that possess descriptive power and are 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 a latitude-longitude parameterization of the sphere which has singularities at each pole, and therefore, variations of the north pole affect significantly the shape function. In this paper we discuss these issues and propose the usage of spherical wavelet transforms 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 threefold; 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 extensions of spherical harmonics and 3D Zernike moments. 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.

1991 Mathematics Subject Classification: 22E46, 53C35, 57S20

1. Introduction

The 21st century is the era of digital media and substantial progress has been made in the acquisition, storage and transmission of different types of information. While text, images, sound and video have been the predominant forms 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 huge amount of 3D data is nowadays available and therefore, significant research effort is required for developing effective techniques for content-based retrieval of 3D data.

Content-based 3D retrieval (CB3DR) implies the indexation of the 3D model database with geometric features extracted from the 3D models. A challenging issue is the description of shapes with suitable numerical representations called *shape descriptors*. In general a shape descriptor should be discriminative by capturing only the salient features, robust to noise, compact, easy to compute, and invariant under similarity transformations such as translation, rotation and scale^{1,2,3}. Other invariant properties may be required for some applications, such as pose invariance for matching articulated shapes^{4,5}.

In this paper we introduce a new 3D content-based retrieval method relying on the spherical wavelet transform (SWT) of the shape function. Spherical Wavelets have been proposed by Schröder and Sweldens⁶ and since, they have been used to solve many geometry processing problems including 3D model compression⁷. Similarly to first generation wavelets⁸, 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 harmonics⁹ and 3D Zernike moments^{10,11}. It offers better feature localization and all the advantages of wavelets over Fourier analysis.

2. Related work

2.1. Shape signatures

Most of 3D 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 of the related literature have been provided by Tangelder and Velkamp¹², and Iyer et al.¹³. Existing techniques 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 (LFD)¹⁴ are reported to be the most effective descriptor¹⁵. View-based techniques are suitable for implementing query interfaces using sketches^{13,9}.

Graph-based techniques are suitable for indexing articulated 3D models. They reduce the problem of 3D shape comparison to the problem of comparing graphs.

Reeb graphs⁴, and skeletons⁵ are among the most popular. Cornea et al.¹⁶ used the skeletal representation of 3D volumetric objects for many-to-many and part matching. Biasotti et al.¹⁷ proposed a matching framework for sub-part correspondences using graph matching. Their method builds the common sub-graphs between the two shapes to match and highlights the maximal sub-parts having similar structure and similar space distribution.

Graph matching is computationally very expensive, especially when the number of nodes is high and when the graphes to match have different number of nodes. Jain et al.¹⁸ avoid explicit comparison of graphes by using spectral techniques. First the shape is embedded into another feature space using spectral embedding, where similar shapes with different poses map to a single point. Shape descriptors can then be extracted and used to compare articulated 3D models.

Feature-based methods aim to extract compact descriptors from the 3D object. Johnson et al.¹⁹ introduced spin images as local features for matching 3D shapes. They have been used for shape retrieval as well as for the matching and registration of 3D scans. Other techniques are based on the distribution of features, such as shape distributions²⁰. Shilane et al.¹⁵ provided a comparison of these techniques and reported that histogram-based methods are the less efficient in terms of discriminative power.

Recently, Reuter et al.²¹ introduced the notion of shape DNA. They proposed fingerprints for shape matching. The fingerprints are computed from the spectra of the Laplace-Beltrami operators. These descriptors are invariant under similarity transformations, and are very efficient in matching 2D and 3D manifold shapes. However, it is not clear how they can be computed for polygon soup models.

2.2. Invariant shape features

The issue of extracting invariant shape features is an important problem in content-based 3D model retrieval. While translation and scale invariance can be easily achieved^{22,9,2}, 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 based on Principal Component Analysis (PCA)²³, such as continuous PCA²⁴, and other extensions for solving for axial ambiguity. However, PCA-based alignment is known to misbehave and therefore, it hampers significantly the retrieval performance².

A popular way to avoid explicit alignment of shapes is to represent the shape using functions defined on the unit sphere. Funckhouser et al.⁹ then use spherical harmonics (SH) to analyze the shape function. They demonstrated later that spherical harmonics can be used to achieve rotation invariance by taking only the power spectrum of the harmonic representation, and therefore, discarding the rotation dependent information². Novotni and Klein¹⁰ use 3D Zernike moments (ZD) as a natural extension of SH. Representing 3D shapes as functions on concentric

spheres has been extensively used. Our descriptors fall into this category and are a natural extension of SH and ZD.

2.3. Overview and contributions

This paper investigates the problem of extracting rotation invariant features and introduces spherical wavelet analysis for content-based 3D Model retrieval. To the best of our knowledge, spherical wavelets have not been applied to content-based retrieval of 3D models so far. We make use of them and propose three new descriptors; (1) spherical wavelet coefficients as feature vector (SWC_d), (2) the L1 energy of the spherical wavelet coefficients (SWEL1), and (3) the 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) The spherical wavelet descriptors we propose can be extracted from any spherical function. In our implementation, we experimented with the Spherical Extent Functions (EXT)³, and the Gaussian Euclidean Distance Transforms (GEDT)².
- (4) We evaluate and compare the performance of the proposed descriptors using the Princeton Shape Benchmark (PSB) evaluation tools.

In the next section we discuss the problems related to shape function sampling and motivate the use of spherical wavelet analysis. Section 4 reviews the general concepts of the spherical wavelet transform of functions on the sphere, and describes how we use them for 3D shape analysis. Section 5 describes in detail the new shape signatures and the similarity estimation method. Section 6 presents some experimental results. Finally, we summarize in Section 7 the main findings of this paper and discuss some issues for future research.

