Incomplete 3D Shape Retrieval via Sparse Dictionary Learning

Lili Wan 1† Jingyu Jiang 1 Hao Zhang 2

1School of Computer and Information Technology, Beijing Jiaotong University, China
2School of Computing Science, Simon Fraser University, Canada

Abstract

How to deal with missing data is one of the recurring questions in data analysis. The handling of significant missing data is a challenge. In this paper, we are interested in the problem of 3D shape retrieval where the query shape is incomplete with moderate to significant portions of the original shape missing. The key idea of our method is to grasp the basis local descriptors for each shape in the retrieved database by sparse dictionary learning and apply them in sparsely coding the local descriptors of an incomplete query. First, we present a method of computing heat kernel signatures for incomplete shapes. Next, for each shape in the database, a set of basis local descriptors, which is called a dictionary, is learned and taken as its representative. Finally, a query incomplete shape’s heat kernel signatures are respectively reconstructed by each dictionary, and the shape similarities are therefore measured by the reconstruction errors. Experimental results show that the proposed method has achieved significant improvements over previous works on retrieving non-rigid incomplete shapes.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Curve, surface, solid, and object representations

1. Introduction

One of the recurring questions in data analysis is how to deal with missing data. Methods aimed at handling complete data may still be applicable if the missing data is not significant and data completion is possible via interpolation. However, when large chunks of data are missing, e.g., entire parts are cut off from a shape, interpolation no longer works and new methods, or significant adjustments to methods designed for complete data, must be developed. Previous attempts on problems such as shape retrieval [DLL∗10] and matching [FS06, GCO06] have revealed that the handling of significant missing data can be quite a challenge.

In this paper, we are interested in the problem of 3D shape retrieval where the query shape is incomplete with moderate to significant portions of the original shape missing. The database on the other hand would consist of complete shapes. Such a problem setting may arise in practice when a modeller wants to create a new 3D shape via part composition and needs to search for one or more missing parts for a partially created shape, which is incomplete. In 3D model reconstruction amid significant missing data, a partial reconstructed shape, which is again incomplete, may be used to query a database for data-driven model completion.

Shape retrieval typically relies on one or more global shape descriptors [RWP06, JZ07]. However, by design, these global descriptors are unlikely to work for retrieving highly incomplete shapes. Local shape descriptors, including the well known shape context [KPNK03] and heat kernel signature (HKS) [SOG09], encode geometry at or from the perspective of a point over a shape’s surface. The entire set of local descriptors for a shape, incomplete or not, needs to be properly organized into a shape signature to allow a sensible comparison between incomplete and complete shapes.

This paper is devoted to incomplete non-rigid 3D shape retrieval. The most similar work is introduced in [DLL∗10]. Their method focuses on critical points detecting, wherein the concept of persistent homology is brought in to help to extract a set of HKS maxima. The HKS descriptors of these

† Corresponding author: llwan@bjtu.edu.cn

© The Eurographics Association 2015.
critical points form a feature set which is called persistent heat signature (PHS), and the similarity between two shapes is then measured based on their feature sets. However, when some parts of a shape are missing, the detection of critical points may be easily impacted. For example, if an original critical point happens to locate on the missing part, another point of the incomplete shape will instead be selected as the critical point.

The key observation we make is the similarity of local descriptors between incomplete and complete shapes. More specifically, if the set of local shape descriptors of a complete shape $S$ had been well sparsely coded by a dictionary, which consists of a few basis signals, via sparse dictionary learning, then the dictionary would also be capable of representing the same local descriptors of incomplete versions of $S$. The overview of our shape similarity measure is shown in Figure 1. In shape retrieval, each shape in the database has its dictionary. When taking an incomplete shape as a query, the shape similarities can be obtained by using these dictionaries respectively to represent the query’s local descriptors.

Our main contributions include:

- To reduce the negative impact of boundary regions to local shape descriptors, we propose a new method to compute heat kernel signatures for a non-rigid incomplete shape, which is also applicable for complete shapes.
- In order to retrieve an incomplete shape, previous work focuses on detecting and matching critical points. Instead, in our approach, the relationship of local descriptors between a complete shape and an incomplete shape is established by a learned dictionary.

