# Photometric Heat Kernel Signatures

Artiom Kovnatsky[1], Michael M. Bronstein[3],
Alexander M. Bronstein[4], and Ron Kimmel[2]

[1]Department of Applied Mathematics,
`art@tx.technion.ac.il`
[2]Department of Computer Science,
`ron@cs.technion.ac.il`
Technion, Israel Institute of Technology, Haifa, Israel

[3]Inst. of Computational Science, Faculty of Informatics,
Università della Svizzera Italiana, Lugano, Switzerland
`michael.bronstein@usi.ch`

[4]Dept. of Electrical Engineering, Tel Aviv University, Israel
`bron@eng.tau.ac.il`

**Abstract.** In this paper, we explore the use of the diffusion geometry framework for the fusion of geometric and photometric information in local heat kernel signature shape descriptors. Our construction is based on the definition of a diffusion process on the shape manifold embedded into a high-dimensional space where the embedding coordinates represent the photometric information. Experimental results show that such data fusion is useful in coping with different challenges of shape analysis where pure geometric and pure photometric methods fail.

## 1 Introduction

In last decade, the amount of geometric data available in the public domain, such as Google 3D Warehouse, has grown dramatically and created the demand for shape search and retrieval algorithms capable of finding similar shapes in the same way a search engine responds to text queries. However, while text search methods are sufficiently developed to be ubiquitously used, the search and retrieval of 3D shapes remains a challenging problem. Shape retrieval based on text metadata, like annotations and tags added by the users, is often incapable of providing relevance level required for a reasonable user experience.

*Content-based shape retrieval* using the shape itself as a query and based on the comparison of geometric and topological properties of shapes is complicated by the fact that many 3D objects manifest rich variability, and shape retrieval must often be *invariant* under different classes of transformations. A particularly challenging setting is the case of non-rigid shapes, including a wide range of transformations such as bending and articulated motion, rotation and translation, scaling, non-rigid deformation, and topological changes. The main challenge in shape retrieval algorithms is computing a *shape descriptor*, that would be unique for each shape, simple to compute and store, and invariant under

different type of transformations. Shape similarity is determined by comparing the shape descriptors.

Broadly, shape descriptors can be divided into *global* and *local*. The former consider global geometric or topological shape characteristics such as distance distributions [21,24,19], geometric moments [14,30], or spectra [23], whereas the latter describe the local behavior of the shape in a small patch. Popular examples of local descriptors include spin images [3], shape contexts [1], integral volume descriptors [12] and radius-normal histograms [22]. Using the bag of features paradigm common in image analysis [25,10], a global shape descriptor counting the occurrence of local descriptors in some vocabulary can be computed [7].

Recently, there has been an increased interest in the use of *diffusion geometry* [11,16] for constructing invariant shape descriptors. Diffusion geometry is closely related to heat propagation properties of shapes and allows obtaining global descriptors, such as distance distributions [24,19,8] and Laplace-Beltrami spectral signatures [23], as well local descriptors such as heat kernel signatures [26,9,7]. One limitation of these methods is that, so far, only *geometric* information has been considered. However, the abundance of textured models in computer graphics and modeling applications, as well as the advance in 3D shape acquisition [35,36] allowing to obtain textured 3D shapes of even moving objects, bring forth the need for descriptors also taking into consideration *photometric* information. Photometric information plays an important role in a variety of shape analysis applications, such as shape matching and correspondence [28,33]. Considering 2D views of the 3D shape [32,20], standard feature detectors and descriptors used in image analysis such as SIFT [18] can be employed. More recently, Zaharescu *et al.* [37] proposed a geometric SIFT-like descriptor for textured shapes, defined directly on the surface.

In this paper, we extend the diffusion geometry framework to include photometric information in addition to its geometric counterpart. The main idea is to define a diffusion process that takes into consideration not only the geometry but also the texture of the shape. This is achieved by considering the shape as a manifold in a higher dimensional combined geometric-photometric embedding space, similarly to methods in image processing applications [15,17]. As a result, we are able to construct geometric and photometric local descriptors (color heat kernel signatures or cHKS).

The rest of this paper is organized as follows. In Section 2, we review the mathematical formalism of diffusion processes and their use in shape analysis. In Section 3, we introduce our approach and in Section 4 its numerical implementation details. Section 5 presents experimental results. Finally, Section 6 concludes the paper.