3. Rotation invariant shape description

One of the main issues in matching 3D models is the lack of proper parameterization. Spherical representations have been introduced in order to overcome this limitation. In this representation, each spherical location (θ, ϕ) , where $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$, encodes some shape properties, $f(\theta, \phi)$, measured at its corresponding location on the shape. The function f is called the *shape function*. The steps commonly used to compare 3D shapes are:

- (1) **Normalization.** Transform the center of mass of the object to the origin, and

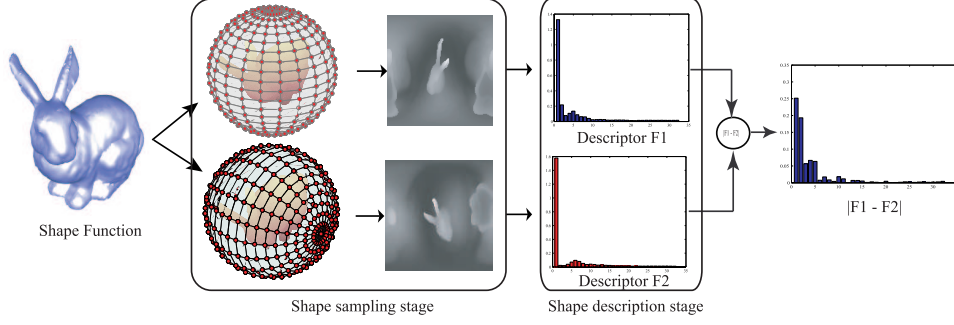


Fig. 1: Problem illustration: latitude-longitude parameterization generates shape functions with singularities at each pole affecting the rotation invariance of the shape descriptor.

scale the object to lie within a unit ball.

- (2) **Parameterization.** Compute the shape function. In the discrete case, the spherical shape function f is constructed by sampling the unit sphere, centered at the shape's center of mass, on a regular grid of size $w \times h$ of angles of azimuth ϕ , and elevation θ . For simplicity, we consider the Spherical Extent Function (EXT) ²². Other types of spherical functions will be considered in the experimental results section. The Spherical Extent Function $f(\theta, \phi)$ measures the extension of the shape in the radial direction (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 the *sampling stage*, and steps 3 and 4 as the *shape description stage*. Step 3 expresses the shape function in terms of its spherical harmonics:

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

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

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

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

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

forms a descriptor that is invariant to all rotations².

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 investigate the rotation invariance of the sampling stage.

3.1. *The irregular sampling problem*

At the sampling stage, the shape function f is sampled on a $w \times h$ regular grid along the azimuthal and elevation angles. Figure 1 shows the Bunny’s power spectrum descriptors^a computed using different poses (rotation of 90 degrees around the X axis), and the L_2 distance between the frequency components of the two descriptors. Note that:

- (1) the sampling is regular in the spherical coordinate frame, but not in the Euclidean space. Consequently, the obtained shape function depends heavily on the alignment of the 3D model. Increasing the sampling rate will alleviate this problem but at the cost of higher computation time.
- (2) the shape function obtained with the latitude-longitude sampling procedure has singularities at the north and south poles of the unit sphere, while the areas near the equator are under-sampled. Consequently, small variations of the shape near the two poles will affect significantly the descriptor.
- (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.

This shows clearly that, 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.

3.2. *Rotation invariant sampling*

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^{7,25}:

- (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 advantages of this representation are two fold; first it guarantees a uniform sampling of the shape since the vertices of the geodesic sphere are uniformly distributed

^aFor illustration purposes 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.

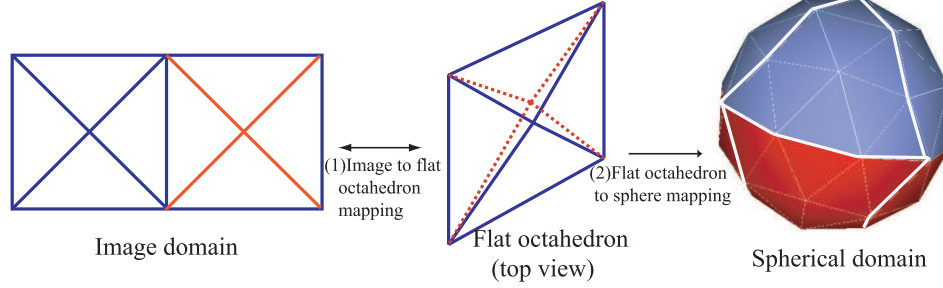


Fig. 2: Flat octahedron parameterization procedure. The flat octahedron is isometrically unfolded onto the image plane. The left half of the image plane is mapped to the top half of the flat octahedron, which is then mapped to the north hemisphere of the geodesic sphere.

on the surface of a unit sphere. Second, it allows a multiresolution analysis of the shape function where the coarsest (level-0) representation is obtained using a basic geodesic dome of 20 vertices, and finer levels are derived by recursive subdivisions.

- (2) **Flat octahedron parameterization.** Hoppe and Praun^{7,25} map 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 rectangular lattice. Therefore, image processing tools can be used with simple boundary extension rules.

These two representations guarantee a uniform sampling of the shape function and eliminate the singularities that appear at each pole in the latitude-longitude parameterization. Therefore, the discrete spherical shape function becomes rotation invariant within the sampling resolution. We make use of these properties to build efficient shape descriptors.

3.3. The geometry image

Our goal is to represent every 3D model O in the database by a geometry image I of size $k = w \times h$. We do this by first mapping the object to a unit sphere then unfolding the sphere onto the image domain using flat-octahedron parameterization (we will justify this choice in Section 4.2). 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 $P = \{p_1, \dots, p_k\}$ on the sphere.

