Eurographics Symposium on Geometry Processing (2003)
L. Kobbelt, P. Schröder, H. Hoppe (Editors)

Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors

Michael Kazhdan, Thomas Funkhouser, and Szymon Rusinkiewicz

Department of Computer Science, Princeton University, Princeton NJ

Abstract

One of the challenges in 3D shape matching arises from the fact that in many applications, models should be considered to be the same if they differ by a rotation. Consequently, when comparing two models, a similarity metric implicitly provides the measure of similarity at the optimal alignment. Explicitly solving for the optimal alignment is usually impractical. So, two general methods have been proposed for addressing this issue: (1) Every model is represented using rotation invariant descriptors. (2) Every model is described by a rotation dependent descriptor that is aligned into a canonical coordinate system defined by the model. In this paper, we describe the limitations of canonical alignment and discuss an alternate method, based on spherical harmonics, for obtaining rotation invariant representations. We describe the properties of this tool and show how it can be applied to a number of existing, orientation dependent descriptors to improve their matching performance. The advantages of this tool are two-fold: First, it improves the matching performance of many descriptors. Second, it reduces the dimensionality of the descriptor, providing a more compact representation, which in turn makes comparing two models more efficient.

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques

1. Introduction

Over the last decade, tools for acquiring and visualizing 3D models have become integral components of data processing in a number of disciplines, including medicine, chemistry, architecture and entertainment. With the proliferation of these tools, we have also witnessed an explosion in the number of available 3D models. As a result, the need for the ability to retrieve models from large databases has gained prominence and a key concern of shape analysis has shifted to the design of efficient and robust matching algorithms.

One of the principal challenges faced in the area of shape matching is that in many applications, a shape and its image under a similarity transformation are considered to be the same. Thus, the challenge in comparing two shapes is to find the best measure of similarity over the space of all transformations. The need for efficient retrieval makes it impractical to test all possible transformations explicitly, and two different solutions have been proposed:

- **Normalization**: Shapes are placed into a canonical coordinate frame (normalizing for translation, scale and rotation) and two shapes are assumed to be optimally aligned when each is in its own frame. Thus, the best measure of similarity can be found without explicitly trying all possible transformations.
- **Invariance**: Shapes are described in a transformation invariant manner, so that any transformation of a shape will be described in the same way, and the best measure of similarity is obtained at *any* transformation.

We have found that while traditional methods for translation and scale normalization provide good matching results, methods for rotation normalization are less robust and hamper the performance of many descriptors.

In this paper we discuss the *Spherical Harmonic Representation*, a method for transforming rotation dependent shape descriptors into rotation independent ones. The general outline of this approach has been described in earlier works[1, 2, 3], and the contribution of our work consists of eval-

uating this method in the context of shape matching. In particular, (1) we provide an analysis of the limitations of traditional methods for rotation normalization, (2) we discuss the mathematical implications, such as information loss, of the Spherical Harmonic Representation when applied to existing shape descriptors, and (3) we present results evaluating the efficacy of the representation in tasks of model retrieval, demonstrating that this representation provides a more robust, concise and efficient method for matching 3D shapes.

The rest of this paper is structured as follows. In Section 2, we review previous work in the area of shape retrieval. We describe the Spherical Harmonic Representation in Section 3, in which we summarize the principal properties of spherical harmonics and provide a method for obtaining rotation invariant representations of spherical-based shape descriptors. In Section 4, we discuss the mathematical properties of the Spherical Harmonic Representation. We provide a generalization of our method to voxel grids in Section 5. In Section 6, we present empirical results comparing matching results of normalized descriptors with their rotation invariant representations, and we provide an analysis of these results in Section 7. Finally, we conclude in Section 8 by summarizing our contribution and discussing topics for future work.

2. Related Work

The problem of shape matching has been well studied in the graphics/vision literature and many methods for evaluating model similarity have been proposed. Motivated by the increased availability and accessibility of 3D models, this paper focuses on the problem of shape retrieval from within large databases of models. In this context, the challenge is to provide a robust and efficient method for computing the similarity between whole shapes.

To address this challenge, many methods have focused on separating the matching problem into two components: (1) an offline step, in which abstracted distinguishing information is extracted from each model independently, and (2) an online step, in which the information between two models is compared. In order to allow for efficient retrieval, the offline step is usually designed to extract information which allows for simple and efficient comparison between models. In particular, many existing methods describe a 3D shape with an abstracted *shape descriptor* that is represented as a function defined on a canonical domain. Shapes are then compared by computing the difference between their descriptors over the canonical domain, so that no explicit establishing of correspondences is necessary, and the online process can be efficient.

However, in the context of shape retrieval, one of the principal difficulties faced by these approaches is that a shape and its image under a similarity transformation are considered to be the same. Thus, the challenge in comparing two shapes is to find the best measure of similarity over the space of all transformations. This challenge has been addressed in two different ways:

- Normalizing the models by finding a canonical transformation for each one.
- Characterizing models with a transformation invariant descriptor so that all transformations of a model result in the same descriptor.

While explicitly solving for the optimal transformation using either exhaustive search or methods such as the ICP[4, 5] algorithm, the Generalized Hough Transform[6], or Geometric Hashing[7], are also possible, these approaches are not applied to database retrieval tasks since the online comparison of models becomes inefficient. Many hybrid methods exists, and a few representative examples are shown in Table 1, which describes how these methods address translation, scale and rotation.

