

# Affine-Invariant Photometric Heat Kernel Signatures

Artiom Kovnatsky<sup>1</sup> Michael M. Bronstein<sup>1</sup> Alexander M. Bronstein<sup>2</sup> Dan Raviv<sup>3</sup> Ron Kimmel<sup>3</sup>

<sup>1</sup>Institute of Computational Science, Faculty of Informatics, Università della Svizzera Italiana, Lugano, Switzerland

<sup>2</sup>School of Electrical Engineering, Tel Aviv University, Israel

<sup>3</sup>Department of Computer Science, Technion, Israel Institute of Technology, Haifa, Israel

---

## Abstract

*In this paper, we explore the use of the diffusion geometry framework for the fusion of geometric and photometric information in local shape descriptors. Our construction is based on the definition of a modified metric, which combines geometric and photometric information, and then the diffusion process on the shape manifold is simulated. Experimental results show that such data fusion is useful in coping with shape retrieval experiments, where pure geometric and pure photometric methods fail. Apart from retrieval task the proposed diffusion process may be employed in other applications.*

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Laplace-Beltrami operator—, diffusion equation, heat kernel descriptors, 3D shape retrieval, deformation invariance

---

## 1. Introduction

The birth of the World Wide Web and the explosive growth of text content has brought the need to organize, index, and search text document, which in turn fueled the development of text search engines. In the past decade, the amount of geometric data available in the public-domain repositories 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. 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.

A common paradigm used in computer vision [SZ03, CPS\*07] is to start with local feature descriptors and aggregate them into a global shape descriptor using the bag of features approach [BBOG09, TCF09]. Popular examples of local descriptors include spin images [ABBP07], shape contexts [ASR07], integral volume descriptors [GMGP05] and radius-normal histograms [PZZY05].

Recently, a family of intrinsic geometric properties broadly known as *diffusion geometry* has become growingly popular. The studies of diffusion geometry are based on the theoretical works by Berard *et al.* [BBG94] and later by Coifman and Lafon [CL06]. Diffusion geometry offers an intuitive interpretation of many shape properties in terms of spatial frequencies and allows to use standard harmonic analysis tools. Also, recent advances in the discretization of the Laplace-Beltrami operator bring forth efficient and robust numerical and computational tools.

One of the first principled practical uses of these methods in the context of shape processing was explored by Lévy [Lév06]. Several attempts have also been made to construct feature descriptors based on diffusion geometric properties

of the shape. Rustamov [Rus07] proposed to construct the global point signature (GPS) feature descriptors closely resembling a diffusion map [CL06]. Fang *et al.* [FSR11] define a temperature distribution descriptor (TD), based on evaluation of temperature distribution after applying a unit heat at each vertex.

Sun *et al.* [SOG09a] introduced the *heat kernel signature* (HKS), based on the fundamental solutions of the heat equation (heat kernels). Scale invariant [BK10], affine-invariant [RBB\*11a, RBB\*], and volumetric [RBBK10] versions of the HKS were subsequently proposed. By applying topology persistence [ELZ00] algorithm on HKS descriptors at some predefined scale, Dey *et al.* [DLL\*10] obtained robust feature points, which are used for shape matching and retrieval. In [ASC11], another physically-inspired descriptor, the *wave kernel signature* (WKS) was proposed as a remedy to the excessive sensitivity of the HKS to low-frequency information. In [Bro], a general family of learnable spectral descriptors generalizing the HKS and WKS was introduced. On existing methods for 3D shape retrieval interested reader referred to a surveys [TV04, IJL\*05].

A major 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 [YPS10, ZBH07] allowing to obtain textured 3D shapes of even moving objects, bring forth the need for descriptors also taking into consideration *photometric* information.

In this paper, we extend the diffusion geometry framework to include photometric information in addition to its geometric counterpart. This way, we incorporate important photometric properties on the one hand, while exploiting a principled and theoretically established approach on the other. The main idea is to define a diffusion process on a manifold in a higher dimensional combined geometric-photometric embedding space, similarly to methods in image processing applications [KMS00, LJ05]. As a result, we are able to construct local descriptors (heat kernel signatures) that incorporate both geometric and photometric data. The proposed data fusion can be useful in coping with different challenges of shape analysis where pure geometric and pure photometric methods fail.

Preliminary results of this study introducing photometric HKS descriptors with Euclidean metric on the photometric space have been published in [KBBK11]. Here, we consider a more generic affine-invariance metric, which is invariant to many important photometric transformations.

## 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*  $g$ . Fix-

ing 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  $f : X \rightarrow \mathbb{R}$  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

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

for any pair of smooth scalar fields  $f, h : X \rightarrow \mathbb{R}$ ; here  $da$  denotes integration with respect to the standard area measure on  $X$ . Such an integral definition is 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*,

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

which describes diffusion processes and heat propagation on the manifold. 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  $u_0(x) = \delta(x, x')$  is called the *heat kernel* and denoted here by  $k_t(x, x')$ .