The mapping function associates to each image pixel a point p_i on the unit sphere.

- (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 the radial direction p_i . f has the domain I as a regular support.

Steps 1 and 2 are common to all objects, therefore, we run them once offline to generate the set of points P . During runtime, all we need is to compute the shape function (step 3), which requires the computation of ray-polygon soup model intersections. This is a well studied problem, and can be performed at interactive rates. In our implementation we used the method proposed in ²⁴.

4. Spherical wavelets for 3D shape description

We now consider the problem of descriptor extraction from the spherical shape function. A straightforward approach is to use a subset of the vector f as shape descriptor ²⁴, but it is well established that the L_p metric is not effective in the spatial domain. On the other hand, spherical harmonics cannot be engaged since the sampling is not uniform in terms of azimuthal and elevation angles. In this paper, we make use of wavelets ^{6,7} to efficiently extract shape descriptors. In the following subsections we will review the general concepts and then describe how we use them to analyze the shape function.

4.1. 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 and Sweldens ⁶ introduced second generation wavelets. The idea 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:

$$\begin{aligned}
 \textbf{Analysis:} \quad & \lambda_{j,k} = \sum_{l \in K(j)} \tilde{h}_{j,k,l} \lambda_{j+1,l} \\
 \text{(forward transform)} \quad & \gamma_{j,k} = \sum_{l \in M(j)} \tilde{g}_{j,m,l} \lambda_{j+1,l} \\
 \\
 \textbf{Synthesis:} \quad & \lambda_{j+1,l} = \sum_{k \in K(j)} h_{j,k,l} \lambda_{j,k} + \\
 \text{(backward transform)} \quad & \sum_{m \in M(j)} g_{j,m,l} \gamma_{j,m}
 \end{aligned}$$

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 correspond to the 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 $j, j = n - 1, \dots, 0$. The coarsest approximation $\lambda_{n-i,\bullet}$ is obtained after i iterations ($0 < i \leq n$). The sets $M(j)$ and $K(j)$ are 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.

4.2. Analysis of the spherical shape function

To analyze a 3D model, we first apply the 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 moments¹⁰ and spherical harmonics^{2,3}, 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 second generation wavelets⁶ including the linear and butterfly spherical wavelets with lifting scheme, and image wavelets with spherical boundary extension rules⁷. In our experiments on the Princeton Shape Benchmark, we found that the performance of both the linear and butterfly spherical wavelets is very low (it is 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 3.3.
 - (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 $C^{(0)}, \dots, C^{(n-1)}$ are collected into a vector F .

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

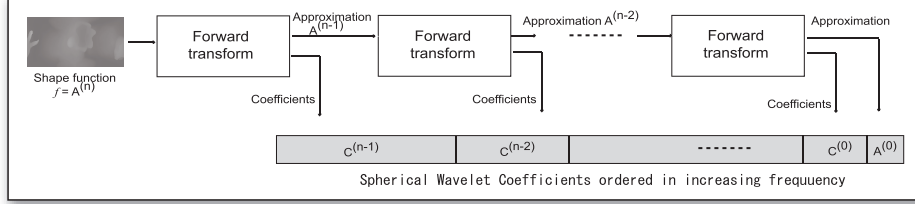


Fig. 3: Spherical wavelet-based shape descriptors computation.

is designed to take the difference between every two samples in the signal. Other wavelet bases can also be used but require further investigations.

5. Spherical wavelet-based descriptors

We now consider the computation of shape descriptors. We propose three methods to compare 3D models 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 spherical wavelet energies: (2) 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.

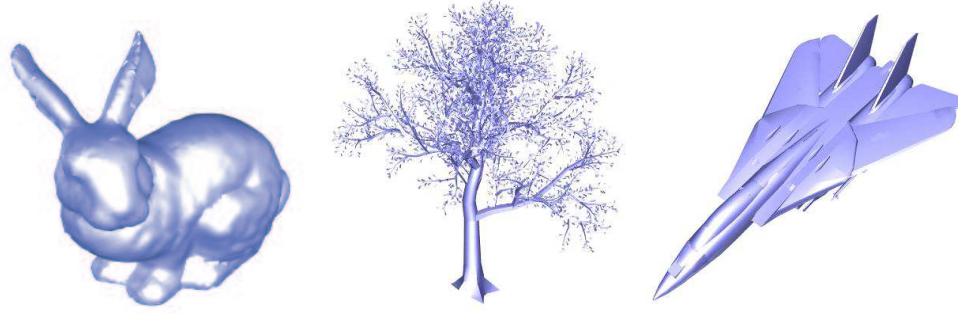
5.1. Wavelet coefficients as shape descriptor

Once the spherical wavelet transform is performed, one may use the wavelet coefficients as a shape descriptor. Using the entire set of 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 $d \in \{0, \dots, n-1\}$. In our implementation we used $d = 3$, therefore we obtain two dimensional feature vectors F of size $N = 2^{d+2} \times 2^{d+1} = 32 \times 16$.

Comparing the wavelet coefficients directly requires efficient alignment of the 3D models prior to wavelet transform. A popular method for finding the reference coordinate frame is pose normalization using Principal Component Analysis (PCA)¹, and continuous PCA²⁴. 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 PCA²⁴. 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 and any point in the surface is equal to $1/2$.

