Learning spectral descriptors for deformable shape correspondence

R. Litman and A. M. Bronstein *Senior Member, IEEE*

Abstract—Informative and discriminative feature descriptors play a fundamental role in deformable shape analysis. For example, they have been successfully employed in correspondence, registration, and retrieval tasks. In the recent years, significant attention has been devoted to descriptors obtained from the spectral decomposition of the Laplace-Beltrami operator associated with the shape. Notable examples in this family are the heat kernel signature (HKS) and the recently introduced wave kernel signature (WKS). Laplacian-based descriptors achieve state-of-the-art performance in numerous shape analysis tasks; they are computationally efficient, isometry-invariant by construction, and can gracefully cope with a variety of transformations. In this paper, we formulate a generic family of parametric spectral descriptors. We argue that in order to be optimized for a specific task, the descriptor should take into account the statistics of the corpus of shapes to which it is applied (the "signal") and those of the class of transformations to which it is made insensitive (the "noise"). While such statistics are hard to model axiomatically, they can be learned from examples. Following the spirit of the Wiener filter in signal processing, we show a learning scheme for the construction of optimized spectral descriptors and relate it to Mahalanobis metric learning. The superiority of the proposed approach in generating correspondences is demonstrated on synthetic and scanned human figures. We also show that the learned descriptors are robust enough to be learned on synthetic data and transferred successfully to scanned shapes.

Index Terms—diffusion geometry, heat kernel signature, wave kernel signature, HKS, WKS, descriptor, deformable shapes, correspondence, retrieval, spectral methods, Laplace-Beltrami operator, metric learning, Wiener filter, Mahalanobis distance

1 INTRODUCTION

The notion of a *feature descriptor* is fundamental in shape analysis. A feature descriptor assigns each point on the shape a vector in some single- or multi-dimensional feature space representing the point's local and global geometric properties relevant for a specific task. This information is subsequently used in higher-level tasks: for example, in shape matching, descriptors are used to establish an initial set of potentially corresponding points [1], [2]; in shape retrieval a global shape descriptor is constructed as a bag of "geometric words" expressed in terms of local feature descriptors [3], [4]; segmentation algorithms rely on the similarity or dissimilarity of feature descriptors to partition the shape into stable and meaningful parts [5].

When constructing or choosing a feature descriptor, it is imperative to answer two fundamental questions: which shape properties the descriptor has to capture, and to which transformations of the shape it shall remain invariant or, at least, insensitive.

• *School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel. Email: bron@eng.tau.ac.il.* *This work was supported by the Israeli Science Foundation and the German-Israeli Foundation.*

1.1 Previous work

Early research on feature descriptors focused mainly on invariance under global Euclidean transformations (rigid motion). Classical works in this category include the shape context [6] and spin image [7] descriptors, as well as integral volume descriptors [8], [1] and multiscale local features [9] just to mention a few out of many.

In the past decade, significant effort has been invested in extending the invariance properties to non-rigid deformations. Some of the classical rigid descriptors were extended to the non-rigid case by replacing the Euclidean metric with its geodesic counterpart [10], [11]. Also, the use of conformal factors has been proposed [12]. Being intrinsic properties of a surface, both are independent of the way the surface is embedded into the ambient Euclidean space and depend only on its metric structure. This makes such descriptors invariant to inelastic bending transformations. However, geodesic distances suffer from strong sensitivity to topological noise, while conformal factors, being a local quantity, are influenced by geometric noise. Both types of noise, virtually inevitable in real applications, limit the usefulness of such descriptors.

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.* [13] and later by Coifman and Lafon [14] who suggested to use the eigenvalues and eigenvectors of the Laplace-Beltrami operator associated with the shape to construct invariant metrics known as diffusion distances. These distances as well as other diffusion geometric constructs have been show significantly more robust compared to their geodesic counterparts [15], [16]. Diffusion geometry offers an intuitive interpretation of many shape properties in terms of spacial frequencies and allows to use standard harmonic analysis tools. Also, recent advances in the discretization of the Laplace-Beltrami operator brought forth efficient and robust numerical and computational tools.