By virtue of the spectral theorem, there exists an orthonormal basis on  $L_2(X)$  consisting of the eigenfunctions  $\phi_0, \phi_1, \dots$  of the Laplace-Beltrami operator (i.e., solutions to  $\Delta_g \phi_i = \lambda_i \phi_i$ , where  $0 = \lambda_0 \leq \lambda_1 \leq \dots$  are the corresponding eigenvalues). This basis can be interpreted analogously to the Fourier basis, and the eigenvalues  $\lambda_i$  as frequencies. Consequently, the heat kernel can be represented as [JMS08]

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

Since the Laplace-Beltrami operator is intrinsic, the diffusion geometry it induces is invariant under isometric deformations of  $X$  (incongruent embeddings of  $g$  into  $\mathbb{R}^3$ ).

## 3. Fusion of geometric and photometric data

The main idea of this paper is to create a modified diffusion operator that combines geometric and photometric properties of the shape by means of definition of a new metric

tensor (and hence the Laplace-Beltrami operator). In modified diffusion process the heat will flow proportionally to changes of color. For this purpose, 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  $\xi : X \rightarrow \mathcal{E}$ . 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  $d\xi : T_x X \rightarrow T_{\xi(x)} \mathcal{E}$  is the differential of  $\xi$ . In coordinate notation, the pullback metric is expressed as  $\hat{g}_{\mu\nu} = (\xi^*h)_{\mu\nu} = h_{ij} \partial_\mu \xi^i \partial_\nu \xi^j$ , where the indices  $i, j = 1, \dots, m$  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 [KMS00]. 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  $(\xi^*h)_{\mu\nu} = \partial_\mu \xi^1 \partial_\nu \xi^1 + \dots + \partial_\mu \xi^3 \partial_\nu \xi^3 = \langle \partial_\mu \xi, \partial_\nu \xi \rangle_{\mathbb{R}^3}$ . In the case considered in this paper, the shape is endowed with photometric information given in the form of a field  $\alpha : X \rightarrow \mathcal{C}$ , where  $\mathcal{C}$  denotes some colorspace (e.g., RGB or Lab). In the following, when required, we tacitly assume that  $\alpha$  is sufficiently smooth.

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 are as before  $\xi_g = (\xi^1, \dots, \xi^3)$ , and the embedding coordinate corresponding to photometric information are given by  $\xi_p = (\xi^4, \dots, \xi^6) = \eta(\alpha^1, \dots, \alpha^3)$ , where  $\eta \geq 0$  is a scaling constant. In addition to trading off between geometry and photometry parts, the scaling constant  $\eta$  has another role of resolving ambiguities of new isometry group, as discussed later in Section 3.3. The Laplace-Beltrami operator  $\Delta_{\hat{g}}$  associated with such a metric gives rise to diffusion geometry that combines photometric and geometric information.

### 3.1. Euclidean color metric

The invariance to different classes of photometric transformations is obtained by selecting the structure of the colorspace  $\mathcal{C}$ . In the simplest case, we assume  $\mathcal{C}$  to have a Euclidean structure.

While being the simplest choice, the Euclidean metric is known to be perceptually meaningful in some colorspace such as the ‘‘color opponent’’ CIE Lab space intended to mimic the nonlinear response of the eye [Jai89]. The photometric coordinates  $\xi_p = (L, a, b)$  in this colorspace represent *lightness* and color differences:  $a$  varies from green to red, and  $b$  varies from blue to yellow. Isometries with respect to the Euclidean metric in the Lab colorspace are shifts (resulting in lightening and hue transformations) and rotations,

$$\xi_p = R \xi_p + c, \quad (4)$$

where  $R$  denotes a  $3 \times 3$  rotation matrix, and  $c$  is a  $3 \times 1$  shift

vector. Such transformations capture many natural color changes the shape can undergo (in Figure 2 two brightness transformations, hue and equi-affine transformations are like this).

The joint metric in this case boils down to  $(\xi^*g)_{\mu\nu} = \langle \partial_\mu \xi_g, \partial_\nu \xi_g \rangle_{\mathbb{R}^3} + \eta \langle \partial_\mu \xi_p, \partial_\nu \xi_p \rangle_{\mathbb{R}^3}$ .

### 3.2. Affine-invariant color metric

A more generic class of photometric transformations can be expressed as *affine transformations* in the Lab colorspace  $\xi_p = A \xi_p + c$ , where  $A$  is an invertible  $3 \times 3$  matrix. In particular, the transformation is called *equi-affine* if  $\det(A) = 1$ .