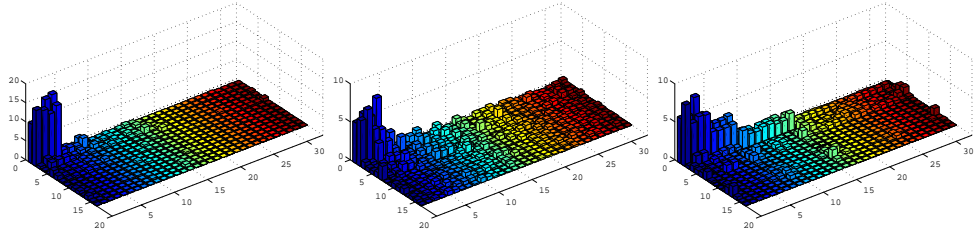
Figure 4c shows the SWC_3 descriptors of the 3D models shown in Figure 4a. Note that, the vector F provides an embedded multi-resolution representation for 3D



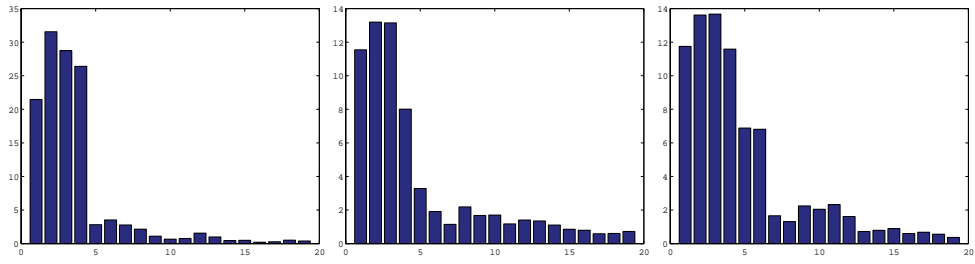
(a) 3D shapes



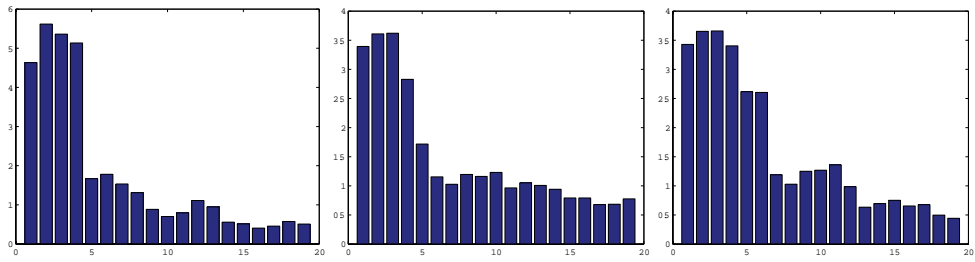
(b) Their associated geometry images of size $w \times h = 256 \times 128$



(c) Spherical wavelet coefficients as descriptor (SWC₃).



(d) L_2 energy descriptor (SWEL2).



(e) L_1 energy descriptor (SWEL1).

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

shape features. This approach performs a filtering of the 3D shape by removing outliers. A major difference to spherical harmonics is that SWT preserves the localization and orientation of local features. However, a feature space of dimension 512 is still computationally very expensive.

5.2. Spherical wavelet energy signatures

The wavelet energy signatures have been proven to be very powerful for texture characterization²⁶. Commonly the L_2 and L_1 norms are used as measures^{27,28}:

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

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

where $x_{l,j}, j = 1 \dots k_l$ are the wavelet coefficients of the l^{th} wavelet sub-band, and k_l is the number of coefficients in the l^{th} wavelet sub-band. To construct the wavelet energy based shape descriptor we first 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 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 L_1 energy and L_2 energy-based descriptors by **SWEL1** and **SWEL2**, respectively.

Observe that rotating a spherical function does not change its energy, therefore, spherical wavelet-based energy descriptors are invariant under any rotation along the axes of the coordinate frame. Since the sampling stage is also rotation invariant, we obtain shape descriptors that are invariant to general rotations. However, similar to the power spectrum², information such as feature localization is lost in the energy spectrum. Finally, the energy descriptor is also very compact. Thus, the storage and computation time required for comparisons are reduced.

Finally, Table 1 summarizes the performance of the proposed descriptors. The SWEL1 and SWEL2 are more efficient in terms of storage requirement and comparison time. They are also rotation invariant, while SWC_d requires pose normalization. The discrimination efficiency of each descriptor will be evaluated and discussed in the experimental results section.

5.3. The similarity metric

Since 3D shapes are now represented in the feature space with N -dimensional vectors of real-valued components a natural way to measure the dissimilarity between two models is to use the vector norms, called also L_p distances. 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:

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

Note 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 this case, the dissimilarity metric should be also adapted. 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 allows different retrieval requirements.

6. Experimental results

We have implemented the algorithms described in this paper and evaluated their performance on the Princeton Shape Benchmark (PSB)¹⁵. 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.

To evaluate the efficiency of spherical wavelet analysis for shape retrieval we considered two types of spherical functions:

- (1) **Spherical Extent Function (EXT)**²²: this is a measure of the extent of the shape in the radial direction. We compute the spherical wavelet descriptors SWC_d , $SWEL1$ and $SWEL2$ of length 512, 19, and 19, respectively. We refer to these descriptors by EXT_SWC_d , EXT_SWEL1 and EXT_SWEL2 respectively.
- (2) **Gaussian Euclidean Distance Transform (GEDT)**²: a 3D function whose value at each point is given by the composition of a Gaussian with the Euclidean distance transform of the surface^{2,15}.

In our implementation the parameter σ of the gaussian is set to 0.5. We compute the GEDT on a $65 \times 65 \times 65$ grid, then compute 32 spherical functions representing the intersection of the voxel grid with concentric spherical shells, in the same manner as in¹⁵. We analyze the spherical functions and extract spherical wavelet descriptors SWC_d , $SWEL1$ and $SWEL2$ of length 32×512 , 32×19 and 32×19 respectively. We refer to these descriptors by $GEDT_SWC_d$, $GEDT_SWEL1$ and $GEDT_SWEL2$, respectively.