These methods were first explored in the context of shape processing by Lévy [17]. Several attempts have also been made to construct feature descriptors based on diffusion geometric properties of the shape. Rustamov [18] proposed to construct the *global point signature* (GPS) feature descriptors by associating each point with an ℓ^2 sequence based on the eigenfunctions and the eigenvalues of the Laplacian, closely resembling a diffusion map [14]. A major drawback of such a descriptor was its ambiguity to sign flips of each individual eigenfunction (or, in the most general case, to rotations and reflections in the eigenspaces corresponding to each eigenvalue).

A remedy was proposed by Sun *et al.* who in their influential paper [19] introduced the *heat kernel signature* (HKS), based on the fundamental solutions of the heat equation (heat kernels). In [20], another physically-inspired descriptor, the *wave kernel signature* (WKS) was proposed as a solution to the excessive sensitivity of the HKS to low-frequency information. As of today, these descriptors achieve state-of-the-art performance in many deformable shape analysis tasks [21], [22] and lie in the foundation of many recent works in shape analysis such as [23], [4], [24].

1.2 Contribution

In this paper, we remain within the diffusion geometric framework and propose a generic family of spectral feature descriptors that generalize the HKS and the WKS. We analyze both descriptors within this framework pointing to their advantages and drawbacks, and enumerate a list of desired properties a descriptor should have.

We argue that in order to construct a good task-specific spectral descriptor, one has to be in the position of defining the spectral content of the geometric "signal" (i.e., the properties distinguishing different classes of shapes from each other) and the "noise" (i.e., the changes of the latter properties due to the deformations the shapes undergo). Both are functions of the corpus of data of interest, and the transformations invariance to which is desired. While it is notoriously difficult to characterize such properties analytically, we propose to learn them from examples in a way resembling the construction of a Wiener filter that passes frequencies containing more signal than noise, while attenuating those where the noise covers the signal. We give experimental evidence to the fact that the proposed construction of descriptors is robust enough to be transferred across very different sets of data.

This study was in part inspired by the insightful paper by Auby *et al.* [20], and in part is a continuation of [25] where we attempted to construct optimal diffusion metrics. However, since diffusion metrics are characterized by a single frequency response, the attempt had a modest success. On the other hand, vector-valued feature descriptors allowing for multiple frequency response functions have, in our opinion, more potential. This paper does not intend to exhaust this potential, but merely to explore a part of it. We focus on deformable shape correspondence problems, and attempt to learn descriptors that improve this task on specific classes of shapes. We believe that approaching correspondence as a learning problem is a novel perspective that can be beneficial in shape analysis and, generally, in computer vision.

1.3 Organization

The rest of the paper is organized as follows: In Section 2 we introduce the mathematical notation of the Laplace-Beltrami operator and its spectrum and briefly overview the state-of-the-art descriptors based on its properties. In Section 3, we indicate several drawbacks of these descriptors and analyze the properties a good descriptor should satisfy. We present a spectral descriptor generalizing the heat and the wave kernel signatures, and show an approach for learning its optimal task-specific parameters from examples. The relation to metric learning is highlighted. In Section 4, the superiority of the proposed learnable descriptor over the fixed ones is shown experimentally on the synthetic TOSCA as well as the scanned SCAPE datasets. Finally, Section 5 concludes the paper.

Since the figures visualizing the experiments in Section 4 are relatively self-explanatory, we decided to incorporate them in the flow as illustrations to the phenomena discussed in the paper even before the exact experimental settings are detailed.

2 SPECTRAL DESCRIPTORS

We model a shape as a compact two-dimensional manifold X, possibly with a boundary ∂X. The manifold is endowed with a Riemannian metric defined as a local inner product $\langle \cdot, \cdot \rangle_x$ on the tangent plane $T_x X$ at each point $x \in X$. Given a smooth scalar field f on the manifold, its gradient $\mathrm{grad}\, f$ is the vector field satisfying $f(x + dr) = f(x) + \langle \mathrm{grad}\, f(x), dr \rangle_x$ for every infinitesimal tangent vector $dr \in T_x X$. The inner product $\langle \mathrm{grad}\, f(x), v \rangle_x$ can be interpreted as the directional derivative of f in the direction v. A directional derivative of f whose direction at every point is defined by a vector field V on the manifold is called the Lie derivative of f along V. The Lie derivative of the manifold volume (area) form along a vector field V is called the divergence of V, $\mathrm{div}\, V$. The negative divergence of the gradient of a scalar field f, $\Delta f = -\mathrm{div}\,\mathrm{grad}\, f$, is called the *Laplacian* of f. The operator Δ is called the *Laplace-Beltrami* operator, and it generalizes the standard notion of the Laplace operator to manifolds. Note that we define the Laplacian with the negative sign to conform to the computer graphics and computational geometry convention.