Raviv *et al.* [RBB\*11a, RBB\*] showed a construction of a metric that is invariant to equi-affine transformations. In our setting, let us be given some parametrization  $\phi(u_1, u_2) : U \subseteq \mathbb{R}^2 \rightarrow X \subset \mathbb{R}^3$  of the shape; the composition of  $\alpha \circ \phi$  gives us a parametrization of the texture. First, allowing some relaxed notation, we denote by  $g_X(u_1, u_2) = \langle \partial_{u_\mu}, \partial_{u_\nu} \rangle$  and  $g_{\alpha(X)}(u_1, u_2) = \langle d\alpha(\partial_{u_\mu}), d\alpha(\partial_{u_\nu}) \rangle$  the first fundamental forms of  $X$  and  $\alpha(X)$ , respectively in matrix representations at point  $\phi(u_1, u_2)$  in our parametrization. Here,  $d\alpha$  is the differential of  $\alpha$  and  $d\alpha(\partial_{u_\mu}) = \frac{\partial}{\partial u_\mu}(\alpha \circ \phi)$ . Second, construct an equi-affine *pre-metric tensor* [Soc01, RBB\*]

$$(\tilde{g}_X(u_1, u_2))_{\mu\nu} = \tilde{g}_{\mu\nu} \det^{-1/4}(\tilde{g}), \quad (5)$$

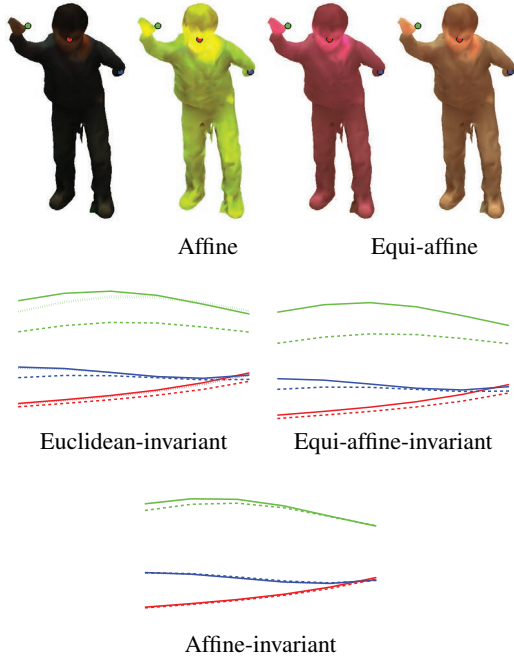
where  $\tilde{g}_{\mu\nu} = \det(d\alpha(\partial_{u_1}), d\alpha(\partial_{u_2}), (\alpha \circ \phi)_{u_\mu u_\nu})$ . Such a normalization tacitly implies that the Gaussian curvature is non-vanishing, otherwise the pre-metric tensor is not defined. Finally, the metric tensor is obtained by forcing  $\tilde{g}$  to have positive eigenvalues. For additional details about derivation and proof of affine invariance, we refer the reader to [RBB\*11a, RBB\*, ARK11, RK12].

The modified geometry and photometry metric tensor with the equi-affine-invariant photometric component is defined in matrix representation with respect to  $(u_1, u_2)$  on  $X$  as

$$\hat{g}(u_1, u_2) = g_X(u_1, u_2) + \eta g_{\alpha(X)}(u_1, u_2). \quad (6)$$

It is possible to use other metrics on the color coordinates (Fig 1). In [ARK11] the authors presented a scale invariant metric by normalizing the induced Euclidean metric according to the Gaussian curvature. This approach provides a new intrinsic distance measurement, which is different than the Euclidean one, but is invariant to local (piecewise linear) scaling. Motivated by [ARK11] the authors of [RK12] detached the scale normalization from the metric, and showed that the equi-affine invariant metric can be further improved and cope the affine group of transformations (similarity and equi-affine) while remaining invariant to non-rigid transformations.

The Gaussian curvature is defined by the ratio between the determinants of the second and the first fundamental forms



**Figure 1:** The influence of metric selection. First row: three photometric transformations of a textured shape. Second and third rows: the value of the heat kernel diagonal  $h_t(x, x)$  at three different points (marked with red, green, and blue on the shape) for different transformations (solid line: null, dashed: affine, dotted: equi-affine). The heat kernel is constructed using Euclidean and equi-affine-invariant and affine-invariant metrics in the colorspace. The more the curves coincide, the better is invariance.

