

# Efficient Similarity Retrieval for Temporal Shape Sequences: A Case Study using Cardiac MR Images

Heng Huang, Fillia Makedon, James Ford, Li Shen, Yuhang Wang, Tilmann Steinberg  
Dartmouth Experimental Visualization Laboratory (DEVLAB)  
Department of Computer Science

Ling Gao, Justin Pearlman  
Dartmouth Medical School  
Dartmouth College, Hanover, NH 03755

{hh, makedon, jford, li, wyh, tilmann}@cs.dartmouth.edu, {Ling.Gao, Justin.D.Pearlman}@dartmouth.edu

## Abstract

The spherical harmonics (SPHARM) approach has been used for the representation of shapes in many types of biomedical image data. We propose a SPHARM-based similarity comparison for shape sequences that allows fast similarity searches for dynamic objects and demonstrate it using 3D images of a beating heart. By using spherical harmonics to extract a small number of features that represent cardiac shape in each sequential state, we enable indexing and pruning of database entries with a multidimensional index tree (e.g. R\*-tree) for fast retrieval. Our approach relies on obtaining selected landmarks to allow normalization within and between sequences. This framework is extensible to other application domains.

## 1. Introduction

As the volume of digitized information objects is rapidly growing, there is a rising demand for effective image searching tools based on similarity search. Shape is a widely used visual feature to describe image content [11] and image change. Identifying small shape changes in an image is of importance in many applications, such as medical imaging, to denote pathology, security, to denote object tracking, biology, to denote molecular changes, and other applications. In spite of the importance of change in recognizing an image, a complete mathematical model of how humans perceive shapes and detect shape similarities is lacking. It would be desirable to have a model that incorporates many different shape measures for shape based information retrieval [4].

Three dimensional shape models are useful for many purposes; in medicine, these include deriving functional in-

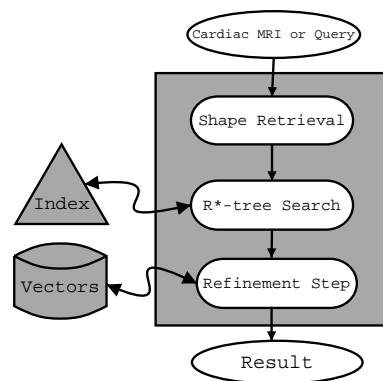


Figure 1. A shape-based similarity retrieval framework.

formation and classifying different pathological symptoms. Our own work has focused on relating the shape sequences of an active heart to pathologies and the effects of therapies. We designed our approach around the need for fast retrieval (Fig.1) for similar shape sequences, with the potential for extension to disease classification based on heart dynamics.

## 2 Shape Retrieval for Cardiac MRI

Cardiac MRI captures 3D images of the heart during its normal operation, with acquisition timed according to heartbeat frequency so that a fixed number of images are acquired during each heartbeat. In this work imaging was performed on a 1.5 Tesla scanner (Genesis Signa, GE Medical systems) with flip angle  $20^\circ$  and slice thickness of 5 mm, producing sequences of cardiac images in DICOM format. Each sequence consists of 20 volume images that together represent one complete heartbeat cycle.

## 2.1 Spherical harmonics (SPHARM) descriptors

Spherical harmonics functions have certain mathematical properties that make them attractive for surface modeling [6, 9]: orthogonality, completeness, and ordering in spatial frequency. Since the shapes of heart and ventricular are close to ellipsoid, spherical harmonics descriptors are reasonably used to describe their shapes [8].

We adopt the SPHARM expansion technique [3] to create the shape description for cardiac 3D surfaces. In this section we give a summary of this method. Since the operation of the main cardiac chamber, known as the left ventricle, is critical, we apply SPHARM analysis to both the external cardiac surface and the surface of the left ventricle.