2.1 Laplacian spectrum and Shape DNA

Being a positive self-adjoint operator, the Laplacian admits an eigendecomposition

$$\Delta \phi = \nu \phi \tag{1}$$

with non-negative eigenvalues ν and corresponding orthogonormal eigenfunctions ϕ. Furthermore, due to the assumption that our domain is compact, the spectrum is discrete, $0 = \nu_1 < \nu_2 < \cdots$.

In physics, (1) is known as the *Helmohltz equation* representing the spatial component of the wave equation. Thinking of our domain as of a vibrating membrane (with appropriate boundary conditions), the ϕ_k's can be interpreted as natural vibration modes of the membrane, while the ν_k's assume the meaning of the corresponding vibration frequencies. In fact, in this setting the eigenvalues have inverse area or squared spatial frequency units.

This physical interpretation leads to a natural question whether the eigenvalues of the Laplace-Beltrami operator fully determine the shape of the domain. The essence of this question was beautifully captured by Mark Kac as *"can one hear the shape of the drum?"* [26]. Unfortunately, the answer to this question is negative as there exist isospectral manifolds that are not isometric. The exact relation between the latter two classes of shapes is unknown, but it is generally believed that most isospectral manifolds are also isometric. Based on this belief, Reuter *et al.* [27] proposed to use truncated sequences of the Laplacian eigenvalues as isometry-invariant shape descriptors, dubbed by the authors as *shape DNA*.

2.2 Heat kernel signature

The Laplace-Beltrami operator plays a central role in the *heat equation* describing diffusion processes on manifolds. In our notation, the heat equation can be written as

$$\left(\Delta + \frac{\partial}{\partial t} \right) u(x,t) = 0 \tag{2}$$

where $u(x,t)$ is the distribution of heat on the manifold at point x at time t. The initial condition is some initial heat distribution $u_0(x)$ at time $t = 0$, and boundary conditions are applied in case the manifold has a boundary.

The solution of the heat equation at time t can be expressed as the application of the *heat operator*

$$u(x,t) = \int h_t(x,y) u_0(y) da(y) \tag{3}$$

to the initial distribution. The kernel $h_t(x,y)$ of this integral operator is called the *heat kernel* and it corresponds to the solution of the heat equation at point x at time t with the initial distribution being a delta function at point y. From the signal processing perspective, the heat kernel can be interpreted as a non shift-invariant "impulse response". It also describes the amount of heat transferred from point x to point y after time t, as well as the transition probability density from point x to point y by a random walk of length t.

According to the spectral decomposition theorem, the heat kernel can be expressed as

$$h_t(x,y) = \sum_{k \geq 1} \exp(-\nu_k t) \phi_k(x) \phi_k(y), \tag{4}$$

where $\exp(-\nu t)$ can be interpreted as its "frequency response" (note that with a proper selection of units in (3), the eigenvalues ν_k assume inverse time or frequency units). The bigger is the time parameter, the lower is the cut-off frequency of the low-pass filter described by this response and, consequently from the uncertainty principle, the bigger is the support of h_t on the manifold. The quantity

$$h_t(x,x) = \sum_{k\geq 1} \exp(-\nu_k t)\phi_k^2(x), \tag{5}$$

sometimes referred to as the *autodiffusivity function* [28], describes the amount of heat remaining at point x after time t. Furthermore, for small values of t is it related to the manifold curvature according to

$$h_t(x,x) = \frac{1}{4\pi t} + \frac{K(x)}{12\pi} + \mathcal{O}(t), \tag{6}$$

where $K(x)$ denotes the Gaussian (in general, sectional) curvature at point x.