In both cases we used spherical functions of size $w \times h = 256 \times 128$.

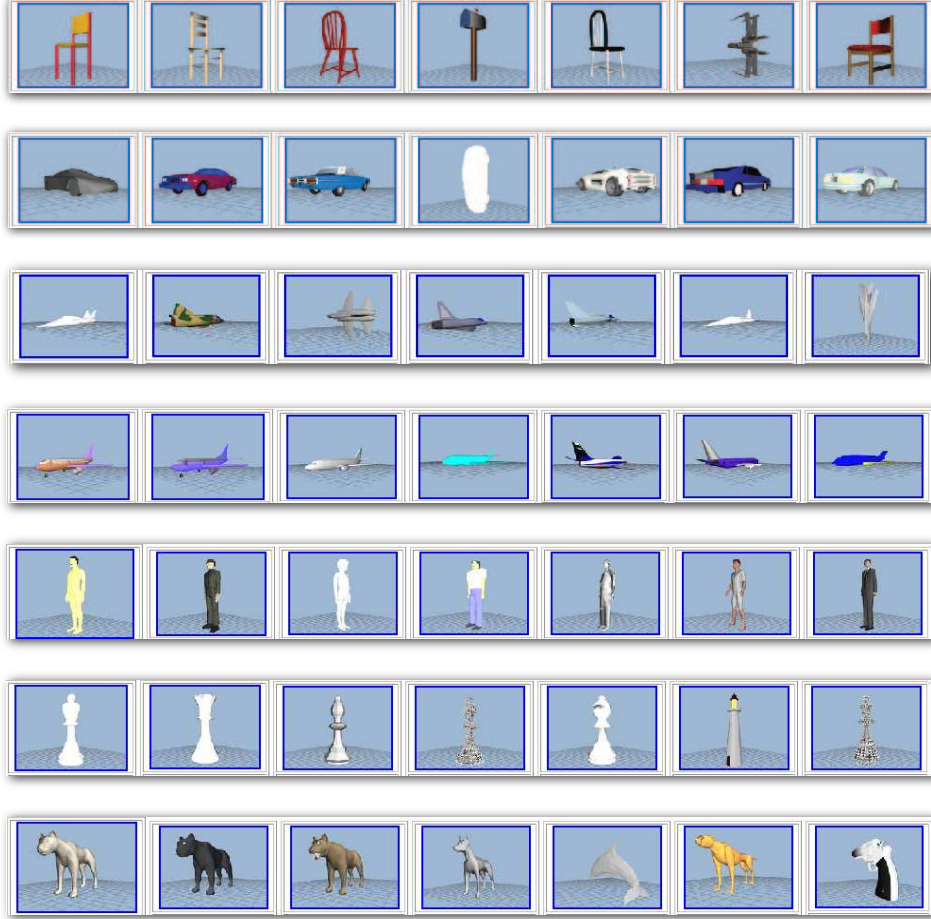


Fig. 5: Some retrieval results from the PSB database¹⁵ using the (SWC_d) descriptor: First column is the query shape followed by top six matches. The first six queries are from the query set provided in the first 3D Shape Retrieval Evaluation Contest (SHREC06)²⁹, while the last row query belongs to the PSB database¹⁵.

6.1. The retrieval results

First we executed series of shape matching experiments on the base test classification of the PSB using Spherical Extent Functions. We select randomly a 3D polygon soup model, and then compare it to the objects in the database. We show in Figures 5, 6, 7 the results of several queries for each of our three descriptors EXT_SWC_d , EXT_SWEL1 and EXT_SWEL2 . The top six matches are displayed. A retrieved model is considered relevant if it belongs to the same class as the query model.

