MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

A 3D Shape Descriptor Based on Hadamard Transform and Spherical Harmonic Transform

Guyu Xin
State Key Laboratory on Machine Perception,
Peking University
xingy@cis.pku.edu.cn

Hongbin Zha
State Key Laboratory on Machine Perception,
Peking University
zha@cis.pku.edu.cn

Abstract

Shape matching plays an important role in many application fields. In this paper, we propose a novel rotation invariant 3D shape descriptor based on Hadamard transform and spherical harmonic transform. In our method, a 3D model is represented as a collection of spherical functions to preserve shape information as much as possible, and the shape similarity is directly defined by the difference of the character functions of 3D models. Retrieval experiments show that our method performs better than many other existing 3D matching methods based on spherical harmonics.

1. Introduction

Over the past decade, the problem of 3D shape matching has been intensively studied in computer vision research due to its relevance to a variety of application fields. These include, among others, robot vision, autonomous navigation, automated inspection and measurement, virtual reality, and retrieval of 3D objects based on shapes. The most common idea of matching 3D shapes is to establish correspondences between the query and target models [1], and then to define similarity measure in terms of distances between corresponding points. Unfortunately, however, the establishing of correspondences is a difficult and time-consuming task that needs to be performed on a model-pair-wise basis. This has motivated a large body of research in the area of shape descriptors: the space of models is mapped into a vector space with fixed dimensions [2, 9, 11], and the measure of similarity between two models is defined as the distance between their corresponding descriptors. This approach has the advantage of addressing the correspondence problem on a model-wise basis, allowing for the computation of the descriptors in an offline process.

With the studies of using spherical harmonics (SH) to extract rotation invariants of spherical functions and progress in fast discrete spherical harmonic computation [6], many methods first represent a 3D shape as either a spherical function or a voxel grid, and then extract their rotation invariants by spherical harmonic transform [3, 4, 8, 12]. Some primary properties of 3D models are naturally spherical functions [2, 11]. If we use the rotation invariants of these primal properties as shape descriptors, one challenge is that the useful information is limited, and then their discriminabilities of 3D models is not good as expected.

The idea to strengthen the discriminability is to extract several groups of rotation invariants of different spherical functions derived from 3D models so that the descriptor conveys shape information more than those extracted from only one spherical function. Many methods adopt this idea [8, 12]. The problem here is that different components have different effects to the shape. Therefore, it is difficult to determine the weight of each component of the rotation invariants.

Unlike the methods introduced above, in our method, a 3D model is considered equivalent to its character function, and so the dissimilarity of two models can be regarded as the difference of their character functions. The character function value is defined to be one on the model and zero otherwise. After a quantification process, we obtain a set of spherical functions by using Hadamard transform along radials from the origin. It is evident that all the spherical functions in the set contribute evenly to the difference of character functions and they are continuous to the surface perturbation. To enhance the robustness of the spherical harmonic transformation, we use the anisotropic scaling technique to normalize the 3D-model.

In this paper, we present a new 3D shape descriptor based on Hadamard transform and spherical harmonic transform. In this method, a 3D model is firstly represented as a collection of spherical functions to preserve shape information as much as possible, and then rotation invariants are extracted. Retrieval experiment results show that our method performs better than many other existing 3D matching methods which are also based on spherical harmonics.

2. Shape Representation and Similarity

2.1 Model representation by spherical functions

First we review the mathematic definition of Hadamard transform. The Hadamard matrix H_N of order N, where N is a positive integral power of 2, is a symmetric square matrix defined as

$$H_N = \begin{bmatrix} H_{N/2} & H_{N/2} \\ H_{N/2} & -H_{N/2} \end{bmatrix},$$

with $H_1 = [1]$. The Hadamard transform maps a N-dimension array g to another N-dimension array $h = H_N g$. Since the elements of H_N are regularly arranged with only two entries ± 1 , the Hadamard transform can be calculated quickly through the algorithm called Fast Hadamard Transform [7].