2. Related work

The literature on shape descriptors, shape matching, and retrieval is vast. In this section, we only cover methods that are most closely related to our work. We refer the readers to a number of surveys on these topics, including [BP06, TV08, vKZHCO11].

Sparse dictionary learning. Sparse dictionary learning methods have been versatile tools in many applications including signal processing and image processing [Ela10, TF11]. Recently, it also attracts some researchers in the field of 3D shape processing. Abdelrahman et al. [AEMF12] utilize sparse representation to reduce the dimensionality of a global shape descriptor. For efficient shape retrieval, local descriptors are usually pooled into a global descriptor by some statistic methods, e.g. bag-of-words (BoW). Sparse dictionary learning can also be applied in this task. In [WLMC13, LBB14], the clustering process in the BoW methods is replaced by sparse dictionary learning. Boscaiani et al. [BC14] investigate the utility of sparse coding to partial shape retrieval, wherein the queries are composites of sub-parts from different classes. However, its use in 3D shape retrieval is still relatively new, and it has not yet been used for incomplete shape retrieval.

Partial matching. Partial shape matching can be applied to compute shape similarity, e.g., [GCO06, FS06], for shape retrieval. Tierny et al. [TVD09] match partial 3D shapes via Reeb pattern unfolding. Shapira et al. [SSS10] execute part-in-whole object queries via hierarchical graph matching. Ferreira et al. [FMA10] propose a part-in-whole matching method for engineering shapes to retrieve 3D models containing a part similar to a query. Itskovich et al. [IT11] present an archaeological application which finds the best match to a specified surface shape within other surfaces. In our work, we compute shape similarity between an incomplete and a complete shape via sparse dictionary learning.

3. Incomplete HKS (I-HKS)

3.1. HKS vs. WKS for incomplete shapes

Heat kernel signature (HKS) [SOG09, BBGO11] and wave kernel signature (WKS) [ASC11b] are two state-of-the-art methods in non-rigid shape analysis. HKS has a foundation of heat diffusion, while WKS is induced from quantum mechanics. They are both invariant under isometric transformations. However, shape missing may influence the energy diffusion on a surface, so we need to analyze which one is more appropriate for incomplete shapes.
The HKS descriptors and WKS descriptors are visualized in Figure 2 for a human shape and its incomplete versions. Our settings to compute HKS and WKS are due to [BBGO11] and [ASC11a] respectively.

From Figure 2, we can see that the HKS descriptors only vary near the cutting boundaries, but the WKS descriptors change significantly in some regions far from the boundaries. Based on the analysis in [ASC11a] and [LB14], HKS can be seen as a collection of low-pass filters, while the responses of WKS are band-pass. Therefore, WKS is good for featuring localization. However, WKS is more sensitive to missing data than HKS.

3.2. HKS for incomplete shapes

For an incomplete shape, we calculate the HKS descriptors for vertices on the largest connected component, and then exclude those descriptors of the boundary vertices and their 1-ring neighbors.

In practice, for a vertex \( x \) on a surface, its HKS descriptor at time \( t \) is approximately computed by

\[
h(x,t) = \sum_{k=0}^{K-1} e^{-\lambda_k t} \phi_k(x),
\]

where \( 0 = \lambda_0 > -\lambda_1 > -\lambda_2, \ldots \) are eigenvalues of the Laplace-Beltrami operator and \( \phi_0, \phi_1, \ldots \) are the corresponding eigenfunctions.

To deal with the subsequent procedure of dictionary learning, the dimension \( M \) of an I-HKS descriptor needs to be set carefully. We utilize the K-SVD algorithm for dictionary learning. As known from [AE06], the sparsity threshold should be small enough relative to the dimension of a signal, because in these circumstances the convergence can be guaranteed. Therefore, the dimension \( M \) can not be too small. Meanwhile, \( M \) should be smaller than the dictionary size. Consequently, in all the experiments of this paper, \( M \) is fixed to 10.