## 2   Background

Throughout the paper, we assume the shape to be modeled as a two-dimensional compact Riemannian manifold $X$ (possibly with a boundary) equipped with a metric tensor $g$. Fixing a system of local coordinates on $X$, the latter can be

expressed as a $2 \times 2$ matrix $g_{\mu\nu}$, also known as the first fundamental form. The metric tensor allows to express the length of a vector $v$ in the tangent space $T_x X$ at a point $x$ as $g_{\mu\nu} v^\mu v^\nu$, where repeated indices $\mu, \nu = 1, 2$ are summed over following Einstein's convention.

Given a smooth scalar field $f : X \to \mathbb{R}$ on the manifold, its *gradient* is defined as the vector field $\nabla f$ satisfying $f(x + dx) = f(x) + g_x(\nabla f(x), dx)$ for every point $x$ and every infinitesimal tangent vector $dx \in T_x X$. The metric tensor $g$ defines the *Laplace-Beltrami operator* $\Delta_g$ that satisfies

$$\int f \Delta_g h \, da = - \int g_x(\nabla f, \nabla h) da \qquad (1)$$

for any pair of smooth scalar fields $f, h : X \to \mathbb{R}$; here $da$ denotes integration with respect to the standard area measure on $X$. Such an integral definition is usually known as the Stokes identity. The Laplace-Beltrami operator is positive semi-definite and self-adjoint. Furthermore, it is an *intrinsic* property of $X$, i.e., it is expressible solely in terms of $g$. In the case when the metric $g$ is Euclidean, $\Delta_g$ becomes the standard Laplacian.

The Laplace-Beltrami operator gives rise to the *heat equation*,

$$\left( \Delta_g + \frac{\partial}{\partial t} \right) u = 0, \qquad (2)$$

which describes diffusion processes and heat propagation on the manifold (note that we use a positive-semidefinite Laplace-Beltrami operator, hence the plus sign in the heat equation). Here, $u(x, t)$ denotes the distribution of heat at time $t$ at point $x$. The initial condition to the equation is some heat distribution $u(x, 0)$, and if the manifold has a boundary, appropriate boundary conditions (e.g. Neumann or Dirichlet) must be specified. The solution of (2) with a point initial heat distribution $u_0(x) = \delta(x, x')$, where $\delta(x', x') = 1$ *o.w.* 0, is called the *heat kernel* and denoted here by $K_t(x, x')$. Using a signal processing analogy, $K_t$ can be thought of as the "impulse response" of the heat equation.

By the spectral decomposition theorem, the heat kernel can be represented as [13]

$$K_t(x, x') = \sum_{i \geq 0} e^{-\lambda_i t} \phi_i(x) \phi_i(x'), \qquad (3)$$

where $0 = \lambda_0 \leq \lambda_1 \leq \ldots$ are the eigenvalues and $\phi_0, \phi_1, \ldots$ the corresponding eigenfunctions of the Laplace-Beltrami operator (i.e., solutions to $\Delta_g \phi_i = \lambda_i \phi_i$). We will collectively refer to quantities expressed in terms of the heat kernel as to *diffusion geometry*. Since the Laplace-Beltrami operator is intrinsic, the diffusion geometry it induces is invariant under isometric deformations of $X$ (incongruent embeddings of $g$ into $\mathbb{R}^3$).

Sun *et al.* [26] proposed using the heat propagation properties as a local descriptor of the manifold. The diagonal of the heat kernel, $K_t(x, x')$, referred to as the *heat kernel signature* (HKS), captures the local properties of $X$ at point $x$ and scale $t$. The descriptor is computed at each point as a vector of

the values $p(x) = (K_{t_1}(x, x), \ldots, K_{t_n}(x, x))$, where $t_1, \ldots, t_n$ are some time values. Such a descriptor is deformation-invariant, easy to compute, and provably informative [26].

Ovsjanikov *et al.* [7] employed the HKS local descriptor for large-scale shape retrieval using the *bags of features* paradigm [25]. In this approach, the shape is considered as a collection of "geometric words" from a fixed "vocabulary" and is described by the distribution of such words, also referred to as a *bag of features* or BoF. The vocabulary is constructed offline by clustering the HKS descriptor space. Then, for each point on the shape, the HKS is replaced by the nearest vocabulary word by means of vector quantization. Counting the frequency of each word, a BoF is constructed. The similarity of two shapes $X$ and $Y$ is then computed as the distance between the corresponding BoFs, $d(X, Y) = \|\mathrm{BoF}_X - \mathrm{BoF}_Y\|$.