The two cardiac surfaces  $r(\theta, \phi)$  are expressed [10] as a linear combination of spherical harmonic basis functions  $Y_l^m(\theta, \phi)$  of varying order  $l$  and degree  $m$  (see the visualization examples of our experiments in Fig. 2):

$$r(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi),$$

where

$$c_l^m = (c_{xl}^m, c_{yl}^m, c_{zl}^m)^T,$$

$\theta \in [0, \pi]$  is the polar angle and  $\phi \in [0, 2\pi)$  is the azimuthal angle. The spherical harmonic basis functions are defined as:

$$Y_l^m(\theta, \phi) \equiv \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\phi},$$

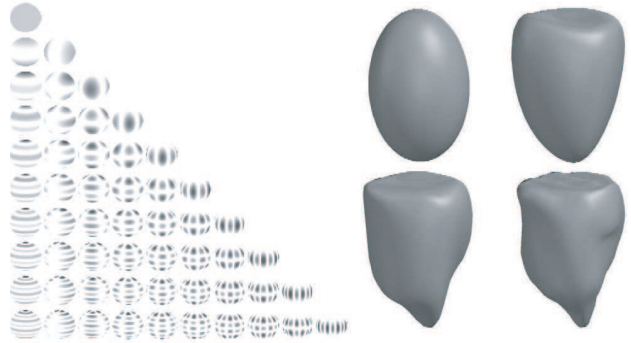
where  $P_l^m(\cos\theta)$  are associated Legendre polynomials (with argument  $\cos\theta$ ) that is defined by the differential equation

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \frac{d^{m+l}}{dx^{m+l}} (x^2-1)^l,$$

where  $l$  and  $m$  are integers with  $-l \leq m \leq l$ .

## 2.2. Normalization

One common problem is that shapes are often described in arbitrary units of measurement and in unpredictable positions and orientations in 3D space. As a prerequisite for any shape similarity calculation, a normalization step is necessary in order to transform each shape into a canonical coordinate system. Using this procedure, we can compare shapes with different translation, rotation (orientation), and scaling of an original model. In our cardiac application, the normalization step serves two purposes. First, it ensures that different 3D cardiac shapes will be comparable regardless of their acquisition (short-axis, long-axis, or other scan orientation). Second, it compensates for size differences and for the translational and rotational motion that occur during the heartbeat cycle, ensuring that different hearts (with different sizes and translational and rotational characteristics) will remain in alignment at every stage when their sequences are compared.



**Figure 2.** SPHARM descriptions. On the left is a visualization of the spherical harmonics basis functions, which are the real parts of  $Y_l^m$ .  $l$  is from 0 (top) to 9 (bottom), and  $m$  increases from 0 (left) to  $l$ . The right figure describes the reconstruction of a left ventricle using degrees 1, 3, 6, and 10 (from top to bottom and left to right).

### 2.2.1 Alignment

During the heart cycle, the heart not only dilates and contracts, but also rotates within certain small angles. Given a set of shape sequences, we need to first align all the shapes within the same sequence, and then align the different sequences. This can be done efficiently if landmarks are defined at the locations of papillary muscles on the surface of the left ventricle; these muscles are the only structures that can be clearly discriminated from the ventricle (and heart) surface. Using the landmarks as guides during a rigid body rotation, the ventricle shapes (including the heart that contains the corresponding left ventricle) are aligned.

### 2.2.2 Scaling

Since we are interested in measuring shape differences, an appropriate scaling method has to be defined in a fast and accurate way. Because the moment at the end of diastole is the most important one in medical analysis during heart cycle, we normalized the diastolic shape volume of every time sequence shape data. After that, every time sequence gain a scaling factor  $f_s$ , and the other new SPHARM coefficients can be achieved by dividing all the coefficients by the scaling factor  $f_s$ . Note that the computation of shape differences without any scale normalization sometime can exhibit difference between small and large objects even though they have the same shape properties, and these results may help us to understand some clinically relevant cases. To achieve this, we save all original object volumes of the sequence in our feature vector in order to do volume similarity or other analysis in the future.

After the above steps, a set of normalized coefficients are obtained that form a comparable shape descriptor for each cardiac surface.

	m=-3	m=-2	m=-1	m=0	m=1	m=2	m=3
l=0				.932			
l=1			.073	.201	.045		
l=2		.277	.028	.304	.029	.281	
l=3	.018	.009	.011	.012	.010	.009	.017

**Table 1.** The coefficients  $c_{xl}^m$  of spherical harmonics descriptor for left ventricle in our experiment.

### 2.3. Feature extraction

As with a Fourier representation, we can approximate  $r(\theta, \phi)$  by truncating the series of spherical harmonic coefficients to a finite number of terms. An upper limit,  $L$ , can be chosen to give a desired level of resolution. Thus

$$r(\theta, \phi) = \sum_{l=0}^L \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi).$$

The coefficients  $c_l^m$  are estimated by solving a set of linear equations. An approximate cardiac surface (using vertices and a mesh) can be reconstructed using any number of these coefficients, with more coefficients providing a more detailed and accurate reconstruction than fewer. An example of the absolute values of the SPHARM coefficients in our experiment (up to  $L = 3$ ) is given in Table 1. Note that each increasing order ( $l$ ) supports more degrees ( $m$ ), as explained in Section 2.