Then, we select the time scales to ensure the invariance of local descriptors as much as possible. The first 100 eigenvalues and eigenfunctions are used to compute the I-HKS descriptors. From [SOG09], we can deduce that the elements of an I-HKS descriptor with \( t > |4 \ln 10/\lambda_j| \) remain almost unchanged, and those elements with \( t < |4 \ln 10/\lambda_j| \) need more eigenvalues and eigenfunctions. For an incomplete shape, small time is more appropriate for representing local attributes. So for each 3D model, we sample \( M \) points between \( t_{\min} = |4 \ln 10/\lambda_{99}| \) and \( t_{\max} = |4 \ln 10/\lambda_{99}| + (|4 \ln 10/\lambda_{99}| - |4 \ln 10/\lambda_{99}|)/10 \), and generate a logarithmically spaced vector. The time scales are then formulated as

\[
t = 10^{t_{\text{min}} + (\lg t_{\text{max}} - \lg t_{\text{min}})/(M-1)}, \quad i = 0, \ldots, M-1.
\]

Finally, all the I-HKS descriptors are normalized to the unit L2 norm for the subsequent matching procedure.

4. Shape similarity

From Figure 2a, we can see that: (1) The local descriptors of a vertex and its neighbors are very close; (2) Two symmetric parts, e.g. left hand and right hand, also have nearly equal local descriptors, and therefore these descriptors are largely redundant. According to the related studies on sparse dictionary learning [TF11,MBPS09], it is appropriate to deal with this kind of redundant information. For a complete shape, taking its I-HKS descriptors as signals, we attempt to learn the basis signals to reconstruct the whole signal set. For its incomplete shape, the I-HKS descriptors can be considered as a signal subset, and they are expected to be also well sparsely represented by the learned basis signals. We, therefore, use the reconstruction error to formulate the shape similarity measure problem.

4.1. Dictionary learning

Sparse dictionary learning has multiple applicable object functions. In our application, the local descriptors vary smoothly along the surface, and thus a vertex’s local descriptor can be interpolated by the descriptors of its nearby vertices. Therefore, we use the sparsity threshold to constrain how many basis signals will contribute to the interpolation,
and naturally choose the object function combined with the sparsity threshold.

For a complete shape \( S_i \) with \( n \) vertices, its I-HKS descriptors \( \{ f_i | i = 1, \ldots, m \} \) are all computed, each of which is taken as a training signal. Let us denote its dictionary as \( D_i \). Each signal \( f_i \) is expected to be approximately represented as a sparse linear combination of basis signals from \( D_i \), which can be described as

\[
f_i \approx D_i \gamma_i \quad \text{Subject To } \| \gamma_i \|_0 \leq T,
\]

where \( \gamma_i \) consists of sparse coefficients and \( T \) is a sparsity threshold.

In the learning process, taking the training signal set \( \{ f_i \} \) and the dictionary size as inputs, the constrained optimization problem can be formulated as

\[
\hat{D}_c = \min_{D_c} \frac{1}{n} \sum_{i=1}^{n} \| f_i - D_c \gamma_i \|_2^2 \quad \text{Subject To } \| \gamma_i \|_0 \leq T. \tag{4}
\]

### 4.2. Sparse representation

In order to ensure the invariance of local descriptors, the descriptors of those greatly impacted regions, e.g. unconnected small parts, boundary vertices and their 1-ring neighbors, are excluded from the descriptor set. For an incomplete shape \( S_i \) with \( m \) vertices, the number of its I-HKS descriptors \( m' \) will be less than \( m \) because of the exclusion. Given a dictionary \( D_c \) for a I-HKS descriptor \( f_i \), the reconstruction error of sparse coding is expressed as

\[
err_{j,c} = \min \| f_i - D_c \gamma_j \|_2^2 \quad \text{Subject To } \| \gamma_j \|_0 \leq T. \tag{5}
\]

Next, given a query \( S_j \) with a descriptor set \( \{ f_j | j = 1, \ldots, m' \} \) and a shape \( S_i \) with a dictionary \( D_c \), the average reconstruction error is utilized to measure their distance, which is formulated as

\[
dist(S_i, S_j) = \frac{1}{m'} \sum_{j=1}^{m'} err_{j,c} \tag{6}.
\]

Each shape in the retrieved database has a dictionary. Then, after using each dictionary to reconstruct the query’s descriptors respectively, we can get and sort the shape similarities. In practice, we use SPAMS (SPArse Modeling Software) [MBPS09,MBPS10] which is an optimization toolbox for solving various sparse estimation problems.