## 3 Photometric Heat Kernel Signatures

Let us further assume that the Riemannian manifold $X$ is a submanifold of some manifold $\mathcal{E}$ ($\dim(\mathcal{E}) = m > 2$) with the Riemannian metric tensor $h$, embedded by means of a diffeomorphism $\xi : X \to \xi(X) \subseteq \mathcal{E}$. A Riemannian metric tensor on $X$ induced by the embedding is the *pullback metric* $(\xi^*h)(r, s) = h(d\xi(r), d\xi(s))$ for $r, s \in T_x X$, where $d\xi : T_x X \to T_{\xi(x)}\mathcal{E}$ is the differential of $\xi$. In coordinate notation, the pullback metric is expressed as $(\xi^*h)_{\mu\nu} = h_{ij}\partial_\mu\xi^i\partial_\nu\xi^j$, where the indices $i, j = 1, \ldots, m$ denote the embedding coordinates.

Here, we use the structure of $\mathcal{E}$ to model joint geometric and photometric information. Such an approach has been successfully used in image processing [15]. When considering shapes as geometric object only, we define $\mathcal{E} = \mathbb{R}^3$ and $h$ to be the Euclidean metric. In this case, $\xi$ acts as a *parametrization* of $X$ and the pullback metric becomes simply $(\xi^*h)_{\mu\nu} = \partial_\mu\xi^1\partial_\nu\xi^1 + \ldots + \partial_\mu\xi^3\partial_\nu\xi^3 = \langle\partial_\mu\xi, \partial_\nu\xi\rangle_{\mathbb{R}^3}$. In the case considered in this paper, the shape is endowed with photometric information given in the form of a field $\alpha : X \to \mathcal{C}$, where $\mathcal{C}$ denotes some colorspace (e.g., RGB or Lab). This photometric information can be modeled by defining $\mathcal{E} = \mathbb{R}^3 \times \mathcal{C}$ and an embedding $\xi = (\xi_g, \xi_p)$. The embedding coordinates corresponding to geometric information $\xi_g = (\xi^1, \ldots, \xi^3)$ are as previously and the embedding coordinate corresponding to photometric information are given by $\xi_p = (\xi^4, \ldots, \xi^6) = \eta(\alpha^1, \ldots, \alpha^3)$, where $\eta \geq 0$ is a scaling constant. Simplifying further, we assume $\mathcal{C}$ to have a Euclidean structure (for example, the Lab colorspace has a natural Euclidean metric). The metric in this case boils down to $(\xi^*h)_{\mu\nu} = \langle\partial_\mu\xi_g, \partial_\nu\xi_g\rangle_{\mathbb{R}^3} + \eta^2\langle\partial_\mu\xi_p, \partial_\nu\xi_p\rangle_{\mathbb{R}^3}$, which hereinafter we shall denote by $\hat{g}_{\mu\nu}$.[1]

---

[1] The joint metric tensor $\hat{g}$ has inherent ambiguities. The diffusion geometry induced by $\hat{g}$ is invariant the joint isometry group $\mathrm{Iso}_{\hat{g}} = \mathrm{Iso}((\xi_g^*h)_{\mu\nu} + \eta^2(\xi_p^*h)_{\mu\nu})$. Ideally, we would like $\mathrm{Iso}_{\hat{g}} = \mathrm{Iso}_g = \mathrm{Iso}((\xi_g^*h)_{\mu\nu}) \times \mathrm{Iso}_p = \mathrm{Iso}((\xi_p^*h)_{\mu\nu})$ to hold. In practice, $\mathrm{Iso}_{\hat{g}}$ is bigger: while every composition of a geometric isometry with a photometric isometry is a joint isometry, there exist some joint isometries which cannot be obtained as a composition of geometric and photometric isometries. Experimental results show that no realistic geometric and photometric transformations lie in $\mathrm{Iso}_{\hat{g}} \setminus (\mathrm{Iso}_g \times \mathrm{Iso}_p)$.

**Fig. 1.** Textured shape (left); values of the heat kernel ($x$ placed on the foot, $t = 1024$) arising from regular purely geometric (middle) and mixed photometric-geometric (right) diffusion process.