Representation	Tr	Sc	Rot
Crease Histograms[8]	I	N	I
Shape Distributions[9]	I	N	I
Extended Gaussian Images[10]	I	N	N
Shape Histograms[11] (Shells)	N	N	I
Shape Histograms[11]	N	N	N
Spherical Extent Functions[12]	N	N	N
Wavelets[13]	N	N	N
Reflective Symmetry Descriptors[14]	N	N	N
Higher Order Moments[15]	N	N	N
Exponentiated EDT[16]	N	N	N

Table 1: *A summary of a number of shape descriptors, showing if they are (N)ormalized or (I)nvariant to each of translation, scale and rotation.*

In general, models are normalized by using the center of mass for translation, the root of the average square radius for scale, and principal axes for rotation. We have found that while the methods for translation and scale normalization are robust for whole object matching[17], rotation normalization via PCA-alignment does not provide a robust normalization for many matching applications. This is due to the fact that PCA-alignment is performed by solving for the eigenvalues of the covariance matrix which captures only second order model information. The assumption in using PCA is that the alignment of higher frequency information is strongly correlated with the alignment of the second order components. (Appendix A provides an analysis of this from a signal processing framework.) We have found that for many shape descriptors this assumption does not hold, and the use of principal axes for alignment hampers the performance of these descriptors.

Many of the descriptors that have used PCA-alignment

represent a 3D shape as either a spherical function or a voxel grid, which rotates with the model. Examples of such descriptors have included:

- The Extended Gaussian Image[10], which describes the distribution of normals across the surface of the model
- Shape Histograms[11], which describe the distribution of points on the model across all rays from the origin
- Spherical Extent Functions[12], which describe the maximal extent of a shape across all rays from the origin
- Reflective Symmetry Descriptors[14], which describe the reflective self-similarity of a shape with respect to reflections about all planes through the origin
- The voxel description of Funkhouser et al.[16], which describes a model by computing the negative exponential of its Euclidean Distance Transform

For this type of descriptor, we propose a solution to the rotation problem by describing a mathematical tool, based on spherical harmonics, for obtaining a rotation invariant representation of the descriptors. The approach is a generalization of the Fourier Descriptor[18] method to the sphere, based on the computation of 0-th order tensors from spherical harmonic decompositions[1, 2, 3]. It characterizes spherical functions by the energies contained at different frequencies, and its application to shape matching was initially proposed in[19, 16]. This paper gives a detailed review of the descriptor, provides a mathematical analysis of its properties, and presents empirical results demonstrating its efficacy in improving the matching performance of a number of existing shape descriptors.

3. Spherical Rotation Invariance

In this paper, we analyze a method for transforming rotation dependent spherical shape descriptors into rotation invariant ones. The key idea of this approach is to describe a spherical function in terms of the amount of energy it contains at different frequencies. Since these values do not change when the function is rotated, the resulting descriptor is rotation invariant. This approach is described in[1, 2, 3] and can be viewed as a generalization of the Fourier Descriptor method[18] to the case of spherical functions.

3.1. Spherical Harmonics

In order to be able to represent a function on a sphere in a rotation invariant manner, we utilize the mathematical notion of spherical harmonics to describe the way that rotations act on a spherical function. The theory of spherical harmonics states that any spherical function $f(\theta,\phi)$ can be decomposed as the sum of its harmonics:

$$f(\theta,\phi)=\sum_{l=0}^{\infty}\sum_{m=-l}^{m=l}a_{lm}Y_l^m(\theta,\phi).$$

(This decomposition is visualized in step (1) of Figure 1.) The key property of this decomposition is that if we restrict

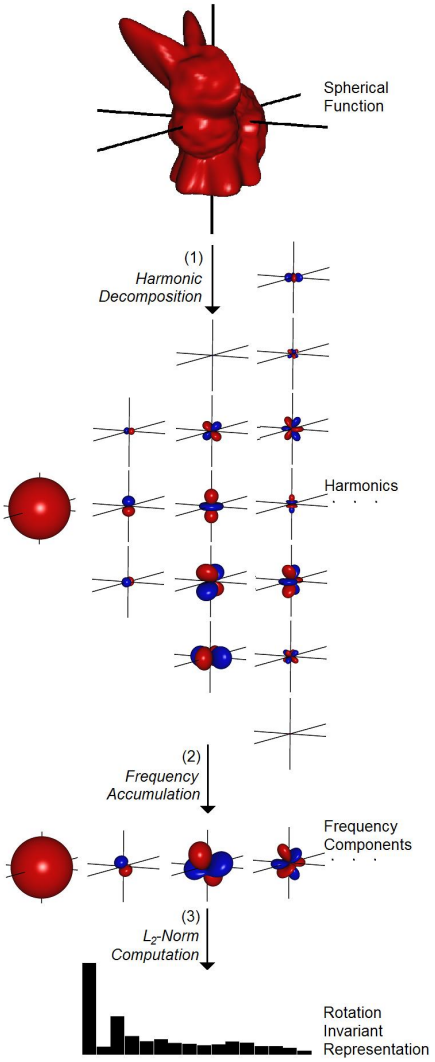

Figure 1: *We compute a rotation invariant descriptor of a spherical function by (1) decomposing the function into its harmonics, (2) summing the harmonics within each frequency, and (3) computing the norm of each frequency component. (Spherical functions are visualized by scaling points on the sphere in proportion to the value of the function at that point, where points with positive value are drawn in light gray and points with negative value are drawn in dark gray.)*

to some frequency l, and define the subspace of functions:

$$V_l=\mathrm{Span}(Y_l^{-l},Y_l^{-l+1},\ldots,Y_l^{l-1},Y_l^{-l})$$

then:

- **$\mathbf{V}_l$ is a Representation for the Rotation Group**: For any function $f\in V_l$ and any rotation R, we have $R(f)\in V_l$. That is, rotating any function described in terms of Y_l^m,

for fixed l, may shuffle information among different m but will never transfer energy to different l; i.e., "frequency" (which corresponds to l) is preserved by rotation. This can also be expressed in the following manner: if π_l is the projection onto the subspace V_l then π_l commutes with rotations:

$$\pi_l(R(f)) = R(\pi_l(f)).$$

- **V_l is Irreducible**: V_l cannot be further decomposed as the direct sum $V_l = V_l' \oplus V_l''$ where V_l' and V_l'' are also (non-trivial) representations of the rotation group. That is, if f is any function of frequency l, then *any* other function of frequency l can be expressed as the sum of rotations of f. Hence, it is impossible to partition the space of spherical harmonic functions further, such that rotations act only within these smaller subspaces.

The first property presents a way for decomposing spherical functions into rotationally independent components, while the second property guarantees that in a linear sense, this decomposition is optimal.

3.2. Rotation Invariant Descriptors

Using the properties of spherical harmonics, and the observation that rotating a spherical function does not change its L_2-norm we represent the energies of a spherical function $f(\theta,\phi)$ as:

$$\mathrm{SH}(f) = \{\|f_0(\theta,\phi)\|, \|f_1(\theta,\phi)\|, \ldots\}$$

where the f_l are the frequency components of f:

$$f_l(\theta,\phi) = \pi_l(f) = \sum_{m=-l}^{m=l} a_{lm} Y_l^m(\theta,\phi)$$

(shown in steps (2) and (3) of Figure 1.)

This representation has the property that it is independent of the orientation of the spherical function. To see this we let R be any rotation and we have:

$$\begin{aligned}\mathrm{SH}(R(f)) &= \{\|\pi_0(R(f))\|, \|\pi_1(R(f))\|, \ldots\} \\ &= \{\|R(\pi_0(f))\|, \|R(\pi_1(f))\|, \ldots\} \\ &= \{\|\pi_0(f)\|, \|\pi_1(f)\|, \ldots\} = \mathrm{SH}(f)\end{aligned}$$

so that applying a rotation to a spherical function f does not change its energy representation.

4. Properties of the Spherical Harmonic Representation

This section provides a mathematical analysis of some of the properties and limitations of the Spherical Harmonic Representation. In particular, we describe how the similarity of spherical descriptors, defined as the optimum over all rotations, relates to the similarity of their Spherical Harmonic Representations. We also describe the way in which information is lost in going from a spherical shape descriptor to its Spherical Harmonic Representation.

1. **Similarity**: The L_2-difference between the Spherical Harmonic Representations of two spherical functions is a lower bound for the minimum of the L_2-difference between the two functions, taken over all possible orientations. To see this, we let $f(\theta,\phi)$ and $g(\theta,\phi)$ be two spherical functions, and observe that:

$$\begin{aligned}\|\mathrm{SH}(f) - \mathrm{SH}(g)\|^2 &= \sum_{l=0}^{\infty} (\|f_l\| - \|g_l\|)^2 \\ &\leq \sum_{l=0}^{\infty} (\|f_l - g_l\|)^2 = \|f(\theta,\phi) - g(\theta,\phi)\|^2.\end{aligned}$$

But as we have shown, the Spherical Harmonic Representation is invariant to rotation, so we get:

$$\|\mathrm{SH}(f) - \mathrm{SH}(g)\| \leq \min_{R \in SO(3)} \|f - R(g)\|.$$

2. **Information Loss**: If a spherical function $f(\theta,\phi)$ is band-limited with bandwidth b, then we can express f as:

$$f(\theta,\phi) = \sum_{l=0}^{b} \sum_{m=-l}^{l} a_{lm} Y_l^m(\theta,\phi).$$

Thus, the space of spherical functions with bandwidth b is of dimension $O(b^2)$. The Spherical Harmonic Representation, however, is of dimension $O(b)$ so that a full dimension of information is lost in going from a spherical function to its representation. This information loss happens in two different ways:

- First, we treat the different frequency components independently. Thus if we write:

$$f = \sum_{l=0}^{b} f_l \qquad \text{and} \qquad g = \sum_{l=0}^{b} R_l(f_l)$$

where R_l are rotations, then the representations of the functions f and g will be the same. That is, the representation is unchanged if we apply different rotations to the different frequency components of a spherical function. Figure 2 shows a visualization of this for two spherical functions. The one on the bottom is obtained from the one on the top by applying a rotation to the rightmost frequency component. Though the two functions differ by more than a single rotation, their spherical harmonic representations are the same. (An analogous form of information loss occurs with Fourier Descriptors where the phases of different frequencies are discarded independently.)
- Second, for each frequency component f_l, the Spherical Harmonic Representation only stores the energy in that component. For $l \geq 2$ it is *not* true that if $\|f_l\| = \|g_l\|$ then there is a rotation R such that $R(f_l) = g_l$. Thus knowing only the norm of the l-th frequency component does not provide enough information to reconstruct the component up to rotation. (This form of information loss does not occur with Fourier Descriptors, as two circular functions with the same amplitude and frequency can only differ by a rotation.)