and is denoted by  $K$ . [RK12] showed that it is possible to compute the Gaussian curvature of the equi-affine invariant metric, and construct a new metric by multiplying the metric elements by  $|K|$ . Specifically, consider the surface  $(X, \tilde{g})$ , where  $\tilde{g}_{ij}$  is the equi-affine invariant metric, and compute the Gaussian curvature  $K_{(X, \tilde{g})}(x)$  at each point. The affine invariant metric is defined by

$$g_{ij} = \left| K_{(X, \tilde{g})}(x) \right| \tilde{g}_{ij}. \quad (7)$$

### 3.3. Invariance of the joint diffusion process

The joint metric tensor  $\hat{g}$  and the diffusion geometry it induces have inherent ambiguities. Let us denote by  $\text{Iso}_g = \text{Iso}((\xi_g^* h)_{\mu\nu})$  and  $\text{Iso}_p = \text{Iso}((\xi_p^* h)_{\mu\nu})$  the respective groups of transformation that leave the geometric and the photometric components of the shape unchanged. We will refer to such transformations as geometric and photometric isometries. The diffusion metric induced by  $\hat{g}$  is invariant to joint isometry group  $\text{Iso}_{\hat{g}} = \text{Iso}((\xi_g^* h)_{\mu\nu} + \eta^2 (\xi_p^* h)_{\mu\nu})$ . Ideally, we would like  $\text{Iso}_{\hat{g}} = \text{Iso}_g \times \text{Iso}_p$  to hold. In practice,  $\text{Iso}_{\hat{g}}$  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.

An example of such transformations is uniform scaling of  $(\xi_g^* h)_{\mu\nu}$  combined with compensating scaling of  $(\xi_p^* h)_{\mu\nu}$ .

It is possible to overcome the ambiguity problem by considering metrics with different values of the scaling factor  $\eta$ . This rules out the compensating scaling situation and ensures that the shapes appear isometric for all values of  $\eta$  only if their geometric and photometric components are isometric.

## 4. Photometric heat kernel signatures

Sun *et al.* [SOG09b] and independently Gebal *et al.* [GBAL09] proposed using the heat propagation properties as a local descriptor of the manifold. The diagonal of the heat kernel,

$$k_t(x, x) = \sum_{i \geq 0} e^{-\lambda_i t} \phi_i^2(x), \quad (8)$$

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  $p(x) = (k_{t_1}(x, x), \dots, k_{t_n}(x, x))$ , where  $t_1, \dots, t_n$  are some time values. The resulting  $n$ -dimensional descriptor is deformation-invariant, easy to compute, and provably informative.

Ovsjanikov *et al.* [BBOG09] employed the HKS local descriptor for large-scale shape retrieval using the *bags of features* paradigm [SZ03]. In this approach, the shape is considered as a collection of “geometric words” from a fixed “vocabulary”  $\{p_1, \dots, p_q\} \subset \mathbb{R}^n$  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  $x$  on the shape, the HKS  $p(x)$  is replaced by the nearest vocabulary word by means of vector quantization,

$$\theta(x) = (\theta_1(x), \dots, \theta_q(x)) = \frac{e^{-\|p(x) - p_i\|_2^2 / 2\sigma^2}}{\sum_{i=1}^q e^{-\|p(x) - p_i\|_2^2 / 2\sigma^2}}, \quad (9)$$

where  $\sigma$  controls the “softness” of quantization. The BOF is constructed as

$$b_X = \int_X \theta(x) da \quad (10)$$

and can be considered as the frequency of different geometric words. The similarity of two shapes  $X$  and  $Y$  is then computed as the distance between the corresponding BOFs,  $d(X, Y) = \|b_X - b_Y\|$ .

Using the proposed approach, we define the *color heat kernel signature* (cHKS), defined in the same way as HKS with the standard Laplace-Beltrami operator replaced by the one resulting from the geometric-photometric embedding.

The photometric information is represented in the Lab color space with the Euclidean, equi-affine or affine-invariant metric.

As discussed in Section 3.3, in order to avoid ambiguities related to the joint metric, we have to compute the cHKS descriptor with multiple values of the scaling parameter  $\eta$ , each value producing a different set of cHKS descriptors  $p_\eta(x)$  and corresponding bags of features  $b_{X,\eta}$ . This set of BOFs can be compared e.g. as

$$d(X, Y) = \sum_{\eta \in H} \eta \|b_{X,\eta} - b_{Y,\eta}\|. \quad (11)$$

## 5. Numerical implementation

Let  $\{x_1, \dots, x_N\} \subseteq X$  denote the discrete samples of the shape, and  $\xi(x_1), \dots, \xi(x_N)$  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  $i, j, k = 1, \dots, N$ ).

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

$$(\hat{\Delta}f)(x_i) = \frac{1}{a_i} \sum_{j \in \mathcal{N}_i} w_{ij}(f(x_i) - f(x_j)), \quad (12)$$

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: a widely-used method is the *cotangent scheme* [WMKG08, DMSB99], Belkin's *et al.* [BSW09b] Mesh Laplacian discretization.

