ABSTRACT
Invariant shape descriptors are instrumental in numerous shape analysis tasks including deformable shape comparison, registration, classification, and retrieval. Most existing constructions model a 3D shape as a two-dimensional surface describing the shape boundary, typically represented as a triangular mesh or a point cloud. Using intrinsic properties of the surface, invariant descriptors can be designed. One such example is the recently introduced heat kernel signature, based on the Laplace-Beltrami operator of the surface. In many applications, however, a volumetric shape model is more natural and convenient. Moreover, modeling shape deformations as approximate isometries of the volume of an object, rather than its boundary, better captures natural behavior of non-rigid deformations in many cases. Here, we extend the idea of heat kernel signature to robust isometry-invariant volumetric descriptors, and show their utility in shape retrieval. The proposed approach achieves state-of-the-art results on the SHREC 2010 large-scale shape retrieval benchmark.

Categories and Subject Descriptors
I.4.7 [Image Processing and Computer Vision]: Feature Measurement—Feature representation, Invariants

General Terms
Algorithms

Keywords
Volumetric Laplacian, Heat Kernel Signature

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

3DOR’10, October 25, 2010, Firenze, Italy.
Copyright 2010 ACM 978-1-4503-0160-2/10/10 ...$10.00.
we formulate the problem and overview the heat kernel descriptors. In Section 4, we propose the volumetric heat kernel descriptors, and address their numerical computation. In Section 5, we outline the construction of bag-of-feature shape descriptors from the local volumetric heat kernel descriptors. Section 6 shows experimental results, and Section 7 concludes the paper.

2. PROBLEM FORMULATION

We model a 3D object as a closed and connected three-dimensional submanifold $X \subset \mathbb{R}^3$, with boundary $\partial X$ represented as a smooth closed two-dimensional surface. A classical and predominant approach for 3D object analysis is presented as a smooth closed two-dimensional surface. A classical and predominant approach for 3D object analysis is by studying the intrinsic geometry of the boundary surface $\partial X$. In these approaches, one of the desired properties of a shape descriptor is invariance to isometric deformations of the surface $\partial X$ (referred here as boundary isometries), or in other words, such transformations that preserve the lengths of shortest paths between any pair of point on $\partial X$. These isometries are opposed to volume isometries, in which shortest paths between points inside $X$ do not change. It is worthwhile noting that exact isometries of $X$ in $\mathbb{R}^3$ are limited to the set of rigid transformations of $X$, that is, no non-trivial isometries of the volume exist. However, non-rigid deformations can still be modeled faithfully as approximate volume isometries. We argue that such nearly isometric deformations are more suitable to model natural non-rigid deformations and articulations of solid objects than approximate boundary isometries. This is illustrated in Figure 1 showing a glove and a human hand. A glove is an illustration of $\partial X$, and its deformations are boundary isometries. The hand illustrates the solid object $X$, and its deformations can be approximated as volume isometries. This example shows how in many cases using $\partial X$ to model $X$ can be wrong: some of the isometries of the glove, such as “deflating” it, are clearly inadmissible for the solid hand object. Approximate boundary isometries constitute a large class of deformations, some of which (e.g., volume-changing deformations) do not model well the natural deformations of the solid.

In this paper, we propose dealing with the object $X$ as with the volume enclosed inside the boundary rather than the $\partial X$ itself, and derive a descriptor based on volumetric heat propagation properties similarly to [36].

3. HEAT KERNEL SIGNATURES

Heat diffusion on the surface $\partial X$ is governed by the heat equation,

$$\left(\Delta_{\partial X} + \frac{\partial}{\partial t}\right) u(t,x) = 0,$$

where the scalar field $u : \partial X \times [0, \infty) \rightarrow \mathbb{R}$ is the value of heat on object boundary surface $\partial X$ at time $t$, and $\Delta_{\partial X}$ is the positive semi-definite Laplace-Beltrami operator, a generalization of the Laplacian to manifolds. The fundamental solution $k_t(x,z) : \partial X \times \partial X \times [0, \infty) \rightarrow \mathbb{R}$ of the heat equation, also called the heat kernel, is the solution of (1) initialized by a point heat distribution at $x$.