The Laplace-Beltrami operator $\Delta_{\hat{g}}$ associated with such a metric gives rise to diffusion geometry that combines photometric and geometric information (Figure 1). We define the *photometric* or *color heat kernel signature* (cHKS) as the diagonal of the heat kernel associated with the joint geometric-photometric diffusion induced by $\Delta_{\hat{g}}$. The cHKS fuses local geometric and photometric information of the shape.

## 4    Numerical Implementation

Let $\{x_1, \ldots, x_N\} \subseteq X$ denote the discrete samples of the shape, and $\xi(x_1), \ldots, \xi(x_N)$ be the corresponding embedding coordinates (three-dimensional in the case we consider only geometry, or six-dimensional in the case of geometry-photometry fusion). We further assume to be given a *triangulation* (simplicial complex), consisting of *edges* $(i, j)$ and *faces* $(i, j, k)$ where each $(i, j), (j, k)$, and $(i, k)$ is an edge (here $i, j, k = 1, \ldots, N$).

A function $f$ on the discretized manifold is represented as an $N$-dimensional vector $(f(x_1), \ldots, f(x_N))$. The discrete Laplace-Beltrami operator can be written in the generic form

$$(\hat{\Delta}f)(x_i) = \frac{1}{a_i} \sum_{j \in \mathcal{N}_i} w_{ij}(f(x_i) - f(x_j)), \qquad (4)$$

where $w_{ij}$ are weights, $a_i$ are normalization coefficients, and $\mathcal{N}_i$ denotes a local neighborhood of point $i$. Different discretizations of the Laplace-Beltrami operator can be cast into this form by appropriate definition of the above constants. For shapes represented as triangular meshes, a widely-used method is the *cotangent scheme*, which preserves many important properties of the continuous Laplace-Beltrami operator, such as positive semi-definiteness, symmetry, and locality [31]. Yet, in general, the cotangent scheme does not converge to the continuous Laplace-Beltrami operator, in the sense that the solution of the discrete eigenproblem does not converge to the continuous one (pointwise convergence exists if the triangulation and sampling satisfy certain conditions [34]).

Belkin *et al.* [5] proposed a discretization which is convergent without the restrictions on "good" triangulation required by the cotangent scheme. In this scheme, $\mathcal{N}_i$ is chosen to be the entire sampling $\{x_1, \ldots, x_N\}$, $a_i = \frac{1}{4\pi\rho^2}$, and $w_{ij} = S_j e^{-\|\xi(x_i)-\xi(x_j)\|^2/4\rho}$, where $\rho$ is a parameter, $S_j$ denotes area of all triangles sharing the vertex $j$. In the case of a Euclidean colorspace, $w_{ij}$ can be written explicitly as

$$w_{ij} = S_j \exp\left\{-\frac{\|\xi_g(x_i) - \xi_g(x_j)\|^2}{4\rho} - \frac{\|\xi_p(x_i) - \xi_p(x_j)\|^2}{4\sigma}\right\} \tag{5}$$

where $\sigma = \rho/\eta^2$, which resembles the weights used in the *bilateral filter* [29]. Experimental results also show that this operator produces accurate approximation of the Laplace-Beltrami operator under various conditions, such as noisy data input and different sampling [27,5].

In matrix notation, equation (4) can be written as $\hat{\Delta}f = A^{-1}Wf$, where $A = \mathrm{diag}(a_i)$ and $W = \mathrm{diag}\left(\sum_{l \neq i} w_{il}\right) - (w_{ij})$. The eigenvalue problem $\hat{\Delta}\Phi = \Lambda\Phi$ is equivalent to the generalized symmetric eigenvalue problem $W\Phi = \Lambda A\Phi$, where $\Lambda = \mathrm{diag}(\lambda_0, \ldots, \lambda_K)$ is the diagonal matrix of the first $K$ eigenvalues, and $\Phi = (\phi_0, \ldots, \phi_K)$ is the matrix of the eigenvectors stacked as columns. Since typically $W$ is sparse, this problem can be efficiently solved numerically. Heat kernels can be approximated by taking the first largest eigenvalues and the corresponding eigenfunctions in (3). Since the coefficients in the expansion of $h_t$ decay as $\mathcal{O}(e^{-t})$, typically a few eigenvalues ($K$ in the range of 10 to 100) are required.