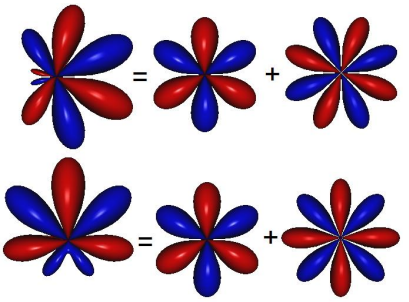

Figure 2: *The bottom spherical function is obtained by rotating the rightmost frequency component of the one on top. Although there is no rotation transforming the function on the top to the one on the bottom, their representations are the same.*

Figure 3 shows a visualization of this for three spherical functions. The functions are all of the same frequency and have the same amplitude, but there is no rotation that can be applied to transform them into each other. In Appendix B we show how this information loss can be resolved for the second order frequency components.

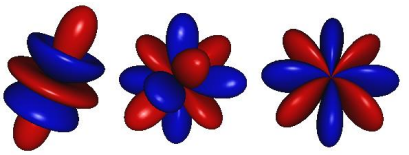

Figure 3: *These three single-frequency ($l = 4$) spherical functions are not related by rotation but have the same spherical harmonic representation.*

5. Extensions to Voxel Grids

In Section 3, we described a method for obtaining rotation invariant representations of spherical functions. In this section, we show how this method can be generalized to obtain rotation invariant representations of regularly sampled 3D functions (voxel grids).

5.1. Rotation Invariant Representations

In order to obtain a rotation invariant representation of a voxel grid we use the observation that rotations fix the distance of a point from the origin. Thus, we can restrict the voxel grid to concentric spheres of different radii, and represent the voxel grid as a collection of spherical functions. Each function is weighted by the radius of the restricting sphere to account for the different areas, and we obtain the Spherical Harmonic Representation of each spherical restriction independently.

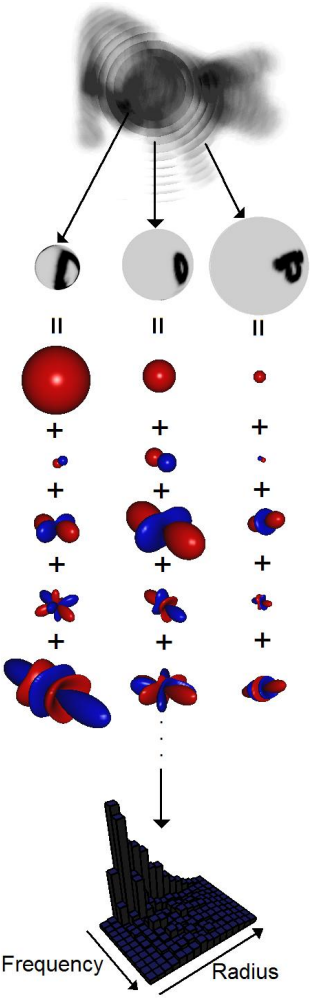

Figure 4: *We compute a rotation invariant descriptor of a voxel grid by intersecting the model with concentric spheres (with darker points corresponding to larger voxel values), computing the frequency decomposition of each spherical function, and computing the norms of each frequency component at each radius. The resultant rotation invariant representation is a 2D grid indexed by radius and frequency.*

This process is demonstrated in Figure 4. First, we restrict the voxel grid to a collection of concentric spheres. Then, we represent each spherical restriction in terms of its frequency decomposition. Finally, we compute the norm of each frequency component, at each radius. The resultant rotation invariant representation is a 2D grid indexed by radius and frequency.

5.2. Properties

In addition to the information loss described in Section 4, the method described above also loses information as a result of the fact that the representation is invariant to independent rotations of the different spherical functions. For example, the plane in Figure 5 (right) is obtained from the one on the left by applying a rotation to the interior part of the model. While the two models are not rotations of each other, the descriptors obtained are the same.

Figure 5: *The model on the right is obtained by applying a rotation to the interior part of the model on the left. While the models are not rotations of each other, their Spherical Harmonic Representations are the same.*

6. Experimental Results

To measure the efficacy of the Spherical Harmonic Representation in tasks of shape analysis, we computed a number of spherical shape descriptors, and compared matching results when the descriptors were aligned by PCA with the results obtained when the Spherical Harmonic Representation was used. The descriptors we used in our experiments were:

- **Extended Gaussian Image**[10]: A description of a surface obtained by binning surface normals.
- **Radial Distribution**: A description of a surface that associates to every ray through the origin, the average distance and standard deviation of points on the intersection of the surface with the ray. (If the intersection is empty then the associated values are zero.)
- **Spherical Extent Function**[12]: A description of a surface associating to each ray from the origin, the value equal to the distance to the last point of intersection of the model with the ray. (If the intersection is empty then the associated value is zero.)
- **Sectors**: A description of a surface associating to each ray from the origin, the amount of surface area that sits over it. This is a continuous implementation of the sectors in Shape Histograms[11], with sectors chosen to correspond to a single cell within the 64×64 representation of the sphere.
- **Shape Histogram**[11]: A finer resolution of the Sector descriptor that breaks up the bounding sphere of the model into a collection of shells and computes the sector descriptor for the intersection of the model with each one.
- **Voxel**[16]: A description of a shape as a voxel grid, where the value at each point is given by the negatively exponentiated Euclidean Distance Transform of the surface.