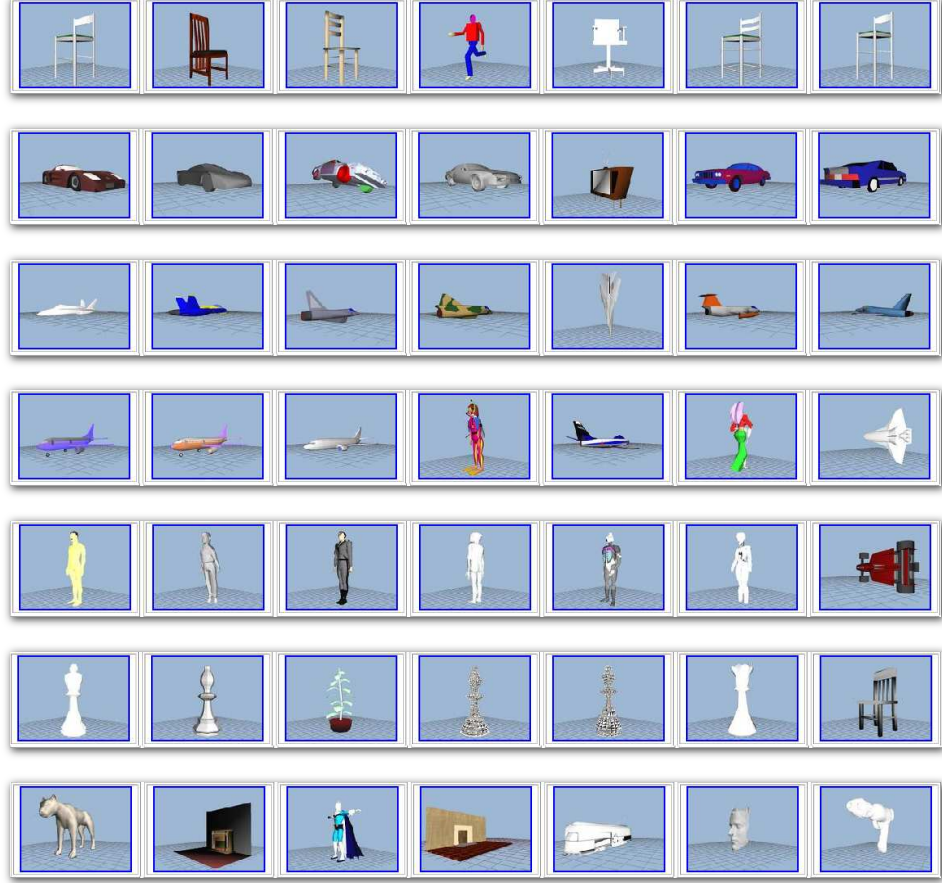


Fig. 6: Some retrieval results from the PSB database¹⁵ using the spherical wavelet L1-energy descriptor (SWEL1): First column is the query shape followed by top six matches. The first six queries are from the query set provided in the first 3D Shape Retrieval Evaluation Contest (SHREC06)²⁹, while the last row query belongs to the PSB database¹⁵.

By visually inspecting these results, we noticed that the EXT_SWC_d descriptor performs better than the others. The L_1 energy of the spherical wavelet coefficients is ranked second.

6.2. Performance evaluation

The precision-recall curves on the base test classifications of the PSB of the spherical wavelet-based shape descriptors are shown in Figure 8. We refer the reader to the Princeton Shape Benchmark paper¹⁵ for comparison with other descriptors

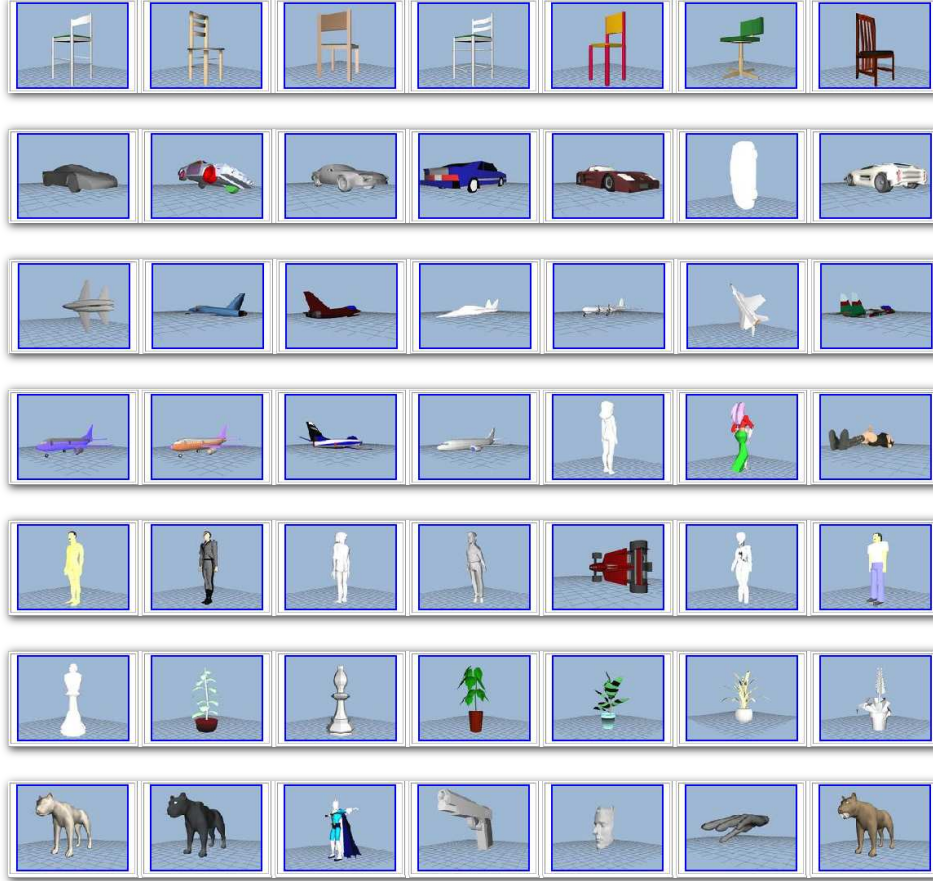


Fig. 7: Some retrieval results from the PSB database¹⁵ using the spherical wavelet L1-energy descriptor (SWEL2): First column is the query shape followed by top six matches. The first six queries are from the query set provided in the first 3D Shape Retrieval Evaluation Contest (SHREC06)²⁹, while the last row query belongs to the PSB database¹⁵.

concerning 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 measures¹⁵. Table 1 summarizes the micro-averaged retrieval statistics of our descriptors. We performed all the experiments on the base test classification of the PSB. Table 1 confirms the visual evaluation, that is, spherical wavelet coefficients perform better, while the L1 and L2 energy come second and third, respectively. Note that the SWC_d requires more storage and comparison time.