### 5. Results

In this section, we present retrieval results under several experimental settings and compare our results to those obtained by HKS and PHS. To make the comparison informative in regards to the work of Dey et al. [DLL*10], we first test the three methods on the small dataset used in [DLL*10]. Then we expand the scale of the experiment significantly by retrieving 150 incomplete shapes from the SHREC 2015 database.

**Figure 3:** Generated incomplete samples. First row shows complete shapes, and the other rows respectively show incomplete shapes in strength 1, 2 and 3.

### 5.1. Experimental setup

**Dataset.** To evaluate the performance of our method for incomplete shape retrieval, we used two publicly-available collections in experiments. One is the PHS dataset from [DLL*10], which consists of two parts: 50 queries and a database of 300 shapes divided into 21 classes, and the other is the newest non-rigid shape benchmark: SHREC 2015 database [LZC*15], which is composed of 1200 models of 50 categories. In all, we establish the following two datasets:

- **Dataset 1:** PHS queries + PHS database. It is the dataset used in [DLL*10]. The queries are 32 incomplete and 18 complete shapes, and the database contains complete and incomplete shapes.

- **Dataset 2:** Generated incomplete shapes + SHREC 2015 database. We manually created 150 incomplete shapes (3 per class) from the SHREC 2015 database as the queries, which appeared in three different incomplete strengths numbered 1-3. Some of them are shown in Figure 3.

**Parameters.** For each model, our I-HKS shape descriptors are computed using the first 100 eigenvalues and eigenfunctions of the Laplace-Beltrami operator. The dimension of each descriptor is 10, and the selection of time scales is introduced in Section 3. For sparse dictionary learning, the dictionary size is fixed to 12, the number of iterations is 1000, and the sparsity threshold \( T \) is 2.

**Assessment criteria.** We utilized the Top-k hit rate presented in [DLL*10] to evaluate the performance of incomplete shape retrieval. Its ideal score is 100%, and higher scores represent better results.
Table 1: Top-3 / Top-5 hit rates on Dataset 1.

<table>
<thead>
<tr>
<th>#queries</th>
<th>Ours</th>
<th>PHS</th>
<th>HKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>32 incomplete</td>
<td>91% / 94%</td>
<td>88% / 91%</td>
<td>56% / 63%</td>
</tr>
<tr>
<td>18 complete</td>
<td>100% / 100%</td>
<td>78% / 83%</td>
<td>83% / 89%</td>
</tr>
<tr>
<td>50 total</td>
<td>94% / 96%</td>
<td>84% / 88%</td>
<td>66% / 72%</td>
</tr>
</tbody>
</table>

Table 2: Top-3 / Top-5 hit rates on Dataset 2.

<table>
<thead>
<tr>
<th>Strength</th>
<th>Ours</th>
<th>PHS</th>
<th>HKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>94% / 94%</td>
<td>78% / 78%</td>
<td>48% / 54%</td>
</tr>
<tr>
<td>≤ 2</td>
<td>87% / 88%</td>
<td>65% / 67%</td>
<td>36% / 41%</td>
</tr>
<tr>
<td>≤ 3</td>
<td>74% / 76%</td>
<td>51% / 56%</td>
<td>30% / 34%</td>
</tr>
</tbody>
</table>

5.2. Evaluations of incomplete shape retrieval

We compared our method with two competitive shape retrieval methods: Persistent Heat Signature (PHS) method [DLL*10], and Heat kernel signature (HKS) method [BBGO11]. We chose these two methods because PHS represents a state-of-the-art technique for incomplete non-rigid shape retrieval, and HKS is a representative spectral method for non-rigid shape retrieval.