We evaluated the performance of each method by measuring how well they classified models within a test database. The database consisted of 1890 “household" objects provided by Viewpoint[20]. The objects were clustered into 85 classes, based on functional similarity, largely following the groupings provided by Viewpoint and classes ranged in size from 5 models to 153 models, with 610 models that did not fit into any meaningful classes[16]. Classification performance was measured using precision/recall plots, which give the percentage of retrieved information that is relevant as a function of the percentage of relevant information retrieved. That is, for each target model in class C and any number K of top matches, “recall” represents the ratio of models in class C returned within the top K matches, while “precision” indicates the ratio of the top K matches that are in class C. Thus, plots that appear shifted up and to the right generally indicate superior retrieval results.

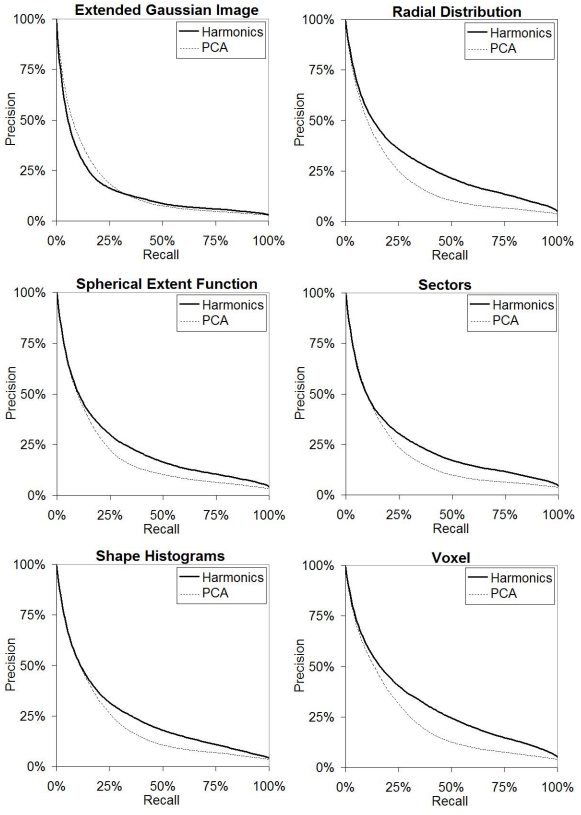

Figure 6: *Precision vs. Recall plots comparing the performance of aligned spherical descriptors with the performance of their Spherical Harmonic Representations. Note that for most of the representations the Spherical Harmonic Representation outperforms the canonically aligned one.*

We computed the spherical representations as 64×64 grids corresponding to regular sampling along the lines of longitude and latitude and we used SpharmonicKit 2.5[21] to obtain the Spherical Harmonic Representation as an array of 32 floating point numbers. Both the spherical descriptors and their Spherical Harmonic Representations were compared using the L_2-difference. The results of the classification experiment are show in Figure 6.

As the results indicate, the application of the Spherical Harmonic Representation improves the performance of most of the descriptors. The improvement of the matching results is particularly meaningful when we consider the fact that the Spherical Harmonic Representation reduces a 2D descriptor into a 1D array of energy values. Thus, the representation not only provides better performance, but it does so with fewer bits of information.

7. Discussion

In this section, we present a discussion of the results in Section 6. In particular, we analyze the case of the Extended Gaussian Image, and discuss how this reflects on the general limitations of the Spherical Harmonic Representation. We also evaluate the implications of the Spherical Harmonic Representation for database retrieval and discuss the dependency of the representation on the center of mass.

7.1. Limitations

The analysis described in Appendix A provides a mathematical interpretation of the failing of PCA-alignment. This analysis makes the assumption that we are looking at the general class of spherical functions, so that frequency components align independently. However, in certain shape applications this may not be the case and the descriptors obtained may fall into a restrictive subset of spherical functions. In these cases, it is possible that the alignment of different frequency components are correlated and PCA alignment performs well.

Such a case may occur when the spherical functions are primarily axis aligned, so that up to rotation they can be described as:

$$\sum a_k x^k + b_k y^k + c_k z^k$$

and the alignments of the different frequency components are strongly correlated. This is the case for the Extended Gaussian Image[10] which describes a polygonal model by the distribution of normal vectors over the unit sphere. When the database of models is restricted to household objects, the obtained descriptors are primarily axis aligned (see Figure 7) and principal axis alignment may provide optimal alignment, (as indicated by the improved performance in Figure 6).