## 5   Results

In order to evaluate the proposed method, we used the SHREC 2010 robust large-scale shape retrieval benchmark methodology [6]. The query set consisted of 270 real-world human shapes from 5 classes acquired by a 3D scanner with real geometric transformations and simulated photometric transformations of different types and strengths, totalling in 54 instances per shape (Figure 2). Geometric transformations were divided into *isometry+topology* (real articulations and topological changes due to acquisition imperfections), and *partiality*

(occlusions and addition of clutter such as the red ball in Figure 2). Photometric transformations included *contrast* (increase and decrease by scaling of the $L$ channel), *brightness* (brighten and darken by shift of the $L$ channel), *hue* (shift in the $a$ channel), *saturation* (saturation and desaturation by scaling of the $a, b$ channels), and *color noise* (additive Gaussian noise in all channels). *Mixed* transformations included isometry+topology transformations in combination with two randomly selected photometric transformations. In each class, the transformation appeared in five different versions numbered 1–5 corresponding to the transformation strength levels. One shape of each of the five classes was added to the queried corpus in addition to other 75 shapes used as clutter (Figure 3).

Retrieval was performed by matching 270 transformed queries to the 75 null shapes. Each query had exactly one correct corresponding null shape in the dataset. Performance was evaluated using the precision-recall characteristic. *Precision* $P(r)$ is defined as the percentage of relevant shapes in the first $r$ top-ranked retrieved shapes. *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. Intuitively, mAP is interpreted as the area below the precision-recall curve. Ideal retrieval performance results in first relevant match with mAP=100%. Performance results were broken down according to transformation class and strength.



**Fig. 2.** Examples of geometric and photometric shape transformations used as queries (shown at strength 5). First row, left to right: null, isometry+topology, partiality, two brightness transformations (brighten and darken), two contrast transformations (increase and decrease contrast). Second row, left to right: two saturation transformations (saturate and desaturate), hue, color noise, mixed.
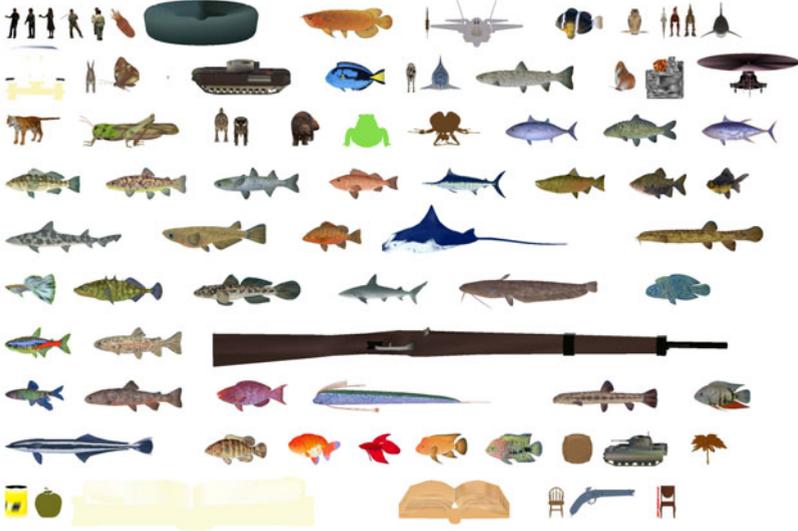
**Fig. 3.** Null shapes in the dataset (shown at arbitrary scale for visualization purposes)

In additional to the proposed approach, we compared purely geometric, purely photometric, and joint photometric-geometric descriptors. As a purely geometric descriptor, we used bags of features based on HKS according to [7]; purely photometric shape descriptor was a color histogram. As joint photometric-geometric descriptors, we used bags of features computed with the MeshHOG [37] and the proposed color HKS (cHKS).

For the computation of the bag of features descriptors, we used the Shape Google framework with most of the settings as proposed in [7]. More specifically, HKS were computed at six scales ($t = 1024, 1351.2, 1782.9, 2352.5$, and $4096$). Soft vector quantization was applied with variance taken as twice the median of all distances between cluster centers. Approximate nearest neighbor method [2] was used for vector quantization. The Laplace-Beltrami operator discretization was computed using the Mesh-Laplace scheme [4] with scale parameter $\rho = 2$. Heat kernels were approximated using the first 200 eigenpairs of the discrete Laplacian. The MeshHOG descriptor was computed at prominent feature points (typically 100-2000 per shape), detected using the MeshDOG detector [37]. The vocabulary size in all the cases was set to 48.

