Affine-Invariant Photometric Heat Kernel Signatures

Artiom Kovnatsky 1 Michael M. Bronstein 1 Alexander M. Bronstein 2 Dan Raviv 3 Ron Kimmel 3

1 Institute of Computational Science, Faculty of Informatics, Università della Svizzera Italiana, Lugano, Switzerland
2 School of Electrical Engineering, Tel Aviv University, Israel
3 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 documents, 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 [ABB07], 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 [RBB11, 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 surveys [TV04, IJL].

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, LIO05]. 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 $g$. 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 $T_xX$ 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 $dx \in T_xX$. The metric tensor $g$ defines the Laplace-Beltrami operator $\Delta_g$ that satisfies

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

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 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,

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

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 $\phi_0, \phi_1, \ldots$ of the Laplace-Beltrami operator (i.e., solutions to $\Delta_g \phi = \lambda \phi$, where $0 = \lambda_0 \leq \lambda_1 \leq \ldots$ are the corresponding eigenvalues). This basis can be interpreted analogously to the Fourier basis, and the eigenvalues $\lambda_k$ as frequencies. Consequently, the heat kernel can be represented as [JMS08]

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

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 \( \xi : X \to \mathcal{E} \). 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 \( r, s \in TX \), where \( d\xi : TX \to T\mathcal{E} \) is the differential of the embedding coordinate. In coordinate notation, the pullback metric is expressed as \( g_{\mu\nu} = (\xi^*h)_{\mu\nu} = h_{ij}\partial\xi^i\partial\xi^j \). 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 [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 \( (\xi^*h)_{\mu\nu} = \delta_{ij}\partial\xi^i\partial\xi^j + \cdots + \delta_{ij}\partial\xi^i\partial\xi^j = (\partial\xi^i, \partial\xi^j) \). In the case considered in this paper, the shape is endowed with photometric information given in the form of a field \( \alpha : X \to C \), where \( C \) denotes some color space (e.g., RGB or Lab). In the following, we require, that tacitly assume that \( \alpha \) is sufficiently smooth.

This photometric information can be modeled by defining \( \mathcal{E} = \mathbb{R}^3 \times C \) and an embedding \( \xi = (\xi_1, \xi_2) \). The embedding coordinates corresponding to geometric information are as before \( \xi_1 = (\xi_1, \ldots, \xi_n) \), and the embedding coordinate corresponding to photometric information are given by \( \xi_2 = (\xi_1, \ldots, \xi^n) = \eta(\alpha_1, \ldots, \alpha^3) \), 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 \( \Delta_1 \) 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 color space \( C \). In the simplest case, we assume \( C \) to have a Euclidean structure.

While being the simplest choice, the Euclidean metric is known to be perceptually meaningful in some color spaces 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 color space 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,

\[
\xi_p = R\xi_p + c, \tag{4}
\]

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 \( (\xi^*g)_{\mu\nu} = (\partial\xi_1\partial\xi_\mu, \partial\xi_\nu) + \eta(\partial\xi_1\partial\xi_p, \partial\xi_p) \).

### 3.2. Affine-invariant color metric

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

Raviv et al. [RBB+11a, RBB+11b] showed a construction of a metric that is invariant to equi-affine transformations. In our setting, let us be given some parametrization \( \Phi(u, u_2) : U \subseteq \mathbb{R}^2 \to X \subseteq \mathbb{R}^3 \) of the shape; the composition of \( \alpha \circ \Phi \) gives us a parametrization of the texture. First, allowing some relaxed notation, we denote by \( g_{\Phi}(u_1, u_2) = (\partial\Phi(u_1, u_2)) \) and \( g_{\Phi}(u_1, u_2) = ((\partial\Phi_{\alpha}(u_1, u_2)) = (((\partial\Phi_{\alpha}(u_1, u_2)) = (d\alpha(u_1, u_2), d\alpha(u_2)) \) the 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 \( d\alpha(u_1, u_2) = \frac{\partial}{\partial u_1}((\alpha \circ \Phi)(u_1, u_2)) \). Second, construct an equi-affine pre-metric tensor [Sec01, RBB+11a]

\[
(\bar{g}(u_1, u_2))_{\mu\nu} = g_{\Phi}\det^{-1/2}(\bar{g}), \tag{5}
\]

where \( \bar{g}_{\mu\nu} = \det(d\alpha(u_1), d\alpha(u_2), (\alpha \circ \Phi)(u_1, u_2)) \). 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 \( \bar{g} \) to have positive eigenvalues. For additional details about derivation and proof of affine invariance, we refer the reader to [RBB+11a, RBB+11b, 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

\[
\hat{g}(u_1, u_2) = g_{\Phi}(u_1, u_2) + \eta g_{\Phi}(u_1, u_2). \tag{6}
\]

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.
Equi-affine-invariant metric is defined by

\[ g(X, \tilde{X}) = \det (\tilde{g}^{-1} g) \]

where \( \tilde{g} \) is the equi-affine invariant metric, and compute the Gaussian curvature \( K(X, \tilde{X}) \) at each point. The affine invariant metric is defined by

\[ g_{ij} = |K(X, \tilde{X})| \tilde{g}_{ij}. \]

3.3. Invariance of the joint diffusion process

The joint metric tensor \( g \) and the diffusion geometry it induces have inherent ambiguities. Let us denote by \( \text{Iso}_g = \text{Iso}(\xi^g h)_{ij} \) and \( \text{Iso}_p = \text{Iso}(\xi^p h)_{ij} \) 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 \( g \) is invariant to joint isometry group \( \text{Iso}_g = \text{Iso}(\xi^g h)_{ij} \). Ideally, we would like \( \text{Iso}_g = \text{Iso}_X \times \text{Iso}_p \) to hold. In practice, \( \text{Iso}_g \) 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 \( (\xi^g h)_{ij}^{\eta} \) combined with compensating scaling of \( (\xi^p h)_{ij}^{\eta} \).