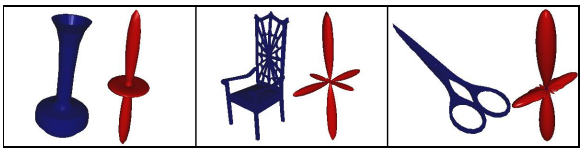

Figure 7: *Images of models of a vase, a chair, and scissors, with their associated Extended Gaussian Images. Note that the EGIs are mainly axial functions and consequently are well aligned by PCA.*

7.2. Implications for Model Databases

Much of the research presented in this paper is guided by the increased proliferation and accessibility of 3D models. These models have been gathered into databases, and one of the challenges has been to design matching implementations that are well suited for database retrieval. The Spherical Harmonic Representation described in this paper addresses this challenge in two ways:

1. While a spherical function of bandwidth b requires $O(b^2)$ space, its Spherical Harmonic Representation is of size $O(b)$. Consequently, the Spherical Harmonic Representations provide a more compact representation of the descriptors, and can be compared more efficiently. (For each method compared in Section 6, Table 2 shows the space requirements of the descriptor and its Spherical Harmonic Representation.)
2. Furthermore, the Spherical Harmonic Representations are based on a frequency decomposition of a spherical function. Consequently, the representation is inherently multiresolutional and this property can be used to guide indexing schemes for efficient retrieval.

Representation	PCA-Aligned	Harmonic
EGI	64×64	32
Spherical Extent Function	64×64	32
Radial Distribution	$2 \times 64 \times 64$	2×32
Sectors	64×64	32
Shape Histogram	$4 \times 64 \times 64$	4×32
Voxel	$32 \times 64 \times 64$	32×32

Table 2: *The number of floating point numbers used to describe each representation. This table demonstrates that the Spherical Harmonic Representation provides a representation that reduces the dimensionality of the space required for storing the descriptor.*

We have taken advantage of these properties of the Spherical Harmonic Representation in the design of a web-based 3D model retrieval system[22]. The system indexes 36,000 3D models using the Voxel descriptor, and performs a query, returning the top 100 results, in under one second. The system

has been publicly available for two years and has seen an average of 550 searches per day. For a full description of the system we refer the reader to[23].

7.3. Translation Invariance

In this paper we have addressed the challenge of matching 3D shapes across different similarity transformations by providing a method for obtaining a rotation invariant representation of many existing shape descriptors. The challenge of matching across different translations is addressed by translating the shape so that its center of mass is at the origin before computing the descriptors. This approach is motivated by the results of Horn et al.[17], which show that for matching pairs of ordered point sets, translation to the center of mass always provides the optimal matching (independent of the order of the point sets).

However, it should be stressed that such an approach only generalizes in the context of whole-object to whole-object matching, where an underlying correspondence between points on the two models can be assumed. In the context of partial-object to whole-object or partial-object to partial-object matching the assumption of complete correspondence fails to hold, and center of mass translation cannot be expected to provide optimal alignment.

An alternative approach to normalizing for translation is providing a representation that is translation invariant. For voxel grids, the phase elimination approach used for obtaining a rotation invariant representation can also be used to obtain a translation invariant representation. In particular, if we consider a voxel grid as the sampling of a real valued function f, then we can express f in terms of its Fourier decomposition:

$$f(v) = \int_{\mathbf{R}^3} \hat{f}(w) e^{i\langle v,w\rangle} dw$$

where $\hat{f}(w)$ are the Fourier coefficients of f. The function $|\hat{f}|$ is a translation invariant representation of f that has the following two properties:

1. For two functions f and g, we have
$$\left\| f-g \right\|_2 \geq \left\| |\hat{f}| - |\hat{g}| \right\|_2$$
so that the L_2-difference of the translation invariant representations of two functions is a lower bound for the L_2-difference of the functions.
2. If R is any rotation and $g(v) = f(R(v))$ then we have:
$$|\hat{g}(w)| = |\hat{f}(R(w))|$$
so that applying a rotation to a function f amounts to applying the same rotation to its translation invariant representation.

Consequently, if we apply the methods of Section 5 to the function $|\hat{f}|$, we obtain a translation and rotation invariant representation of the function f that satisfies the L_2 lower bound property described in Section 4.

8. Conclusion and Future Work

In this paper we have described the Spherical Harmonic Representation, a rotation invariant representation of spherical functions in terms of the energies at different frequencies. We have presented a mathematical analysis of its properties and demonstrated its efficacy in shape matching by empirically showing that it provides a more concise and robust representation for many existing shape descriptors.

In the future, we would like to explore extensions of the Spherical Harmonic Representation that completely define a spherical function up to rotation. That is, we would like to consider methods for eliminating the rotation independence of the different frequency components, and we would like to capture enough information in the higher frequency components to reconstruct them up to rotation. We would also like to consider the implications and limitations of extending this method to obtain translation and scale invariant representations. These extensions would not depend on translation and scale normalization, and would be better suited for partial-object matching tasks.