Shilane et al.¹⁵ summarized the performance on the PSB of several shape

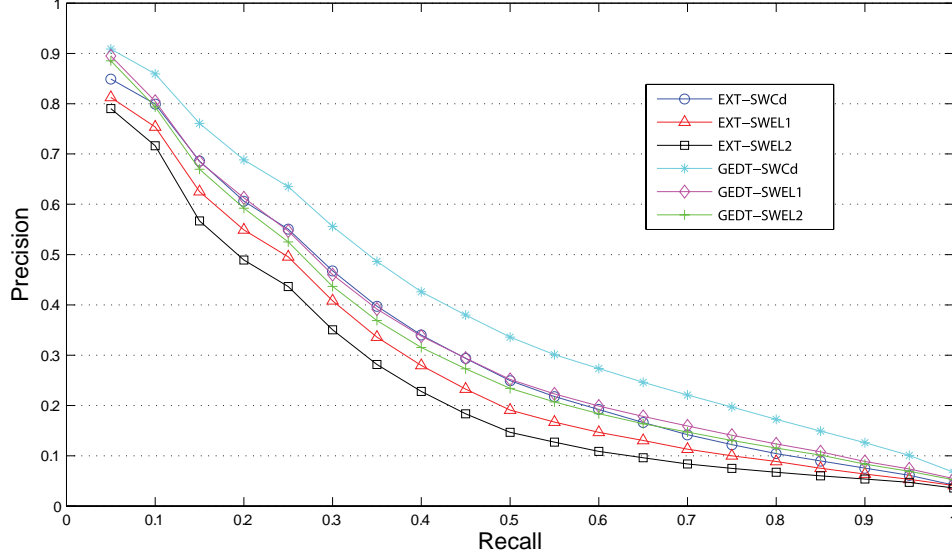


Fig. 8: Precision-recall curves for SW-based descriptors.

descriptors and we use their results to compare with our descriptors. In this paper, we show the performance of six descriptors, but we refer the reader to the original paper for a complete evaluation. More precisely, we consider the:

- (1) **Lightfields descriptors (LFD)** ¹⁴: 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) **Gaussian Euclidean Distance Transform (GEDT)** ²: Each spherical shell of the GEDT is represented by its spherical harmonic coefficients up to order 16 ^{2,15}. It uses the latitude-longitude parameterization.
- (3) **Spherical Harmonic Descriptor (SHD)** ²: a rotation invariant representation of the GEDT obtained by computing the restriction of the function to concentric spheres and storing the norm of each harmonic frequency ^{2,15}.
- (4) **Spherical Extent Function (EXT)** ²²: It was computed on a 64×64 spherical grid using the latitude-longitude parameterization and then represented by its harmonic coefficients up to order 16. We obtain feature vectors of 153 floating point numbers.
- (5) **Harmonics of the Spherical Extent Function (H-EXT)** ²: 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 ²) obtaining feature vectors of 33 floating point

Table 1: Performance of SW descriptors on the PSB base test classification. Size refers to the dimension of the feature space.

		Size	Nearest Neighbor	First tier	Second tier	E-measure	DCG
Spherical Extent Function (EXT)	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%
Gaussian Euclidean Distance Transform (GEDT)	SWC _d .	16384	53.6%	37.7%	47.1%	27.9%	69.8%
	SWEL1	608	42.2%	31.8%	40.7%	22.3%	65.5%
	SWEL2	608	40.2%	30.7%	38.8%	21.3%	64.5%

Table 2: Performance of the LFD, SHD, GEDT, EXT, H-EXT and D2 on the PSB base classification. Size refers to the dimension of the feature space.

	Size	Nearest Neighbor	First tier	Second tier	E-measure	DCG
LFD	4500	65.7%	38.0%	48.7%	28.0%	64.3%
SHD	544	55.6%	30.9%	41.1%	24.1%	58.4%
GEDT	4896	60.3%	31.3%	40.7%	23.7%	58.4%
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%

numbers. We used geometry images of size 128×128 .

- (6) **Osada’s D2 shape distribution (D2)**²⁰: a one dimensional histogram that measures the distribution of the pairwise distance between pairs of random points on the shape surface. Similar to¹⁵, we used histograms of 64 bins.

In the literature, the LFD is considered to be 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 ones reported in the original paper¹⁵, while the results of H-EXT are from our implementation). Table 1 shows that the GEDT-based wavelet descriptors outperform significantly the spherical extension function-based wavelet descriptors. This was predictable since the GEDT takes into account interior details of the shape.

Now comparing to other methods, spherical wavelet descriptors perform better than the LFD, shape distributions and spherical harmonic descriptors according to the DCG measure. An interesting observation is that the lightfield descriptor, which is considered a very good signature¹⁴, 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.

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

	Nearest Neighbor	First tier	Second 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%

Table 4: Evaluating the retrieval performance for the EXT_SWEL1 descriptor on different classes using the PSB coarse2 test classification (6 classes).

	Nearest Neighbor	First tier	Second 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%

An interesting result is that the GEDT-based wavelet descriptors outperform most of the existing descriptors on all measures. The GEDT_SWC_d is very expensive in terms of memory storage. However GEDT_SWEL1 and GEDT_SWEL2, which are very compact and rotation invariant, achieved very good performance compared to SHD, EXT, GEDT, H-EXT and D2, and outperform the LFD on the DCG and precision-recall measures.

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 size 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 invariance without pose normalization, and (3) they are easy to compute.

6.3. Performance on different shape classes

Finally, we evaluated the performance of the EXT_SWC_d and EXT_SWEL1 descriptors on different shape classes. Table 3 and Table 4 summarize the micro-averaged

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.

7. Conclusions and future work

We proposed in this paper a spherical wavelet-based framework for the search and retrieval of 3D shapes represented by functions on the sphere. We developed and tested using the Spherical Extent Function (EXT) and the Gaussian Euclidean Distance Transform (GEDT) three new shape descriptors. Our results on the Princeton Shape Benchmark show that the new framework outperforms, in terms of the Discount Cumulative Gain measure, the spherical harmonic based descriptors, 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, performs similarly to the LightField descriptor on the DCG measure when applied to EXT and outperforms the Lightfields when applied to the GEDT.