Sun et al. [36] proposed using the diagonal of the heat kernel $k_t(x,x)$ at multiple scales as a local descriptor, referred to as the heat kernel signatures (HKS). The HKS is invariant under isometric deformations of $\partial X$, insensitive to topological noise at small scales, and informative in the sense that under certain assumptions one could reconstruct the surface (up to an isometry) from it.

Furthermore, the computation of the HKS descriptor relies on the computation of the first eigenfunctions and eigenvalues of the Laplace-Beltrami operator, which can be done efficiently and across different shape representations.

A disadvantage of the HKS is its dependence on the global scale of the shape, manifested as scaling of the descriptor and the time parameter. As a remedy, a local normalization based on logarithmic scale space and magnitude of the Fourier transform, dubbed the scale-invariant HKS (SI-HKS), was proposed in [10].

4. VOLUMETRIC HEAT KERNEL SIGNATURES

In order to generalize heat kernel signatures to volume, we consider the heat equation in the volume $X$ with Neumann boundary conditions on $\partial X$,

$$\left(\Delta + \frac{\partial}{\partial t}\right) U(x,t) = 0 \quad x \in \text{int}(X),$$

$$\langle \nabla U(x,t), n(x) \rangle = 0 \quad x \in \partial X$$

where $n$ is the normal to the boundary surface $\partial X$, $\Delta$ is the standard Laplacian in $\mathbb{R}^3$, and $U : (X \subset \mathbb{R}^3) \times [0, \infty) \rightarrow \mathbb{R}$ is the volumetric heat distribution in $X$.

We denote by $K_t(x,y) : X \times X \times [0, \infty) \rightarrow \mathbb{R}$ the heat kernel of (2). By the spectral decomposition theorem, $K_t(x,y)$ can be written as [18]

$$K_t(x,y) = \sum_{l=0}^{\infty} e^{-\lambda_l t} \Phi_l(x) \Phi_l(y).$$

where $\lambda_l$, $\Phi_l$ are the eigenvalues and eigenfunctions of the Laplacian operator with the above boundary conditions,

$$\Delta \Phi_l(x) = \lambda_l \Phi_l(x),$$

$$\langle \nabla \Phi_l(x), n(x) \rangle = 0 \quad x \in \partial X$$

Figure 1: Deformations of a glove (left) and a solid hand (right) are an illustration of the difference between boundary and volume isometries.
The heat kernel of (2) gives rise to volumetric heat kernel signatures (VHKS) which we define exactly in the same way as HKS. Namely, for each point \( x \in X \),
\[
    h(x) = (K_{t_1}(x,x), \ldots, K_{t_n}(x,x)),
\]
is used as an \( n \)-dimensional local descriptor, where \( t_1, \ldots, t_n \) are different time-scales.

VHKS are related to a geometric quantity called the scalar curvature or the trace of the Ricci curvature by
\[
    K_i(x,x) = \frac{1}{(4\pi t)^{3/2}} \left(1 + \frac{1}{6} \sigma(x)^2\right),
\]
where \( \sigma(x) \) is the scalar curvature. \( \sigma(x) \) can be defined as the ratio of the volume of a small ball centered at \( x \) to the volume of a three-dimensional Euclidean ball of the same radius.

### 5. VOLUMETRIC BAGS-OF-FEATURES

In order to aggregate local point descriptors into a single global volume descriptor, we follow the ShapeGoogle framework introduced in [9]. Given a volume \( X \), a dense point descriptor \( h(x) \in \mathbb{R}^n \) is first computed for every \( x \in X \). The second step consists of vector quantization of the descriptors. Given a vocabulary \( V \) of \( \{h_1, \ldots, h_V\} \), representative descriptors, \( h(x) \) at each point is replaced by the \( V \)-dimensional distribution
\[
    \theta_k(x) = \exp\left(-\frac{||h_k - h(x)||^2}{2\sigma^2}\right),
\]
where \( \sigma \) is a parameter. In the limit \( \sigma \to 0 \), the process boils down to the standard hard vector quantization assigning each descriptor the index of its nearest neighbor in the vocabulary. The vocabulary is created offline by clustering a training set of descriptors.

After vector quantization, the point-wise distributions \( \theta(x) \) are integrated over the entire volume, resulting in the volumetric bag-of-features
\[
    f(x) = \int \theta(x) dx.
\]

Volumetric bags-of-features are compared using a standard metric in \( \mathbb{R}^V \), e.g., the \( L_1 \) metric.