References

1. Burel, G., Henocq, H.: Three-dimensional invariants and their application to object recognition. Signal Processing **45** (1995) 1–22
2. Lo, C., Don, H.: 3D moment forms: Their construction and application to object identification and positioning. IEEE PAMI **11** (1989) 1053–1064
3. Makadia, A., Daniildis, K.: Direct 3D-rotation estimation from spherical images via a generalized shift theorem. IEEE CVPR (2003)
4. Besl, P., McKay, N.: A method for registration of 3D shapes. IEEE PAMI **14** (1992) 239–256
5. Zhang, Z.: Iterative point matching for registration of free-form curves and surfaces. IJCV **13** (1994) 119–152
6. Ballard, D.: Generalized Hough Transform to detect arbitrary patterns. IEEE PAMI **12** (1981) 111–122
7. Lamdan, Y., Wolfson, H.: Geometric hashing: A general and efficient model-based recognition scheme. ICCV (1988) 238–249
8. Besl, P.: Triangles as a primary representation. Object Recognition in Computer Vision **LNCS 994** (1995) 191–206
9. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Matching 3D models with shape distributions. Shape Matching International (2001) 154–166
10. B. Horn, B.: Extended Gaussian images. IEEE **72** (1984) 1656–1678
11. Ankerst, M., Kastenmüller, G., Kriegel, H.P., Seidl, T.: 3D shape histograms for similarity search and classification in spatial databases. In: SSD. Volume 1651. (1999)

12. Vranic, D., Saupe, D.: 3D model retrieval with spherical harmonics and moments. Proceedings of the DAGM (2001) 392–397

13. Gain, J., Scott, J.: Fast polygon mesh querying by example. SIGGRAPH Technical Sketches (1999) 241

14. Kazhdan, M., Chazelle, B., Dobkin, D., Finkelstein, A., Funkhouser, T.: A reflective symmetry descriptor. ECCV (2002) 642–656

15. Elad, M., Tal, A., Ar, S.: Content based retrieval of VRML objects - an iterative and interactive approach. EG Multimedia (2001) 97–108

16. Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., Dobkin, D., Jacobs, D.: A search engine for 3D models. ACM TOG **22** (2003) 83–105

17. Horn, B., Hilden, H., Negahdaripour, S.: Closed form solution of absolute orientation using orthonormal matrices. Journal of the Optical Society **4** (1988) 1127–1135

18. Zahn, C., Roskies, R.: Fourier descriptors for plane closed curves. IEEE Transaction on Computers **21** (1972) 269–281

19. Kazhdan, M., Funkhouser, T.: Harmonic 3D shape matching. SIGGRAPH Sketches and Applications (2002)

20. Viewpoint Data Labs: `http://www.viewpoint.com` (2001)

21. SpharmonicKit 2.5: `http://www.cs.dartmouth.edu/~geelong/sphere/` (1998)

22. Princeton 3D Model Search Engine: `http://shape.cs.princeton.edu/` (2001)

23. Min, P., Halderman, A., Kazhdan, M., Funkhouser, T.: Early experiences with a 3D model search engine. Proc. Web3D Symposium (2003) 7–18

Appendix A: A Signal Processing Framework for PCA

This appendix presents a signal processing framework for analyzing the implications and limitations of model alignment via PCA. In particular, we show that PCA-alignment only considers second order shape information and consequently does not guarantee optimal alignment at other frequencies.

To do this, we define a spherical function characterizing the radial variance of a shape along different rays from the origin. In particular, for a model S and a direction v we set:

$$\mathrm{RV}(S,v) = \lim_{\alpha\to 0} \int_{C(v,\alpha)\cap S} \frac{\|x\|^2}{2\pi(1-\cos(\alpha))}\, dA$$

where $C(v,\alpha)$ is the cone with apex at the origin, angle α and direction v, and $2\pi(1-\cos(\alpha))$ is the area of the intersection of the cone with the unit sphere (Figure 8 (left)). That is, $\mathrm{RV}(S,v)$ gives the sum of the square of the distances of the points lying on the intersection with S and the ray, from the origin, with direction v. A visualization of this function for a cube is shown in Figure 8 on the right. Note that the function

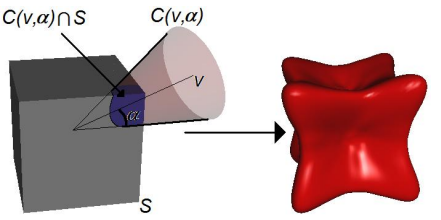

Figure 8: *The value of the* Radial Variance *in the direction v is defined by intersecting the model with a cone, in the direction v, with small angle α, and integrating the square of the distance over the intersection of the model with the cone (left). The resultant spherical function is shown on the right by scaling points on the sphere in proportion to their value.*

scales the points at the corners of the cube more drastically because: (1) we integrate the *square* of the distance to the origin over each patch, and (2) the angle between the point on the sphere and the surface normal is large, so that more surface area projects onto a spherical patch.

What is valuable about this function is that for any surface S, the function has the property that:

$$M^S_{i,j} = \int_S x_i x_j\, dA = \langle \mathrm{RV}(S,v), x_i x_j \rangle_{S^2}$$

where M^S is the covariance matrix of the shape S. That is, the second (and 0-th) order components of the radial variance function completely define the principal axes of the model. This function gives a representation of the initial model in a signal processing framework that allows us to make two observations:

1. Because of the orthogonality of the frequency components, principal axis alignment does not take into account information at non second-order frequencies and hence makes no guarantees as to how they align.
2. Aligning two models using their principal axes provides the optimal alignment for their second order components, as will be shown in the following theorem:

Theorem: If f and g are two spherical functions consisting of only constant and second order harmonics, then the L_2-difference between the two is minimized when each is aligned to its own principal axes.