We collect the first  $L$  orders (*i.e.*,  $L + 1$  rows) of coefficients of every object as vectors  $r_i^j$  in the spatial domain. We assume that every shape sequence is composed of a particular number,  $s$ , of shapes: in our cardiology application,  $s = 20$  as 20 successive heart shapes are recorded on the time axes during a heart cycle. The feature vector representing the shape sequence,  $r_i$ , is created from from all  $s$  shapes in a sequence. This entails combining  $s$  heart shape models, including all the coefficients vectors  $r_i^j$ , together with the volume  $r_i^v$ :

$$r_i = (r_i^1, r_i^2, \dots, r_i^{20}, r_i^v)^T$$

Feature extraction of these shape sequence vectors is efficient. During our experiment, on a PC with an 2.40 GHz pentium IV processor and 512 MB RAM running Windows XP, the average time for extracting feature vectors for a single sequence with  $L = 128$  is about 23 seconds.

### 3. Description of indexing approach

Fig.1 illustrates our similarity retrieval framework for the spatio-temporal cardiac shape data. We use the Euclidean distance as the distance function between surfaces. Formally, for two surfaces given by  $v_1(s)$  and  $v_2(s)$ , we define their distance  $D(v_1, v_2)$  as

$$D(v_1, v_2) \equiv \left( \oint \|v_1(s) - v_2(s)\|^2 ds \right)^{1/2}$$

If this distance is below a threshold  $\varepsilon$  defined by user, we consider the two shapes identical with respect to this threshold.

Each cardiac shape sequence of our dataset is transformed to a point in an  $f$ -dimensional space (described in section 3.1). When a query is given, it is also mapped to a point into the  $f$ -dimensional space. And the most similar shape sequences are found by the Euclidean distance computations.

---

#### Algorithm 1 The generic algorithm

---

**R\*-tree search:** R\*-trees provide us the fast search. The result is a set of matching shape sequence (possibly including some incorrectly matched sequences).

**False filtering:** For each of the above obtained shapes set,

1. if query is a shape sequence, the distances (length is  $L$ , not  $f$ ) between those shape sequences and the query are computed;
2. if user only focuses on the diastolic shape, the distances between those shapes and the query shape are computed.

If the distance is less than the tolerance  $\varepsilon$ , the corresponding shape or shape sequence is in the answer set.

---

### 3.1. SPHARM-index structure

During the cardiac shape study, the end of diastole case is considered in particular, as many pathological symptoms appear at this moment (*e.g.* dilated, constricted, and restrictive cardiomyopathy). In other terms, the most differences between two shape sequences exhibit at diastolic shape. And a query always is a shape sequence or a shape of the end of diastole (only such shapes are used in user's study). Approximately we index the shape sequence by the diastolic feature vectors. Many methods can be used for our index structure, such as R-tree [7], R\*-tree [1], X-tree [2]. We only keep the first few ( $f \ll L$ ) coefficients of diastolic feature vectors, and each sequence is mapped into a point in an  $f$ -dimensional space. The R\*-tree holds an  $f$ -dimensional vector for each shape sequence. Because the first  $f$  spherical harmonics coefficients are used for indexing, we call this method as SPHARM-index structure.

### 3.2. Search strategy

Queries of both shape sequences and representative shapes (*e.g.*, diastolic shape in cardiology) are of interest. Here, we focus on finding shape sequences within a tolerance  $\varepsilon$  of a given sequence.

Given two shape sequences and their representations under the SPHARM-index structure of Section 3.1, it is desirable to show that the index structure does not introduce any problems in calculating similarities. Lemma 1 shows that the distance in feature space (*i.e.*, using a SPHARM-index representation with limited components) should match or underestimate the Euclidean distance between the original sequence of cardiac surfaces. Thus, our data reduction may introduce false alarms, which must be dealt with in post-processing, but will not omit correct matches [5]! The whole search process is outlined in Algorithm 1. A similar proof can be used to demonstrate that corresponding subsequences of a similar (within  $\varepsilon$ ) sequence are also similar.