In cHKS, in order to avoid the choice of an arbitrary value $\eta$, we used a set of three different weights ($\eta = 0, 0.05, 0.1$) to compute the cHKS and the corresponding BoFs. The distance between two shapes was computed as the sum of the distances between the corresponding BoFs for each $\eta$, weighted by $\eta$, and 1 in case of $\eta = 0$, $d(X,Y) = \|\mathrm{BoF}_X^0 - \mathrm{BoF}_Y^0\|_1^2 + \sum_\eta \eta \|\mathrm{BoF}_X^\eta - \mathrm{BoF}_Y^\eta\|_1^2$.

Tables 1–4 summarize the results of our experiments. Geometry only descriptor (HKS) [7] is invariant to photometric transformations, but is somewhat sensitive to topological noise and missing parts (Table 1). On the other hand, the

**Table 1.** Performance (mAP in %) of ShapeGoogle using BoFs with HKS descriptors

| Transform. | Strength | | | | |
|---|---|---|---|---|---|
| | **1** | **≤2** | **≤3** | **≤4** | **≤5** |
| *Isom+Topo* | 100.00 | 100.00 | 96.67 | 95.00 | 90.00 |
| *Partial* | 66.67 | 60.42 | 63.89 | 63.28 | 63.63 |
| *Contrast* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Brightness* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Hue* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Saturation* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Noise* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Mixed* | 90.00 | 95.00 | 93.33 | 95.00 | 96.00 |

**Table 2.** Performance (mAP in %) of color histograms

| Transform. | Strength | | | | |
|---|---|---|---|---|---|
| | **1** | **≤2** | **≤3** | **≤4** | **≤5** |
| *Isom+Topo* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Partial* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Contrast* | 100.00 | 90.83 | 80.30 | 71.88 | 63.95 |
| *Brightness* | 88.33 | 80.56 | 65.56 | 53.21 | 44.81 |
| *Hue* | 11.35 | 8.38 | 6.81 | 6.05 | 5.49 |
| *Saturation* | 17.47 | 14.57 | 12.18 | 10.67 | 9.74 |
| *Noise* | 100.00 | 100.00 | 93.33 | 85.00 | 74.70 |
| *Mixed* | 28.07 | 25.99 | 20.31 | 17.62 | 15.38 |

**Table 3.** Performance (mAP in %) of BoFs using MeshHOG descriptors

| Transform. | Strength | | | | |
|---|---|---|---|---|---|
| | **1** | **≤2** | **≤3** | **≤4** | **≤5** |
| *Isom+Topo* | 100.00 | 95.00 | 96.67 | 94.17 | 95.33 |
| *Partial* | 75.00 | 61.15 | 69.93 | 68.28 | 68.79 |
| *Contrast* | 100.00 | 100.00 | 100.00 | 98.33 | 94.17 |
| *Brightness* | 100.00 | 100.00 | 100.00 | 100.00 | 99.00 |
| *Hue* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Saturation* | 100.00 | 100.00 | 100.00 | 98.75 | 99.00 |
| *Noise* | 100.00 | 100.00 | 88.89 | 83.33 | 78.33 |
| *Mixed* | 100.00 | 100.00 | 100.00 | 93.33 | 83.40 |

color-only descriptor works well only for geometric transformations that do not change the shape color. Photometric transformations, however, make such a descriptor almost useless (Table 2). MeshHOG is almost invariant to photometric transformations being based on texture gradients, but is sensitive to color noise (Table 3). The fusion of the geometric and photometric data using our approach (Table 4) achieves nearly perfect retrieval for mixed and photometric transformations and outperforms other approaches. Figure 4 visualizes a few examples of the retrieved shapes ordered by relevance, which is inversely proportional to the distance from the query shape.