For the PHS method, the parameter settings were the same as [DLL*10]. For the HKS method, we used the first 100 eigenvalues and eigenfunctions as [BBGO11] to compute the HKS descriptors, which were all 6D vectors. In order to be fit for the model scales in our datasets, the time scales were chosen by \( t = \alpha^{-1}t_0 \) with \( t_0 = 0.006 \), \( \alpha = 2 \) and \( i = 1, \ldots, 6 \). For Dataset 1, all the models in the database were used to train a vocabulary, and the word number was 64. For Dataset 2, 50 models (1 per class) were selected to train a vocabulary, and the word number was 192, because there were more classes in the SHREC 2015 database.

Table 1 shows the Top-3 and Top-5 hit rates on Dataset 1. PHS performs better than HKS for the queries of incomplete shapes, and HKS achieves better performance than PHS for complete shapes as the queries. However, our method has obviously better performance than PHS and HKS in these two circumstances. Since the database has both incomplete and complete shapes, we conclude that our method can also deal with the shape similarity measure between two complete shapes or between two incomplete shapes.

Next, we assess our method under different incomplete strengths using Dataset 2, and present the results in Table 2. Each row shows the hit rates using the queries of the specified incomplete strength. Our method also performs much better than PHS and HKS in this dataset.

5.3. Running time

To evaluate the running time, we used the SHREC 2015 database. All the experiments in this section were carried out using MATLAB R2010b on a laptop with a 2.5GHz dual-core 4-thread CPU and 8.00 GB RAM.

The retrieval time is shown in Table 3. HKS has the fastest retrieval speed. The retrieval time of PHS and HKS algorithms are nearly equal for any model, while ours increases with the number of vertices. To accelerate our algorithm, we utilized Matlab parallel computing, and the results are shown in the Ours(PC) column. From them, we find that the parallelism can obviously reduce the retrieval time, but our algorithm will still be slower than PHS and HKS. Although the time needed for retrieval is increased, the performance of accuracy achieves significant improvements, as shown in Section 5.2.

<table>
<thead>
<tr>
<th>Query</th>
<th>#v</th>
<th>Ours</th>
<th>Ours(PC)</th>
<th>PHS</th>
<th>HKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>T684</td>
<td>2959</td>
<td>5.8</td>
<td>3.2</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>T784</td>
<td>5971</td>
<td>10.7</td>
<td>5.5</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>T470</td>
<td>9999</td>
<td>17.8</td>
<td>10.8</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>T837</td>
<td>14718</td>
<td>25.8</td>
<td>13.4</td>
<td>0.8</td>
<td>0.5</td>
</tr>
</tbody>
</table>

6. Conclusion, limitation, and future work

We propose a dictionary learning and sparse coding based framework for non-rigid incomplete shape retrieval. Different from previous work of detecting and matching critical points, we measure shape similarities based on reconstructing the local shape descriptors of a query under sparse constraints. We also present a method of computing heat kernel signatures for incomplete shapes. The proposed method has achieved significant improvements over previous works on retrieving non-rigid shapes amid significant missing data.

One major limitation of our current method is that the query shape is assumed to be connected. If the input shape is disconnected but with some large connected components, then the retrieval can simply be conducted on the largest component. Another limitation is that we assume that the boundary regions are easy to detect. While this assumption often holds when a complete model is being cut, in practice, particularly for partial surface reconstruction, boundary detection is not always an easy task.

For future work, we would like to deal with incomplete point clouds, incomplete topology-varying man-made shapes, etc. Our framework is not restricted to a particular local shape descriptor, and the I-HKS shape descriptors in our method are possible to be replaced by other descriptors. With the progress of local shape descriptors, it may be applicable to those more complex cases in retrieving incomplete shapes.

Acknowledgment

The work is supported in part by China Scholarship Council, National Natural Science Foundation of China (61370127),
References