Our best results have been obtained using SWC_3 after efficient pose normalization. We explain this improvement in the performance by the fact that the 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 as 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 only on the Princeton Shape Benchmark, we plan to evaluate them on other 3D model databases. Another issue is to experiment with different spherical wavelet bases and compare their performance on different classes of shapes. Finally, none of the developed descriptors perform equally well 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.

Acknowledgements

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. This work is partially supported by the Japan Society for the Promotion of Science (JSPS).

References

1. Eric Paquet, Mark 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.
2. Michael Kazhdan, Thomas Funkhouser, and Szymon 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.
3. Dejan V. Vranic. An improvement of rotation invariant 3D-shape based on functions on concentric spheres. In *Proceedings of the IEEE International Conference on Image Processing (ICIP 2003)*, pages 757–760, 2003.
4. Masaki Hilaga, Yoshihisa Shinagawa, Taku Kohmura, and Tosiyasu 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.
5. H. Sundar, Deborah Silver, Nikhil Gagvani, and Sven J. Dickinson. Skeleton based shape matching and retrieval. In *SMI'03: Proceedings of the Shape Modeling International 2003*, pages 130–142, Washington, DC, USA, 2003. IEEE Computer Society.
6. Peter Schroder and Wim 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.
7. Hugues Hoppe and Emil Praun. Shape compression using spherical geometry images. In *Advances in Multiresolution for Geometric Modelling, N. Dodgson, M. Floater, M. Sabin (eds.), Springer-Verlag*, number 2, pages 27–46, 2003.
8. Ingrid Daubechies. *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
9. Thomas A. Funkhouser, Patrick Min, Michael M. Kazhdan, Joyce Chen, J. Alex Halderman, David P. Dobkin, and David Pokrass Jacobs. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1):83–105, 2003.
10. Marcin Novotni and Reinhard 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.
11. Marcin Novotni and Reinhard Klein. Shape retrieval using 3D Zernike descriptors. *Computer Aided Design*, 36(11):1047–1062, 2004.
12. Johan W.H Tangelder and Remco C. Veltkamp. A survey of content based 3D shape retrieval. In *Shape Modeling International 2004, Genova, Italy*, pages 145–156, June 2004.
13. Natraj Iyer, Subramaniam Jayanti, Kuiyang Lou, Yagnanarayanan Kalyanaraman, and Karthik Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, 37(5):509–530, 2005.
14. Ding-Yun Chen, Xiao-Pei Tian, Yu-Te Shen, and Ming Ouhyoung. On visual similarity based 3D model retrieval. *Computer Graphics Forum*, 22(3):223–232, 2003.
15. Philip Shilane, Patrick Min, Michael Kazhdan, and Thomas 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.
16. Nicu D. Cornea, M. Fatih Demirci, Deborah Silver, Ali Shokoufandeh, Sven J. Dickin-

- son, and Paul B. Kantor. 3D object retrieval using many-to-many matching of curve skeletons. In *SMI '05: Proceedings of the International Conference on Shape Modeling and Applications 2005 (SMI' 05)*, pages 368–373, Washington, DC, USA, 2005. IEEE Computer Society.
17. Silvia Biasotti, Simone Marini, Michela Spagnuolo, and Bianca Falcidieno. Subpart correspondence by structural descriptors of 3D shapes. *Computer-Aided Design*, 38(9):1002–1019, 2006.
18. Varun Jain and Hao Zhang 0002. Shape-based retrieval of articulated 3D models using spectral embedding. In Myung-Soo Kim and Kenji Shimada, editors, *Geometric Modeling and Processing - GMP 2006, 4th International Conference, Pittsburgh, PA, USA, July 26-28, 2006, Proceedings*, volume 4077 of *Lecture Notes in Computer Science*, pages 299–312. Springer, 2006.
19. Andrew Johnson. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, August 1997.
20. Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Matching 3D models with shape distributions. In *Shape Modeling International*, pages 154–166, Genova, Italy, May 2001.
21. Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace-Beltrami spectra as "shape-DNA" of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.
22. Dietmar Saupe and Dejan V. Vranic. 3D model retrieval with spherical harmonics and moments. In Bernd Radig and Stefan Florczyk, editors, *DAGM-Symposium*, volume 2191 of *Lecture Notes in Computer Science*, pages 392–397. Springer, 2001.
23. I. T. Jolliffe. *Principal Component Analysis*. Springer, 2nd edition edition, 2002.
24. Dejan V. Vranic. *3D Model Retrieval*. PhD dissertation, Universitat Leipzig, Institut Fur Informatik, 2003.
25. Emil Praun and Hugues Hoppe. Spherical parametrization and remeshing. *ACM Transactions on Graphics*, 22(3):340–349, 2003.
26. Van de Wouwer Gert, Paul Scheunders, and Van Dyck Dirk. Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing*, 8(4):592–598, April 1999.
27. Minh N. Do and Martin Vetterli. Texture similarity measurement using Kullback-Leibler distance on wavelet subbands. In *International Conference on Image Processing ICIP2000*, pages 730–733, 2000.
28. Minh N. Do and Martin Vetterli. Wavelet-based texture retrieval using generalized gaussian density and Kullback-Leibler distance. *IEEE Transactions on Image Processing*, 11(2):146–158, February 2002.
29. Remco C. Velkamp, Remco Ruijsenaars, Michela Spagnuolo, Roelof van Zwol, and Frank ter Haar. SHREC2006: 3D Shape Retrieval Contest. Technical Report UU-CS-2006-030, Department of Information and Computing Sciences, Utrecht University, 06 2006.