**Table 4.** Performance (mAP in %) of ShapeGoogle using w-multi-scale BoFs with cHKS descriptors

|              |        | Strength |        |        |        |
| ------------ | ------ | ------ | ------ | ------ | ------ |
| Transform.   | 1      | ≤2     | ≤3     | ≤4     | ≤5     |
| *Isom+Topo*  | 100.00 | 100.00 | 96.67  | 97.50  | 94.00  |
| *Partial*    | 68.75  | 68.13  | 69.03  | 67.40  | 67.13  |
| *Contrast*   | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Brightness* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Hue*        | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Saturation* | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Noise*      | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| *Mixed*      | 100.00 | 100.00 | 96.67  | 97.50  | 98.00  |



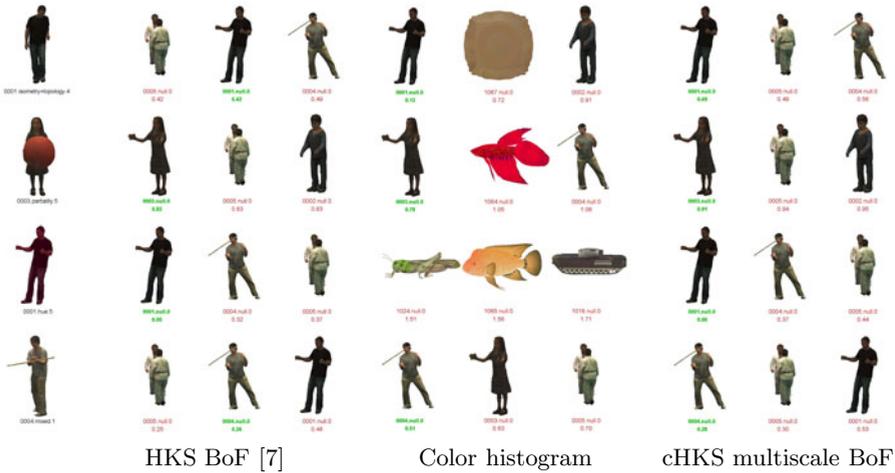HKS BoF [7]          Color histogram          cHKS multiscale BoF

**Fig. 4.** Retrieval results using different methods. First column: query shapes, second column: first three matches obtained with HKS-based BoF [7], third column: first three matches obtained using color histograms, fourth column: first three matches obtained with the proposed method (cHKS-based multiscale BoF). Shape annotation follows the convention *shapeid.transformation.strength*; numbers below show distance from query. Only a single correct match exists in the database (marked in green), and ideally, it should be the first one.

## 6   Conclusions

In this paper, we explored a way to fuse geometric and photometric information in the construction of shape descriptors. Our approach is based on heat propagation on a manifold embedded into a combined geometry-color space. Such diffusion processes capture both geometric and photometric information and give rise to local and global diffusion geometry (heat kernels and diffusion distances), which can be used as informative shape descriptors. We showed experimentally

that the proposed descriptors outperform other geometry-only and photometry-only descriptors, as well as state-of-the-art joint geometric-photometric descriptors. In the future, it would be important to formally characterize the isometry group induced by the joint metric in order to understand the invariant properties of the proposed diffusion geometry, and possibly design application-specific invariant descriptors.

# References

1. Amores, J., Sebe, N., Radeva, P.: Context-based object-class recognition and retrieval by generalized correlograms. Trans. PAMI 29(10), 1818–1833 (2007)
2. Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R., Wu, A.Y.: An optimal algorithm for approximate nearest neighbor searching. J. ACM 45, 891–923 (1998)
3. Assfalg, J., Bertini, M., Bimbo, A.D., Pala, P.: Content-based retrieval of 3-d objects using spin image signatures. IEEE Transactions on Multimedia 9(3), 589–599 (2007)
4. Belkin, M., Sun, J., Wang, Y.: Constructing Laplace operator from point clouds in Rd. In: Proc. Symp. Discrete Algorithms, pp. 1031–1040 (2009)
5. Belkin, M., Sun, J., Wang, Y.: Discrete Laplace operator on meshed surfaces. In: Proc. Symp. Computational Geometry, pp. 278–287 (2009)
6. Bronstein, A.M., Bronstein, M.M., Castellani, U., Falcidieno, B., Fusiello, A., Godil, A., Guibas, L.J., Kokkinos, I., Lian, Z., Ovsjanikov, M., Patané, G., Spagnuolo, M., Toldo, R.: Shrec 2010: robust large-scale shape retrieval benchmark. In: Proc. 3DOR (2010)
7. Bronstein, A.M., Bronstein, M.M., Ovsjanikov, M., Guibas, L.J.: Shape google: a computer vision approach to invariant shape retrieval. In: Proc. NORDIA (2009)
8. Bronstein, M.M., Bronstein, A.M.: Shape recognition with spectral distances. Trans. PAMI (2010) (to appear)
9. Bronstein, M.M., Kokkinos, I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In: Proc. CVPR (2010)
10. Chum, O., Philbin, J., Sivic, J., Isard, M., Zisserman, A.: Total recall: Automatic query expansion with a generative feature model for object retrieval. In: Proc. ICCV (2007)
11. Coifman, R.R., Lafon, S.: Diffusion maps. Applied and Computational Harmonic Analysis 21, 5–30 (2006)
12. Gelfand, N., Mitra, N.J., Guibas, L.J., Pottmann, H.: Robust global registration. In: Proc. SGP (2005)
13. Jones, P.W., Maggioni, M., Schul, R.: Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels. PNAS 105(6), 1803 (2008)
14. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Proc. SGP, pp. 156–164 (2003)
15. Kimmel, R., Malladi, R., Sochen, N.: Images as embedded maps and minimal surfaces: movies, color, texture, and volumetric medical images. IJCV 39(2), 111–129 (2000)