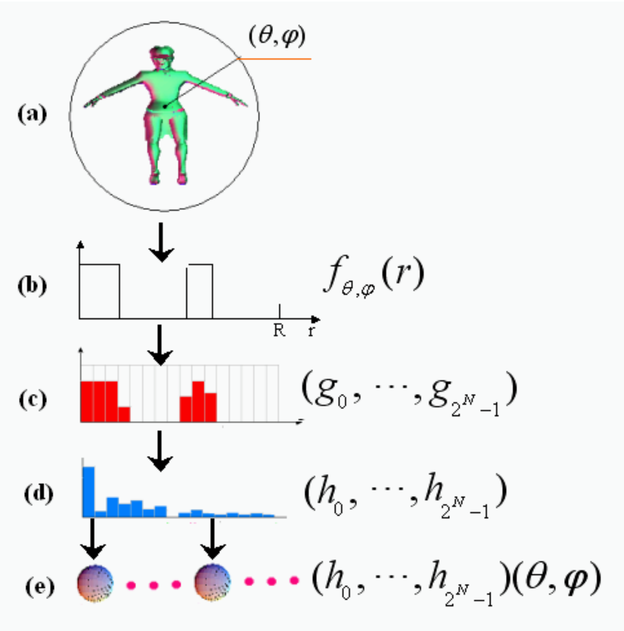

Figure 1. The process of constructing spherical functions. (a) A 3D model in a ball. (b) Rectangle wave function $f_{\theta,\varphi}(r)$ on a casting ray. (c) $f_{\theta,\varphi}(r)$ is digitalized. (d) The Hadamard frequencies of $(g_0,g_1,\cdots,g_{2^N-1})$. (e) All the Hadamard frequencies form a spherical function vector.

A 3D volume model M can be represented by a character function

$$f(P)=\begin{cases}1, & P\in M\cup\partial M,\\ 0, & \text{otherwize.}\end{cases}$$

Then f on a casting ray with the direction (θ,φ) defines a rectangle wave function $f_{\theta,\varphi}(r)$. In practice, a 3D model is volume finite and we can put it in an envelope ball $B(0,R)$ with the radius R. Hence $\operatorname{supp} f_{(\theta,\varphi)}\subset[0,R]$ for all (θ,φ), and $f_{\theta,\varphi}(r)$ can be characterized with an array $g=\{g_0,g_1,\cdots,g_{2^N-1}\}$, where

$$g_k=\int_{kR/2^N}^{(k+1)R/2^N} f_{\theta,\varphi}(r)dr\ .$$

Then its Hadamard transform $h=H_{2^N}g$ forms a spherical function vector as (θ,φ) are taken over all the directions:

$$h(\theta,\varphi)=(h_0(\theta,\varphi),h_1(\theta,\varphi),\cdots,h_{2^N-1}(\theta,\varphi))\ .$$

Figure 1 shows the process of constructing the spherical function vector. The property of Hadamard transform assures that the L_2-distance between two character functions f and f' can be approximately represented the L_2-distance between their spherical function vectors h and h'.

2.2 Rotation invariants and similarity measure

According to the theory of spherical harmonics, a spherical function $f(\theta,\varphi)$ can be decomposed as the sum of its harmonics:

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

where $a_{lm}, l=0,1,\cdots,\infty, m=-l,\cdots,l-1,l,$ are the coefficients in frequency domain. It is proved that the summation $A_l=\sum_{m=-l}^{l}a_{lm}^2$ is a rotational invariant. In practice, the Fast SHT method is used for deriving the coefficients in the first B frequencies by sampling a $2B\times 2B$ grid per latitude and longitude for a spherical function [6]. Given a 3D model M, it can be featured by its rotation invariants:

$$A_{k,l}=\sqrt{\sum_{m=-l}^{l}a_{k,lm}^2}\ ,k=0,1,\cdots,2^N-1,l=0,1,\cdots B\ .$$

Here $a_{k,lm}$ is the spherical harmonics coefficients defined by spherical harmonic decompositions:

$$h_k(\theta,\varphi)=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}a_{k,lm}Y_l^m(\theta,\varphi), k=0,1,\cdots,2^N-1\ ,$$

where $h(\theta,\varphi)=(h_0(\theta,\varphi),h_1(\theta,\varphi),\cdots,h_{2^N-1}(\theta,\varphi))$ are a series of spherical functions of M as defined in section 2.1. For two models M and M', the shape distance between them is measured by L_2-distance of their features:

$$D_f(M,M')=\sum_{k=0}^{2^N}\sum_{l=0}^{B}(A_{k,l}-A'_{k,l})^2\ ,$$

where $A_{k,l}, A'_{k,l}$ are rotation invariants for M and M' respectively.