**Lemma 1** The SPHARM-index introduces no false dismissals.

**Proof:** We define the Euclidean distance of two surfaces  $v_1$  and  $v_2$  as  $D_{object}$ , and the Euclidean distance for a corresponding feature space with a selected number of dimensions as  $D_{feature}$ . Let  $H(v_1)$  and  $H(v_2)$  be the portions of  $v_1$  and  $v_2$  that are cut for indexing. We wish to show that for

$$D_{object}(v_1, v_2) \leq \varepsilon.$$

it is also true that

$$D_{feature}(H(v_1), H(v_2)) \leq \varepsilon.$$

The spherical harmonics basis functions are similar to a Fourier basis, and their orthogonality allows the use of Parseval's theorem [10]:

$$\oint \|v_1(s) - v_2(s)\|^2 ds = \sum_{l=0}^{\infty} \sum_{m=-l}^l |c_{l1}^m - c_{l2}^m|^2,$$

and

$$\begin{aligned} \sum_{l=0}^L \sum_{m=-l}^l |c_{l1}^m - c_{l2}^m|^2 &\leq \sum_{l=0}^{\infty} \sum_{m=-l}^l |c_{l1}^m - c_{l2}^m|^2 \\ &= \oint \|v_1(s) - v_2(s)\|^2 ds. \end{aligned}$$

Because we only use  $f \ll L$  coefficients to calculate the feature space distance, we have:

$$\begin{aligned} |D_{feature}(H(v_1), H(v_2))|^2 &= \sum_{l=0}^f \sum_{m=-l}^l |c_{l1}^m - c_{l2}^m|^2 \\ &< \oint \|v_1(s) - v_2(s)\|^2 ds. \end{aligned}$$

The Euclidean distance of two surfaces

$$D(v_1, v_2) = (\oint \|v_1(s) - v_2(s)\|^2 ds)^{1/2}$$

implies

$$D_{feature}(H(v_1), H(v_2)) \leq D(v_1, v_2) \leq \varepsilon$$

and therefore the SPHARM-index will return all feature vectors, plus potentially some false matches. That is to say, the query result is a superset of the correct. ■

## 4. Conclusion

This paper presents a framework for similarity measurements between sequences of shape descriptors, with particular reference to automated analysis of the heart. The required normalization and feature extraction steps are currently slow, but searches on indexed data are quite fast. Future work will accelerate normalization and feature extraction and focus on complete data mining tools, combining the shape feature vectors used here with additional features and with clinical outcome data.

## References

- [1] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger. The R\*-tree: an efficient and robust access method for points and rectangles. In *Proceedings of 1990 ACM SIGMOD Conference*, pages 322–331, Atlantic City, NJ, 1990.
- [2] S. Berchtold, D. A. Keim, and H.-P. Kriegel. The X-tree: An index structure for high-dimensional data. In *Proceedings of the 22nd International Conference on Very Large Databases*, pages 28–39, San Francisco, CA, 1996.
- [3] B. Bhanu. Representation and shape matching of 3-D objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(3):340–351, May 1984.
- [4] A. Bimbo and P. Pala. Shape indexing by structural properties. In *Proceedings of IEEE Multimedia System*, pages 370–375, 1997.
- [5] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos. Fast subsequence matching in time-series databases. In *Proceedings of 1994 ACM SIGMOD Conference*, pages 419–429, Minneapolis, MN, 1994.
- [6] G. Gerig, M. Styner, M. E. Shenton, and J. A. Lieberman. Shape versus size: Improved understanding of the morphology of brain structures. *Lecture Notes in Computer Science*, 2208:24–31, 2001.
- [7] A. Guttman. R-trees: a dynamic index structure for spatial searching. In *Proceedings of 1984 ACM SIGMOD Conference*, pages 47–57, Boston, MA, 1984.
- [8] H. Huang, L. Shen, F. Makedon, L. Gao, and J. Pearlman. Three-dimensional analysis of cardiac magnetic resonance imaging using spherical harmonics model. In *Supplement to the Journal of the American College of Cardiology for ACC'04*, March 7–10, 2004.
- [9] L. Shen, J. Ford, F. Makedon, and A. Saykin. A surface-based approach for classification of 3D neuroanatomic structures. *Intelligent Data Analysis*, 8(5), 2004.
- [10] E. W. Weisstein. From MathWorld – A Wolfram Web Resource. <http://mathworld.wolfram.com/>.
- [11] D. S. Zhang and G. Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19, 2004.