### 6. RESULTS

In order to test our approach, we used local feature descriptors to construct global shape descriptors, which were used to compare shapes. Global shape descriptors were constructed using bags of geometric words proposed in [9]. A geometric vocabulary of size 64 was built using clustering in the space of descriptors. We used the first 150 eigenpairs of the Laplacian to approximate the heat kernel and 6 time scales \( (t = 1024, 1351, 1783, 2353, 3104, 4096) \) in the construction of the VHKS. Shapes were represented as volumes; the VHKS at each point in the volume of the shape were replaced by the closest geometric word from the vocabulary using soft vector quantization. The distribution of geometric words (64-dimensional bag of features) was used as the shape descriptor. \( L_1 \) distance was used to compare the bags of features.

#### 6.1 Boundary vs Volume Isometries

In the first experiment, we designed a special dataset to illustrate the difference between boundary and volume isometries and thus emphasize the problem of approaches considering 3D shapes as 2D boundary surfaces. Our dataset contained objects from three classes (humans, dogs, and horses). In each class, the null pose of the object underwent five approximate volume isometries (being also boundary isometries) and five boundary isometries which change the volume significantly (such as inverting the human head inside the torso). For all the shapes we computed local HKS descriptors on the boundary following [9], and volumetric HKS descriptors proposed here. These local descriptors were aggregated into a global bag of features as described above.

Figure 2 shows the distance matrices between the shapes using HKS (left) and VHKS (right). HKS fail to distinguish between volume-changing deformation, though they render the shapes visually very different. With VHKS, on the other hand, such deformations are distant from the volume-preserving ones. This fact is further visible from Figure 3 depicting the two descriptors under volume-changing and volume-preserving isometries of the boundary. Observe that
while both descriptors are capable of discriminating between non-isometric shapes, HKS are invariant under volume-changing isometries of the boundary, while VHKS are not.

### 6.2 Shape retrieval

In order to evaluate the performance of the proposed descriptors for shape retrieval, we used the SHREC 2010 robust large-scale shape retrieval benchmark, simulating a retrieval scenario, in which the queries include multiple modifications and transformations of the same shape [6].

**Dataset.** The dataset used in this benchmark was aggregated from three public domain collections: TOSCA shapes [8], Robert Sumner’s collection of shapes [35], and Princeton shape repository [33]. The shapes in the original SHREC dataset are given as triangular meshes with the number of vertices ranging approximately between 300 and 30,000. We converted the shapes into a volumetric representation using rasterization, with fixed voxel size. Typical volume size was 150×150×150. The original dataset consisted of two parts: 715 shapes from 13 shape classes with simulated transformation (55 per shape) used as queries and the rest of 456 shapes, totalling in 1184 shapes.

**Queries.** The query set consisted of 13 shapes taken from the dataset (null shapes), with simulated transformations applied to them. An insignificant number of shapes containing thin structures was removed due to our inability to accurately convert them into a volumetric representation. For each null shape, transformations were split into 10 classes shown in Figure 4 (we removed the partial view transformation present in the original benchmark, as it represents a partial occlusion of the boundary and not what is expected to be a partial occlusion of the volume). In each class, the transformation appeared in five different versions numbered 1–5. The total number of transformations per shape was 50, and the total query set size was 600. Each query had one correct corresponding null shape in the dataset.

**Evaluation Criteria.** Evaluation simulated matching of transformed shapes to a database containing untransformed (null) shapes. As the database, all 469 shapes with null transformations were used. Multiple query sets according to transformation class and strength were used. For transformation $x$ and strength $n$, the query set contained all the shapes with transformation $x$ and strength $\leq n$.

Performance was evaluated using precision/recall characteristic. Precision $P(r)$ is defined as the percentage of relevant shapes in the first $r$ top-ranked retrieved shapes. In the present benchmark, a single relevant shape existed in the database for each query. Mean average precision (mAP), defined as

$$mAP = \sum_r P(r) \cdot rel(r),$$

(where $rel(r)$ is the relevance of a given rank), was used as a single measure of performance. Ideal performance retrieval performance results in first relevant match with mAP=100%.