For computation of the spectrum of Laplace-Beltrami operator the finite elements methods (FEM) may be adopted as well. This approach is more suitable for our calculations, since we work with metric tensors, and the spectrum is sufficient for further processing. Considering the FEM, the eigenvalue problem is formulated in weak form [Dzi87]:

$$\Delta_X \phi_X = \lambda \phi_X \quad (13)$$

$$\langle \Delta_X \phi_X, h \rangle_{L_2(X, \mathbb{R})} = \lambda \langle \phi_X, h \rangle_{L_2(X, \mathbb{R})} \quad (14)$$

for all  $h \in L_2(X, \mathbb{R})$ . Assume  $h(x) = c_1 \alpha_1(x) + c_2 \alpha_2(x) + \dots + c_m \alpha_m(x)$ , where  $\{\alpha_i(x)\}_{i=1}^m$  is a basis of some subspace of  $L_2(X, \mathbb{R})$  (for example, a set of some linearly independent polynomials). Substituting this into Equation 14 we get:

$$\sum_{j=1}^m c_j \langle \Delta_X \alpha_j, h \rangle_{L_2(X)} = \lambda \sum_{j=1}^m c_j \langle \alpha_j, h \rangle_{L_2(X)} \quad (15)$$

Taking  $h = \alpha_r(x)$ ,  $r = 1, \dots, m$  we obtain the  $r$  equations:

$$\sum_{j=1}^m c_j \langle \Delta_X \alpha_j, \alpha_r \rangle_{L_2(X)} = \lambda \sum_{j=1}^m c_j \langle \alpha_j, \alpha_r \rangle_{L_2(X)} \quad (16)$$

The above linear system of equations can be written as a generalized eigenvalue problem

$$\mathbf{A}\mathbf{c} = \lambda \mathbf{B}\mathbf{c} \quad (17)$$

where  $A$  and  $B$  are  $m \times m$  matrices with elements  $a_{rj} = \langle \Delta_X \alpha_j, \alpha_r \rangle_{L_2(X, \mathbb{R})}$  and  $b_{rj} = \langle \alpha_j, \alpha_r \rangle_{L_2(X, \mathbb{R})}$ .

For heat kernel approximation a few eigenvalues are required, since the coefficients in the expansion of  $h_t$  (8) decay as  $\mathcal{O}(e^{-t})$ .

For non-triangulated meshes other different methods may be adopted [BSW09a, GLS\*10].

## 6. Results

In order to evaluate the proposed method, we used the SHREC 2010 robust large-scale shape retrieval benchmark methodology [BBC\*10]. The query set consisted of 560 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 95 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), *saturation* (saturation and desaturation by scaling of the  $a, b$  channels), and *color noise* (additive Gaussian noise in all channels), *equi-affine* (rotation and scaling channels  $L$  and  $a, b$  s.t. the scaling matrix will have determinant 1), *affine* (multiplying by matrix  $A$  of determinant value according to strength). *Mixed* transformations included isometry+topology transformations in combination with two randomly selected photometric transformations, and *Mixed-EaAff* and *Mixed-Aff*, with the same isometry-topology transformation and applied on it *equi-affine* and *affine* photometry transformation respectively (the geometry is constant through all strength, only photometry transformation changes). 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 85 shapes used as clutter.

Retrieval was performed by matching 475 transformed queries to the 85 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.



**Figure 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; Figure 1 illustrates equi-affine and affine transformations.

We compared purely geometric and joint photometric-geometric descriptors. As a purely geometric descriptor, we used bags of features based on HKS according to [BBOG09]; as joint photometric-geometric descriptors, we used bags of features computed with the proposed color HKS (cHKS) resulting from different fusion processes.

For the computation of the bag of features descriptors, we used the Shape Google framework with most of the settings as proposed in [BBOG09]. 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 [AMN\*98] was used for vector quantization. The Laplace-Beltrami operator spectrum was computed using the FEM approach discussed in Section 5, [RBB\*11b]. Heat kernels were approximated using the first 200 eigenpairs of the discrete Laplacian. 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 ( $\eta = 0, 0.1, 0.2$ ) 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  $\eta = 0$ ,  $d(X, Y) = \|\text{BoF}_X^0 - \text{BoF}_Y^0\|_1^2 + \sum_{\eta} \eta \|\text{BoF}_X^\eta - \text{BoF}_Y^\eta\|_1^2$ .