2.3 Anisotropic normalization

In many applications, a model and its image under a similarity transformation are considered to be the same. Though our method of feature extraction is rotation invariant, we need to normalize the models with respect to the translation and scale before extracting the rotation invariants. In general methods, models are normalized in an object coordinate by using the center of mass for the translation, and the root of the average square radius for the scale. Here we adopt the optimal anisotropic scale normalization [5], which will improve the representation ability of the features and take advantage of assigning importance of anisotropy. A model is decomposed into an anisotropy vector $\lambda_M=(\lambda_1^M,\lambda_2^M,\lambda_3^M)$ and an isotropic model $\tilde{M}$ by anisotropic scaling. We compute the feature vector $v_{\tilde{M}}$ of the isotropic model $\tilde{M}$. As shown in figure 2, the model M can be represented by a new feature vector $\{v_{\tilde{M}}, \lambda_M\}$. Then we can define the measure of similarity between two models Q and M as:

$$D_\gamma(M,Q) = \left\|v_{\tilde{M}}\right\|^2 + \left\|v_{\tilde{Q}}\right\|^2 - 2(v_{\tilde{M}} \cdot v_{\tilde{Q}})(\lambda_M \cdot \lambda_Q)^\gamma .$$

If $\gamma=1$, then $D_\gamma(M,Q)$ is the L_2-difference between the vectors $v_{\tilde{M}} \times \lambda_M$ and $v_{\tilde{Q}} \times \lambda_Q$. More generally, γ can be treated as a constant representing the importance of anisotropy information in the context of shape matching.

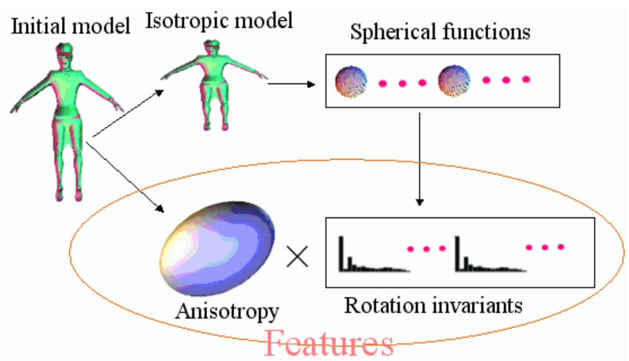

Figure 2. A model is decomposed into an isotropic model and an anisotropy by the anisotropic scaling. The model feature is represented by the anisotropy and the rotation invariants of the isotropic model.

There is still another reason why we make the anisotropy rescaling before extracting rotation invariants. If the model is just scaled by the root of its average square radius, the max distance of its extent to its origin may be very large. To ensure that all the models are enclosed in the repository, the envelope ball need to be very large, and the margin volume between the envelope ball and the convex hull of the model will consume too much storing space in the computation. After the rescaling by anisotropic normalization, almost all the isotropic models in the repository can be enveloped in a ball with radius R=2 excluding some extreme cases (eg. an insect model with very long antennas), and the proportion of the margin space in the envelope ball is not so high as in the case of anisotropic models (as shown in figure 3).