Proof: Because f and g consist of only constant and second order terms, we can represent the functions by symmetric matrices A and B where

$$f(v) = v^t A v \qquad \text{and} \qquad g(v) = v^t B v.$$

If we assume that A and B are already aligned to their principal axes we get:

$$A = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{pmatrix}$$

Thus, if R is any rotation we get:

$$\langle R^t(f), g\rangle = (\alpha-\beta)\text{Trace}(ARBR^t) + \beta \sum_{i,j=1}^{3} a_i b_j$$

where $\alpha = \int_{S^2} x^4 dx$ and $\beta = \int_{S^2} x^2 y^2 dx$ define the lengths and angles between the functions x_i^2 on the unit sphere. We would like to show that the dot product is maximized when R is a permutation matrix so that RAR^t is diagonal.

Using the fact that the differentials of a rotation R are defined by RS where S is a skew-symmetric matrix, it suffices to solve for:

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{t=0} \text{Trace}(A(R+tRS)B(R^t - tSR^t)) \\ &= \text{Trace}(R^t AR(SB-BS)) \end{aligned}$$

But S is a skew-symmetric matrix so that, $SB-BS$ is a symmetric matrix with 0's along the diagonal:

$$SB-BS = \begin{pmatrix} 0 & (b_2-b_1)S_{12} & (b_3-b_1)S_{13} \\ (b_2-b_1)S_{12} & 0 & (b_3-b_2)S_{23} \\ (b_3-b_1)S_{13} & (b_3-b_2)S_{23} & 0 \end{pmatrix}$$

Thus, if R^tAR is a diagonal matrix then the derivative is zero, independent of the choice of S. Conversely, if the b_i are distinct and R^tAR is not diagonal, we can always choose values for S_{12}, S_{13}, and S_{23} such that the derivative is non-zero, implying that if R^tAR is not diagonal it cannot maximize the dot product. (Note that if $b1 = b2 = b3$ then B is a constant multiple of the identity so that the dot product is independent of the choice of rotation. Similarly, if $b_i = b_j$ then rotations in the plane spanned by x_i and x_j also do not change the dot product.)

This shows that the L_2-difference between f and g is at an extremum if and only if A and B are diagonal matrices. The minimum L_2-difference is then attained when $\sum a_i b_i$ is maximal. So, if $a_1 \geq a_2 \geq a_3$ then we must also have $b_1 \geq b_2 \geq b_3$, and the L_2-difference between f and g is minimized precisely when f and g are aligned to their principal axes.

Appendix B: Further Quadratic Invariants

One of the limitations of only storing the energies at the different frequencies of a spherical function is that it does not allow us to reconstruct the frequency components uniquely, up to rotation. In the past, this problem has been addressed by using *algebraic* methods to obtain additional rotation invariants for the different frequency components[1, 2, 3]. The difficulty with these approaches is that the derived 0-th order tensor invariants are often redundant and consequently, cannot be directly compared to obtain a lower bound for the minimum L_2-difference between two spherical functions.

In this appendix, we present a new *geometric* approach for computing orthogonal invariants and describe an implementation for the quadratic components of a spherical function. This approach is based on the idea that for a spherical function f, we can generate the manifold M_f, defined as the image of f under all rotations. For two spherical functions f and g, the minimum difference between f and g, taken over the space of all rotations, is precisely the distance between the two manifolds M_f and M_g, (Figure 9). The goal then, is to be able to index these manifolds in such a way that the L_2-difference between two sets of indices is exactly the distance between the two manifolds.

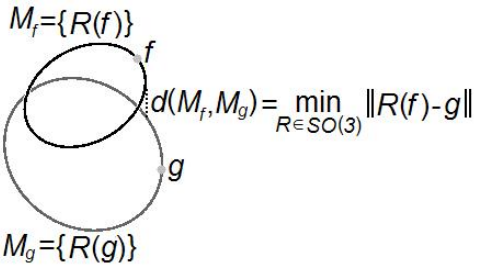

Figure 9: *The minimum distance between two functions f and g, taken over the space of all rotations, is equal to the distance between the two manifolds M_f and M_g, where M_f and M_g are obtained by applying all rotations to the functions f and g.*

Using the results from Appendix A we know that the L_2-difference between the quadratic components of two spherical functions is minimized when the two functions are aligned with their principal axes. Thus, we can represent the constant and quadratic components by the three scalars $a_1 \leq a_2 \leq a_3$, where after alignment to principal axes:

$$f_0 + f_2 = a_1 x^2 + a_2 y^2 + a_3 z^2.$$

The indices (a_1, a_2, a_3) uniquely define the constant and quadratic function up to rotation, but because the functions $\{x^2, y^2, z^2\}$ are not orthogonal, they do not have the property that the L_2-difference between two sets of indices is the minimum of the L_2-distance between the two functions. To address this, we fix an orthonormal basis $\{v_1, v_2, v_3\}$ for the span of $\{x^2, y^2, z^2\}$ and represent the function $f = f_0 + f_2$ by the three scalars $R^{-1}(a_1, a_2, a_3)$, where R is the matrix whose columns are the orthonormal vectors v_i. This provides the desired orthogonal indexing for the constant and quadratic components of a spherical function, which define the components uniquely, up to rotation.