16. Lévy, B.: Laplace-Beltrami eigenfunctions towards an algorithm that understands geometry. In: Proc. Shape Modeling and Applications (2006)
17. Ling, H., Jacobs, D.W.: Deformation invariant image matching. In: ICCV, pp. 1466–1473 (2005)
18. Lowe, D.: Distinctive image features from scale-invariant keypoint. IJCV (2004)
19. Mahmoudi, M., Sapiro, G.: Three-dimensional point cloud recognition via distributions of geometric distances. Graphical Models 71(1), 22–31 (2009)
20. Ohbuchi, R., Osada, K., Furuya, T., Banno, T.: Salient local visual features for shape-based 3d model retrieval, pp. 93–102 (June 2008)
21. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. TOG 21(4), 807–832 (2002)
22. Pan, X., Zhang, Y., Zhang, S., Ye, X.: Radius-normal histogram and hybrid strategy for 3d shape retrieval, pp. 372–377 (June 2005)
23. Reuter, M., Wolter, F.-E., Peinecke, N.: Laplace-spectra as fingerprints for shape matching. In: Proc. ACM Symp. Solid and Physical Modeling, pp. 101–106 (2005)
24. Rustamov, R.M.: Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In: Proc. SGP, pp. 225–233 (2007)
25. Sivic, J., Zisserman, A.: Video google: A text retrieval approach to object matching in videos. In: Proc. CVPR (2003)
26. Sun, J., Ovsjanikov, M., Guibas, L.J.: A concise and provably informative multi-scale signature based on heat diffusion. In: Proc. SGP (2009)
27. Thangudu, K.: Practicality of Laplace operator (2009)
28. Thorstensen, N., Keriven, R.: Non-rigid shape matching using geometry and photometry. In: Proc. CVPR (2009)
29. Tomasi, C., Manduchi, R.: Bilateral fitering for gray and color images. In: Proc. ICCV, pp. 839–846 (1998)
30. Vranic, D.V., Saupe, D., Richter, J.: Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics. In: Proc. Workshop Multimedia Signal Processing, pp. 293–298 (2001)
31. Wardetzky, M., Mathur, S., Kälberer, F., Grinspun, E.: Discrete Laplace operators: no free lunch. In: Conf. Computer Graphics and Interactive Techniques (2008)
32. Wu, C., Clipp, B., Li, X., Frahm, J.-M., Pollefeys, M.: 3d model matching with viewpoint-invariant patches (vip), pp. 1–8 (June 2008)
33. Wyngaerd, J.V.: Combining texture and shape for automatic crude patch registration, pp. 179–186 (October 2003)
34. Xu, G.: Convergence of discrete Laplace-Beltrami operators over surfaces. Technical report, Institute of Computational Mathematics and Scientific/Engineering Computing, China (2004)
35. Yoon, K.-J., Prados, E., Sturm, P.: Joint estimation of shape and reflectance using multiple images with known illumination conditions (2010)
36. Zaharescu, A., Boyer, E., Horaud, R.P.: Transformesh: a topology-adaptive mesh-based approach to surface evolution (November 2007)
37. Zaharescu, A., Boyer, E., Varanasi, K.: R Horaud. Surface feature detection and description with applications to mesh matching. In: Proc. CVPR (2009)