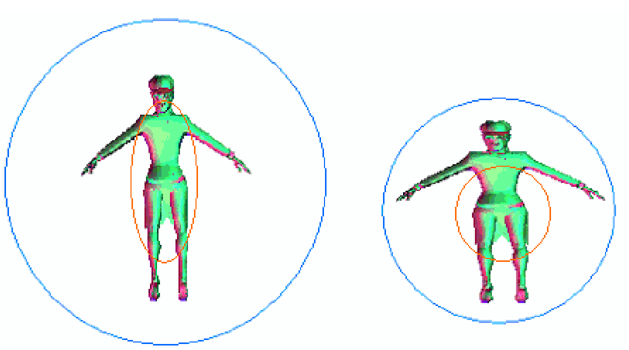

Figure 3. An anisotropic model (left) requires a large envelope ball and a large portion of the envelope ball is margin. The envelope ball of its isotropic model (right) can be chosen smaller.

3. Results

The method is implemented for experiments to show its performance in 3D retrieval. Specifically, every model was normalized in size by isotropically rescaling it so that the average distance from the points on its surface to the center of mass is 0.5. The length of Hadamard transformation is 32. The spherical harmonic transformation is computed on a 64×64 spherical grid, and the first 16 rotation invariants are computed and reserved for our descriptor. As a result, the descriptor is a 16×32 dimensional array from the spherical functions and a triplet of isotropic scalar information. The similarity distance is the L_2-difference between the vectors $v_{\tilde{M}} \times \lambda_M$ and $v_{\tilde{Q}} \times \lambda_Q$ (here $\gamma=1$). In order to test the effectiveness of the proposed method, we designed a series of shape matching experiments on the database provided by the Princeton Shape Benchmark (PSB) [10].

3.1 Results on robustness

In the first experiment, we test the robustness of our dissimilarity measure to transformations and perturbations of 3D models. Specifically, fifty models are randomly selected from the PSB. We design the experiments from three aspects to test the robustness:

1) **Translation:** Translate the origin of the coordinates for a small bias d from the centroid of the 3D model;
2) **Rotation:** Rotate the 3D model by a random angle;
3) **Perturbation:** Perturb each vertex randomly by d along its normal direction.

We compute the relative error of the transformations by

$$\varepsilon = \frac{\|v - v'\|}{\|v\|},$$

where v and v' are the shape descriptors of the original model and the transformed model respectively. Let's denote the average value and max value of ε as $\varepsilon_{ave}, \varepsilon_{\max}$ respectively. The test results are shown in Table 1.

Table 1. Results for the robustness testing

Tests	Results			
Translation	d	0.01	0.02	0.05
	ε_{ave}	0.0138	0.0354	0.0863
Rotation	Error	ε_{ave}	0.0907	
		$\varepsilon_{\max}$	0.1540	
Perturbation	d	0.01	0.02	0.03
	ε_{ave}	0.0084	0.0193	0.0463

3.2 Retrieval results

In order to evaluate the performance of the shape similarity measure, we design experiments performing 3D model retrieval on the PSB database. Figure 4 shows some of the retrieval examples. We randomly select a model in the database as the query, and then our system returns a list of outputs ranking on the degree of similarity to the input (because the query model in each retrieval test is still from the PSB database, the most similar model is itself). We can find that the first several models in the retrieval list are really shape-like to the query model in most cases. Especially, many classes of 3D models with complex structures, such as potted plants, bicycles, etc, have good retrieval results by using our retrieval scheme.

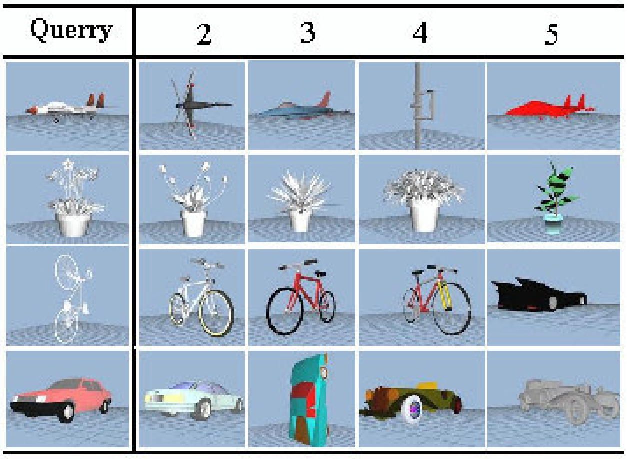

Figure 4. Some retrieval examples.