**Comparison to other methods.** Each shape in the benchmark was represented as a bag of VHKS features in a vocabulary of size 64, as described above. Table 1 show the performance (mAP in %) across transformation classes and strengths of shape retrieval using bags of features based on VHKS local descriptors. For comparison, in Table 3 we show the performance of ShapeGoogle [9] using local HKS descriptors computed on a mesh using cotangent weight approximation of the Laplace-Beltrami operator. One can observe that our approach is more robust under strong isometric and topological transformations and is less sensitive to noise and resampling. Since the descriptor is not scale-invariant, the performance on global and local scaling and mixed transformations is poor.

Table 1: Performance (mAP in %) of our approach using VHKS descriptors.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>≤1</th>
<th>≤2</th>
<th>≤3</th>
<th>≤4</th>
<th>≤5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isometry</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Topology</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Holes</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>98.75</td>
</tr>
<tr>
<td>Micro holes</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Scale</td>
<td>0.61</td>
<td>11.94</td>
<td>8.81</td>
<td>6.74</td>
<td>5.46</td>
</tr>
<tr>
<td>Local scale</td>
<td>100.00</td>
<td>93.35</td>
<td>81.86</td>
<td>69.04</td>
<td>60.81</td>
</tr>
<tr>
<td>Sampling</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Noise</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Shot noise</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Mixed</td>
<td>100.00</td>
<td>61.14</td>
<td>41.65</td>
<td>31.47</td>
<td>25.29</td>
</tr>
</tbody>
</table>

Table 2 shows the performance of volumetric ShapeDNA [29], which represents the shapes as vectors of the corresponding Laplacian eigenvalues. We used the first 150 eigenvalues. Our approach significantly outperforms this method.

The complexity of calculating VHKS is higher than that of HKS. In surfaces, several thousand points are enough for constructing a distinguishable signature, while in volumes that number increases by a factor of ten. Since the Laplacian is represented by a sparse matrix, calculating several dozen eigenvalues and eigenvectors remains a fast procedure. Under core 2 Duo 3GHz computer, it takes approximately several seconds to calculate each VHKS. Training the dictionary is done off-line and takes several minutes.

Figure 3: RGB representation of the first three components of the HKS descriptor (first row) and VHKS descriptor on the shape boundary (second row). Second and third columns show, respectively, a volume-preserving and a volume changing boundary isometries.
Figure 4: Transformations of the human shape used as queries (shown in strength 5, left to right): null, isometry, topology, sampling, local scale, scale, holes, micro holes, noise, shot noise, and mixed.

Table 3: Performance (mAP in %) of the 2D HKS descriptor computed with cotangent weight discretization of the Laplacian.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>1</th>
<th>≤2</th>
<th>≤3</th>
<th>≤4</th>
<th>≤5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isometry</td>
<td>100.00</td>
<td>100.00</td>
<td>98.61</td>
<td>98.96</td>
<td>99.17</td>
</tr>
<tr>
<td>Topology</td>
<td>100.00</td>
<td>96.04</td>
<td>93.34</td>
<td>95.01</td>
<td>96.01</td>
</tr>
<tr>
<td>Holes</td>
<td>100.00</td>
<td>96.43</td>
<td>94.91</td>
<td>85.81</td>
<td>77.41</td>
</tr>
<tr>
<td>Micro holes</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Scale</td>
<td>80.03</td>
<td>90.02</td>
<td>93.34</td>
<td>95.01</td>
<td>96.01</td>
</tr>
<tr>
<td>Local scale</td>
<td>100.00</td>
<td>100.00</td>
<td>95.90</td>
<td>86.43</td>
<td>78.38</td>
</tr>
<tr>
<td>Sampling</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Noise</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Shot noise</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Mixed</td>
<td>40.84</td>
<td>39.95</td>
<td>46.38</td>
<td>47.79</td>
<td>44.79</td>
</tr>
</tbody>
</table>