Tables 1–3 summarize the results of our experiments. Ge-

ometry only descriptor (HKS) [BBOG09] is invariant to photometric transformations, but is somewhat sensitive to topological noise and missing parts; the recognition rate of the shapes underwent geometry transformation in the *mixed* transformation is 75% (Table 1). The fusion of the geometric and photometric data using Euclidean metric for color embedding (Table 2) does not improve the results (even we see small reduction for *isometry+photometry*), whereas employing our approach with Affine metric for color embedding gives improvement in *Mixed* type of transformations. Figure 3 visualizes a few examples of the retrieved shapes ordered by relevance, which is inversely proportional to the distance from the query shape.

| Transformation     | Strength |        |        |        |        |
|--------------------|----------|--------|--------|--------|--------|
|                    | 1        | ≤2     | ≤3     | ≤4     | ≤5     |
| <i>Isom+Topo</i>   | 86.67    | 80.00  | 76.51  | 72.80  | 76.24  |
| <i>Partial</i>     | 79.17    | 59.49  | 52.52  | 54.35  | 52.43  |
| <i>Contrast</i>    | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Brightness</i>  | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Hue</i>         | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Saturation</i>  | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Noise</i>       | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>EqAff</i>       | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Aff</i>         | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Mixed</i>       | 75.00    | 75.00  | 75.00  | 75.00  | 75.00  |
| <i>Mixed-EqAff</i> | 75.00    | 75.00  | 75.00  | 75.00  | 75.00  |
| <i>Mixed-Aff</i>   | 75.00    | 75.00  | 75.00  | 75.00  | 75.00  |

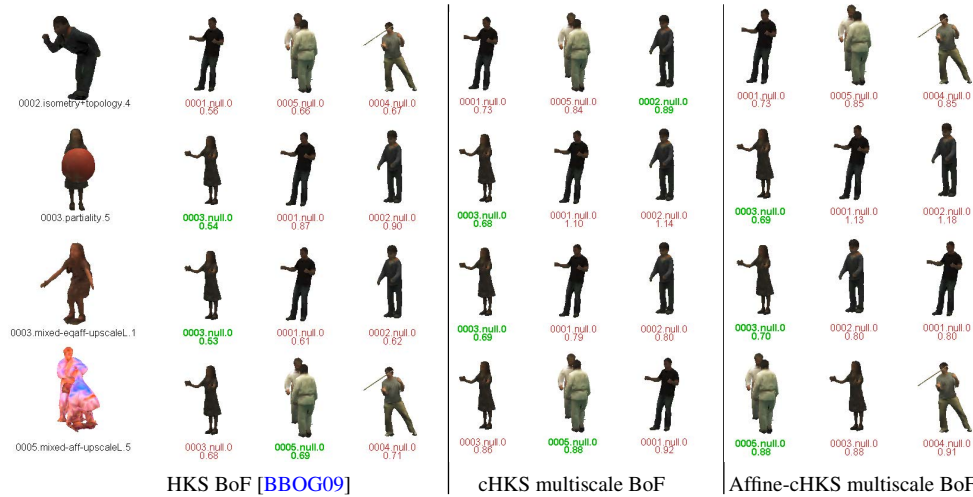
**Table 1:** Performance (mAP in %) of BOFs with purely geometric HKS descriptors.

| Transformation     | Strength |        |        |        |        |
|--------------------|----------|--------|--------|--------|--------|
|                    | 1        | ≤2     | ≤3     | ≤4     | ≤5     |
| <i>Isom+Topo</i>   | 83.33    | 73.10  | 69.84  | 70.30  | 71.24  |
| <i>Partial</i>     | 79.17    | 60.54  | 53.21  | 55.39  | 53.27  |
| <i>Contrast</i>    | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Brightness</i>  | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Hue</i>         | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Saturation</i>  | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Noise</i>       | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>EqAff</i>       | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Aff</i>         | 100.00   | 100.00 | 100.00 | 100.00 | 100.00 |
| <i>Mixed</i>       | 75.00    | 75.00  | 75.00  | 75.00  | 75.00  |
| <i>Mixed-EqAff</i> | 75.00    | 75.00  | 75.00  | 75.00  | 75.00  |
| <i>Mixed-Aff</i>   | 75.00    | 75.00  | 75.00  | 75.00  | 75.00  |

**Table 2:** Performance (mAP in %) of BOFs with multiscale Euclidean metric cHKS descriptors.

## 7. 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. The choice of the metric in the joint geometric-photometric space gives rise to



**Figure 3:** Retrieval results using different methods. First column: query shapes, second column: first three matches obtained with HKS-based BoF [BBOG09], third column: first three matches obtained using cHKS-based multiscale BoF, fourth column: first three matches obtained with the proposed method (Affine-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.