We compared the classification performance of the proposed method with many other previous methods based on spherical harmonics. 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. Precision-recall plots of five methods are shown in Figure 5. We can see that the proposed shape descriptors (HSH) performs better than the Extended Gaussian Image (EGI) [2], the Spherical Extent Function (EXT) [11], the Robust definition of Spherical functions (RSF) [8], and the Radialized Spherical Extent Functions (REXT) [12].

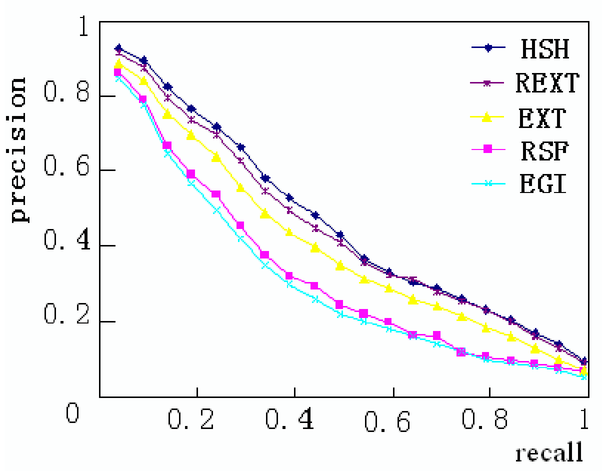

Figure 5. Precision-recall curves of some shape descriptors on PSB database.

4. Conclusion

In this paper, we present a novel 3D shape descriptor based on Hadamard transform and spherical harmonic transform. The similarity measure of the descriptor is derived directly from the L_2-distance of character function of 3D models. The proposed descriptor is robust and rotation invariant. It can describe 3D models of complex structures. The anisotropic scaling normalization before extracting rotation invariants can improve the shape representation ability of the proposed descriptor.

References

[1] P. Besl and N. Mckay: "A Method for Registration of 3-d Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.14, no.2, pp.239-256, 1992.

[2] B. Horn: "Extended Gaussian Images," *Proceedings of the IEEE*, vol.72, no.12, pp.1671-1686, 1984.

[3] M. Kazhdan, B. Chazelle, D.P. Cobkin, et al: "A Reflective Symmetry Descriptor," *Proceedings of European Conference on Computer Vision*, pp.642-656, 2002.

[4] M. Kazhdan and T. Funkhouser: "Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors," *Proceedings of Symposium on Geometry processing*, pp.156-165, 2003.

[5] M. Kazhdan, T. Funkhouser, S. Rusinkiewicz: "Shape Matching and Anisotropy," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 623–629, 2004.

[6] P.J. Kostelec, D.K. Maslen, et al: "Computational Harmonic Analysis for Tensor Fields on the Two-sphere," *Journal of Computational Physics*, vol.162, no.2, pp.515-534, 2000.

[7] M.H. Lee and M. Kaveh: "Fast Hadamard Transform on a Simple Matrix Factorization," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol.34, no.6, pp.1666-1667, 1986.

[8] Y. Liu and J. Pu, G. Xin, H. Zha, et al: "A Robust Method for Shape-based 3D Model Retrieval," *Proceedings of the 12th Pacific Conference on Computer Graphics and Applications*, pp.3-9, 2004.

[9] R. Osada, T. Funkhouser, B. Chazelle, D. Dobkin: "Matching 3d Models with Shape Distributions," *Proceedings of the International conference on Shape Modeling and Applications*, pp.154-166, 2001.

[10] Princeton Shape Benchmark (PSB): http: //shape.cs.princeton.edu/benchmark/

[11] D. Saupe and D. V. Vranic: "3D Model Retrieval with Spherical Harmonics and Moments," *Proceedings of the 23rd DAGM-Symposium on Pattern Recognition*, pp. 392-397, 2001.

[12] D. V. Vranic: "An Improvement of Rotation Invariant 3D-shape Descriptor Based on Functions on Concentric Spheres," *IEEE International Conference on Image Processing*, pp.757-760, 2003.