Table 2: Performance (mAP in %) of volumetric ShapeDNA.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>1</th>
<th>≤2</th>
<th>≤3</th>
<th>≤4</th>
<th>≤5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isometry</td>
<td>56.97</td>
<td>53.65</td>
<td>49.82</td>
<td>47.97</td>
<td>48.91</td>
</tr>
<tr>
<td>Topology</td>
<td>61.97</td>
<td>57.96</td>
<td>58.83</td>
<td>56.92</td>
<td>55.44</td>
</tr>
<tr>
<td>Holes</td>
<td>53.89</td>
<td>53.72</td>
<td>51.91</td>
<td>45.40</td>
<td>41.41</td>
</tr>
<tr>
<td>Micro holes</td>
<td>51.20</td>
<td>55.05</td>
<td>56.44</td>
<td>56.15</td>
<td>56.06</td>
</tr>
<tr>
<td>Scale</td>
<td>0.91</td>
<td>1.93</td>
<td>1.52</td>
<td>1.24</td>
<td>1.08</td>
</tr>
<tr>
<td>Local scale</td>
<td>49.69</td>
<td>37.11</td>
<td>33.63</td>
<td>28.15</td>
<td>23.41</td>
</tr>
<tr>
<td>Sampling</td>
<td>58.26</td>
<td>60.18</td>
<td>59.45</td>
<td>60.70</td>
<td>58.84</td>
</tr>
<tr>
<td>Noise</td>
<td>50.74</td>
<td>51.18</td>
<td>48.76</td>
<td>43.17</td>
<td>38.50</td>
</tr>
<tr>
<td>Shot noise</td>
<td>58.93</td>
<td>57.04</td>
<td>50.53</td>
<td>42.54</td>
<td>36.88</td>
</tr>
<tr>
<td>Mixed</td>
<td>55.25</td>
<td>30.09</td>
<td>21.00</td>
<td>16.23</td>
<td>13.26</td>
</tr>
</tbody>
</table>

7. CONCLUSIONS

Our starting point was the fact that not all boundary isometries of solids correspond to meaningful deformations. We argued that volume isometries better model natural non-rigid deformations and constructed local descriptors based on volumetric heat kernels. Such descriptors readily lend themselves to the bag-of-features representation used in shape retrieval. Experimental results on the SHREC 2010 benchmark demonstrate that the proposed descriptors are more robust to topological and geometric noise than previously proposed approaches, achieving state-of-the-art results in several deformation categories. In future studies, we intend to construct a scale-invariant version of the VHKS descriptors.

8. ACKNOWLEDGEMENTS

This research was supported in part by The Israel Science Foundation (ISF) grant number 623/08, and by The USA Office of Naval Research (ONR) grant.

9. REFERENCES

NORDIA*, 2009.


detection and local classification for surfaces based on
moment analysis. *Trans. Visualization and Computer

eigendecomposition of the Laplace-Beltrami operator.
In *Proc. 3DPVT*, 2010.

representations for surfaces. In *Proc. CVPR*, pages

shape signature. *Trans. Visualization and Computer

SGP*, 2005.

Topology matching for fully automatic similarity

Kropatsch, and L. Cohen. Shape matching using the
geodesic eccentricity transform—a study. In *Proc.
Workshop of the Austrian Association for Pattern

parametrizations by eigenfunctions of the Laplacian

Rotation invariant spherical harmonic representation
of 3D shape descriptors. In *Proc. SGP*, pages 156–164,
2003.


[22] D. Lowe. Distinctive image features from

[23] F. Memoli and G. Sapiro. A theoretical and
computational framework for isometry invariant
recognition of point cloud data. *Foundations of

Probabilistic fingerprints for shapes. In *Proc. SGP*,
2006.


[27] E. Paquet, M. Rioux, A. Munching, T. Naveen, and
A. Tabatabai. Description of shape information for
2-D and 3-D objects. *Signal Processing: Image

[28] D. Raviv, A. M. Bronstein, M. M. Bronstein, and


[31] R. M. Rustamov. Laplace-Beltrami eigenfunctions for
deformation invariant shape representation. In *Proc.

shape descriptors for similarity retrieval. In *Proc.
Shape Modelling and Applications*, 2006.

The Princeton shape benchmark. In *Proc. SMI*, pages

[34] J. Sivic and A. Zisserman. Video google: A text


and provably informative multi-scale signature based


[38] A. Tal, M. Elad, and S. Ar. Content based retrieval of
VRML objects - an iterative and interactive approach.

matching using geometry and photometry. In *Proc.
CVPR*, 2009.

vocabulary signature for 3D object retrieval and

Surface feature detection and description with

[42] C. Zhang and T. Chen. Efficient feature extraction for
2D/3D objects in mesh representation. In *Proc. ICIP*,