| Transformation     | Strength |          |          |          |          |
|--------------------|----------|----------|----------|----------|----------|
|                    | 1        | $\leq 2$ | $\leq 3$ | $\leq 4$ | $\leq 5$ |
| <i>Isom+Topo</i>   | 86.67    | 80.33    | 76.73    | 72.96    | 74.37    |
| <i>Partial</i>     | 79.17    | 60.54    | 53.41    | 55.01    | 52.55    |
| <i>Contrast</i>    | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>Brightness</i>  | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>Hue</i>         | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>Saturation</i>  | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>Noise</i>       | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>EqAff</i>       | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>Aff</i>         | 100.00   | 100.00   | 100.00   | 100.00   | 100.00   |
| <i>Mixed</i>       | 86.67    | 86.67    | 86.67    | 86.67    | 86.67    |
| <i>Mixed-EqAff</i> | 86.67    | 84.17    | 85.00    | 85.42    | 85.67    |
| <i>Mixed-Aff</i>   | 86.67    | 86.67    | 86.67    | 86.67    | 86.67    |

**Table 3:** Performance (mAP in %) of BoFs with multiscale Affine metric cHKS descriptors.

different invariance properties both to geometric and photometric transformations. 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, testing them on database with a wide variety of shapes with multiple classes.

### Acknowledgements

This research was supported by European Community's FP7- ERC program, grant agreement no. 267414.

### References

- [ABBP07] ASSFALG J., BERTINI M., BIMBO A., PALA P.: Content-based retrieval of 3-d objects using spin image signatures. *Multimedia, IEEE Transactions on* 9, 3 (apr. 2007), 589–599. 1
- [AMN\*98] ARYA S., MOUNT D. M., NETANYAHU N. S., SILVERMAN R., WU A. Y.: An optimal algorithm for approximate nearest neighbor searching. *J. ACM* 45 (1998), 891–923. 6
- [ARK11] AFLALO Y., RAVIV D., KIMMEL R.: *Scale invariant Geometry for Non-Rigid Shapes*. Tech. Rep. CIS-2011-2, Dept. of Computer Science, Technion, Israel, 2011. 3
- [ASC11] AUBRY M., SCHLICKEWEI U., CREMERS D.: The wave kernel signature—a quantum mechanical approach to shape analysis. In *Proc. CVPR* (2011). 2
- [ASR07] AMORES J., SEBE N., RADEVA P.: Context-based object-class recognition and retrieval by generalized correlograms. *Trans. PAMI* 29, 10 (2007), 1818–1833. 1
- [BBC\*10] 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). 5
- [BBG94] BÉRARD P., BESSON G., GALLOT S.: Embedding Riemannian manifolds by their heat kernel. *Geometric and Functional Analysis* 4, 4 (1994), 373–398. 1
- [BBOG09] 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). 1, 4, 6, 7
- [BK10] BRONSTEIN M. M., KOKKINOS I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In *Proc. CVPR* (2010). 2
- [Bro] BRONSTEIN A. M.: Spectral descriptors of deformable shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 2