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,

\[ k_t(x, x) = \sum_{i \geq 0} e^{-\lambda_i^2 t} \theta_i^2(x), \]

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 \( p(x) = (k_t(x, x), \ldots, k_t(x, x)) \), where \( t_1, \ldots, t_n \) 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” \( \{p_1, \ldots, p_q\} \subset \mathbb{R}^d \) 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,

\[ \theta(x) = \theta_1(x), \ldots, \theta_q(x) \]

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-Belrami 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 $p_i(x)$ and corresponding bags of features $b_i(\eta)$. This set of BOFs can be compared e.g. as

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

5. 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 $(i, j), (j, k), (i, k)$ 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

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

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 [WMK08, 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]:

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

$$\langle \Delta \phi_X, h \rangle_{L^2(X, \mathbb{R})} = \lambda \langle \phi_X, h \rangle_{L^2(X, \mathbb{R})} \quad (14)$$

for all $h \in L^2(X, \mathbb{R})$. Assume $h(x) = c_1 \alpha_1(x) + c_2 \alpha_2(x) + \ldots + c_m \alpha_m(x)$, 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:

$$\sum_{j=1}^m c_j \langle \Delta \alpha_j, h \rangle_{L^2(X)} = \lambda \sum_{j=1}^m c_j \langle \alpha_j, h \rangle_{L^2(X)} \quad (15)$$

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

$$\sum_{j=1}^m c_j \langle \Delta \alpha_j, \alpha_r \rangle_{L^2(X)} = \lambda \sum_{j=1}^m c_j \langle \alpha_j, \alpha_r \rangle_{L^2(X)} \quad (16)$$

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

$$Ac = \lambda Bc \quad (17)$$

where $A$ and $B$ are $m \times m$ matrices with elements $a_{ij} = \langle \Delta \alpha_i, \alpha_j \rangle_{L^2(X, \mathbb{R})}$ and $b_{ij} = \langle \alpha_i, \alpha_j \rangle_{L^2(X, \mathbb{R})}$.

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

For non-triangulated meshes other different methods may be adopted [BSW09a, GLS10].

6. Results

In order to evaluate the proposed method, we used the SHREC 2010 robust large-scale shape retrieval benchmark methodology [BBC10]. 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), 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), 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 \( mAP = \sum_i P(i) \cdot \text{rel}(i) \) where \( \text{rel}(i) \) 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.

We compared purely geometric and joint photometric-geometric descriptors. As a purely geometric descriptor, we used bags of features based on HKS according to [BBBG09]; 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 HKS (cHKS) resulting from different fusion processes. Used bags of features computed with the proposed color HKS according to [AMN].

We 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.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>1</th>
<th>( \leq 2 )</th>
<th>( \leq 3 )</th>
<th>( \leq 4 )</th>
<th>( \leq 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isom+Topo</td>
<td>79.17</td>
<td>59.49</td>
<td>52.52</td>
<td>54.35</td>
<td>52.43</td>
</tr>
<tr>
<td>Partial</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Contrast</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Brightness</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Hue</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Saturation</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>EqAff</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Aff</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transformation</th>
<th>1</th>
<th>( \leq 2 )</th>
<th>( \leq 3 )</th>
<th>( \leq 4 )</th>
<th>( \leq 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isom+Topo</td>
<td>83.33</td>
<td>73.10</td>
<td>69.84</td>
<td>70.30</td>
<td>71.24</td>
</tr>
<tr>
<td>Partial</td>
<td>79.17</td>
<td>60.54</td>
<td>53.21</td>
<td>55.19</td>
<td>53.27</td>
</tr>
<tr>
<td>Contrast</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Brightness</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Hue</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Saturation</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>EqAff</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Aff</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

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

HKS BoF [BBOG09]

Affine-cHKS multiscale BoF

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.

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

<table>
<thead>
<tr>
<th>Transformation</th>
<th></th>
<th>≤2</th>
<th>≤3</th>
<th>≤4</th>
<th>≤5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isom+Topo</td>
<td>86.67</td>
<td>80.33</td>
<td>70.73</td>
<td>72.96</td>
<td>74.37</td>
</tr>
<tr>
<td>Partial</td>
<td>79.17</td>
<td>60.54</td>
<td>53.41</td>
<td>55.01</td>
<td>52.55</td>
</tr>
<tr>
<td>Contrast</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Brightness</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Hue</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Saturation</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>EqAff</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Aff</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>86.67</td>
<td>86.67</td>
<td>86.67</td>
<td>86.67</td>
<td>86.67</td>
</tr>
<tr>
<td>Mixed-EqAff</td>
<td>86.67</td>
<td>84.17</td>
<td>85.00</td>
<td>85.42</td>
<td>85.67</td>
</tr>
<tr>
<td>Mixed-Aff</td>
<td>86.67</td>
<td>86.67</td>
<td>86.67</td>
<td>86.67</td>
<td>86.67</td>
</tr>
</tbody>
</table>

Acknowledgements

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

References