- [BSW09a] BELKIN M., SUN J., WANG Y.: Constructing Laplace operator from point clouds in Rd. In *Proc. Symp. Discrete Algorithms* (2009), pp. 1031–1040. 5
- [BSW09b] BELKIN M., SUN J., WANG Y.: Discrete Laplace operator on meshed surfaces. In *Proc. Symp. Computational Geometry* (2009), pp. 278–287. 5
- [CL06] COIFMAN R., LAFON S.: Diffusion maps. *Applied and Computational Harmonic Analysis* 21, 1 (2006), 5–30. 1, 2
- [CPS\*07] 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). 1
- [DLL\*10] DEY T. K., LI K., LUO C., RANJAN P., SAFA I., WANG Y.: Persistent heat signature for pose-oblivious matching of incomplete models. *Comput. Graph. Forum* (2010), 1545–1554. 2
- [DMSB99] DESBRUN M., MEYER M., SCHRODER P., BARR A. H.: Implicit fairing of irregular meshes using diffusion and curvature flow. *Proc. SIGGRAPH* (1999), 317–24. 5
- [Dzi87] DZIUK G.: *Finite elements for the Beltrami operator on arbitrary surfaces*. Preprint // Sonderforschungsbereich 256. Nichtlineare partielle Differentialgleichungen. Rheinische Friedrich-Wilhelms-Universität. Univ., Sonderforschungsbereich 256, 1987. 5
- [ELZ00] EDELSBRUNNER H., LETSCHER D., ZOMORODIAN A.: Topological persistence and simplification. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on* (2000), pp. 454–463. 2
- [FSR11] FANG Y., SUN M., RAMANI K.: Temperature distribution descriptor for robust 3D shape retrieval. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on* (June 2011), pp. 9–16. 2
- [GBAL09] GEBAL K., BÆRENTZEN J. A., AANÆS H., LARSEN R.: Shape analysis using the auto diffusion function. *Comput. Graph. Forum* 28, 5 (2009), 1405–1413. 4
- [GLS\*10] GAO W., LAI R., SHI Y., DINOVI I., TOGA A. W.: A narrow band approach for approximating the laplace-beltrami spectrum of 3D shapes. In *ICNAAM 2010: International Conference of Numerical Analysis and Applied Mathematics* (2010), pp. 454–463. 5
- [GMGP05] GELFAND N., MITRA N. J., GUIBAS L. J., POTTMANN H.: Robust global registration. In *Proc. SGP* (2005). 1
- [IJL\*05] IYER N., JAYANTI S., LOU K., KALYANARAMAN Y., RAMANI K.: Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design* 37, 5 (2005), 509–530. 2
- [Jai89] JAIN A. K.: *Fundamentals of digital image processing*. Prentice-Hall information and system sciences series. Prentice Hall, 1989. 3
- [JMS08] JONES P. W., MAGGIONI M., SCHUL R.: Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels. *PNAS* 105, 6 (2008), 1803. 2
- [KBBK11] KOVNATSKY A., BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: Photometric heat kernel signatures. In *Proc. Conf. on Scale Space and Variational Methods in Computer Vision (SSVM)* (2011). 2
- [KMS00] KIMMEL R., MALLADI R., SOCHEN N.: Images as embedded maps and minimal surfaces: movies, color, texture, and volumetric medical images. *IJCV* 39, 2 (2000), 111–129. 2, 3
- [Lév06] LÉVY B.: Laplace-Beltrami eigenfunctions towards an algorithm that “understands” geometry. In *Proc. Shape Modeling and Applications* (2006). 1
- [LJ05] LING H., JACOBS D. W.: Deformation invariant image matching. In *In ICCV* (2005), pp. 1466–1473. 2
- [PZZY05] PAN X., ZHANG Y., ZHANG S., YE X.: Radius-normal histogram and hybrid strategy for 3d shape retrieval. pp. 372–377. 1
- [RBB\*] RAVIV D., BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R., SOCHEN N.: Affine-invariant geodesic geometry of deformable 3d shapes. *Computers & Graphics* 35, 3, 692. 2, 3
- [RBB\*11a] RAVIV D., BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R., SOCHEN N.: Affine-invariant diffusion geometry for the analysis of deformable 3d shapes. In *Proc. CVPR* (2011). 2, 3
- [RBB\*11b] RAVIV D., BRONSTEIN M. M., BRONSTEIN A. M., KIMMEL R., SOCHEN N. A.: Affine-invariant geodesic geometry of deformable 3d shapes. *Computers & Graphics* 35, 3 (2011), 692–697. 6
- [RBBK10] RAVIV D., BRONSTEIN M. M., BRONSTEIN A. M., KIMMEL R.: Volumetric heat kernel signatures. In *Proc. ACM Multimedia Workshop on 3D Object Retrieval* (2010). 2
- [RK12] RAVIV D., KIMMEL R.: *Affine invariant non-rigid shape analysis*. Tech. Rep. CIS-2012-01, Dept. of Computer Science, Technion, Israel, 2012. 3, 4
- [Rus07] RUSTAMOV R. M.: Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *Proc. SGP* (2007), pp. 225–233. 2
- [Soc01] SOCHEN N.: On affine invariance in the beltrami framework for vision. *VLSM* (2001). 3
- [SOG09a] SUN J., OVSJANIKOV M., GUIBAS L.: A Concise and Provably Informative Multi-Scale Signature Based on Heat Diffusion. In *Computer Graphics Forum* (2009), vol. 28, pp. 1383–1392. 2
- [SOG09b] SUN J., OVSJANIKOV M., GUIBAS L. J.: A concise and provably informative multi-scale signature based on heat diffusion. In *Proc. SGP* (2009). 4
- [SZ03] SIVIC J., ZISSERMAN A.: Video google: A text retrieval approach to object matching in videos. In *Proc. CVPR* (2003). 1, 4
- [TCF09] TOLDO R., CASTELLANI U., FUSIELLO A.: Visual vocabulary signature for 3D object retrieval and partial matching. In *Proc. 3DOR* (2009). 1
- [TV04] TANGELDER J. W. H., VELTKAM R.: A survey of content based 3D shape retrieval methods. In *Shape Modeling Applications, 2004. Proceedings* (June 2004), pp. 145–156. 2
- [WMKG08] WARDETZKY M., MATHUR S., KÄLBERER F., GRINSPUN E.: Discrete Laplace operators: no free lunch. In *Conf. Computer Graphics and Interactive Techniques* (2008). 5
- [YPS10] YOON K.-J., PRADOS E., STURM P.: Joint estimation of shape and reflectance using multiple images with known illumination conditions, 2010. 2
- [ZBH07] ZAHARESCU A., BOYER E., HORAUD R. P.: Transforms: a topology-adaptive mesh-based approach to surface evolution, November 2007. 2