In [19], Sun *et al.* showed that under mild technical conditions, the sequence $\{h_t(x,x)\}_{t>0}$ contains *full* information about the metric of the manifold. The authors proposed to associate each point x on the manifold with a vector

$$\mathbf{p}(x) = \left(h_{t_1}(x,x),\ldots,h_{t_n}(x,x)\right)^{\mathrm{T}}, \tag{7}$$

of the autodiffusivity functions sampled at a finite set of times $t_1,\ldots,t_n$. The authors dubbed such a feature descriptor as the *heat kernel signature*. In [4], an HKS-based bag-of-features approach was introduced under the name of Shape Google and was shown to achieve state-of-the-art results in deformable shape retrieval. In [24], a scale-invariant version of the HKS was proposed, and [29] extended the descriptor to volumes.

Despite its success, the heat kernel descriptor suffers from several drawbacks. First, being a collection of low-pass filters (Figure 1, top), the descriptor is dominated by low frequencies conveying information mostly about the global structure of the shape. While being important to discriminate between distinct shapes (which usually differ greatly at coarse scales), this emphasis of low frequencies damages the ability of the descriptor to precisely localize features. This phenomenon can be observed in Figure 2 (top). In fact, the distance between HKS computed at a point x and HKS of neighboring points increases slowly, while for good localization a steeper increase is required.

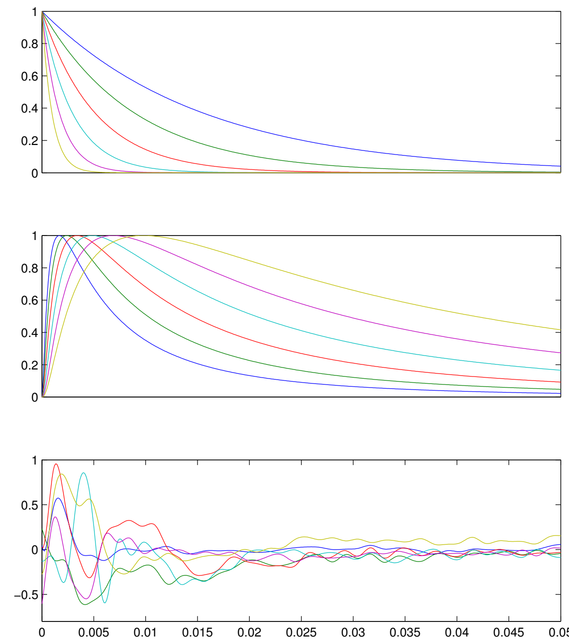

Fig. 1. Examples of (unnormalized) kernels used for the computation of the heat kernel (first row), wave kernel (second row), and trained optimized kernel (last row) descriptors.

2.3 Wave kernel signature

A remedy to the poor feature localization of the heat kernel descriptor was proposed by Aubry *et al.* [20]. The authors proposed to replace the heat diffusion model that gives rise to the HKS by a different physical model in which one evaluates the probability of a quantum particle with a certain energy distribution to be located at a point x. The behavior of a quantum particle on a surface is governed by the Schrödinger equation

$$\left(i\Delta + \frac{\partial}{\partial t}\right)\psi(x,t) = 0 \tag{8}$$

where $\psi(x,t)$ is the complex wave function. Despite an apparent similarity to the heat equation, the multiplication of the Laplacian by the complex unity in the Schrödinger equation has a dramatic impact on the dynamics of the solution. Instead of representing diffusion, ψ now has oscillatory behavior.

Let us assume that the quantum particle has an initial energy distributed around some nominal energy and described by the probability density function $f(e)$. Since energy is directly related to frequency, we will use $f(\nu)$ instead in order to stick to the previous notation.

Fig. 2. Normalized Euclidean distance between the descriptor at a reference point on the right wrist, belly, and chest (white dots pointed by red arrows) and descriptors computed at rest of the points of the same synthetic shape from the TOSCA set (left shape in each group), its approximate isometry (middle shape in each group), and a scanned human shape from the SCAPE set (right shape in each group). 16-dimensional descriptors based on the heat kernel (first row), wave kernel (second row), and trained kernel (last row) are shown. Dark blue stands for small distance; red represents large distance. To improve visual rendering, a common color map scale is used in each row for each descriptor, and is saturated at the median distance on the rightmost shape in each group (i.e. at least half of a shape is always red).

The solution of the Schrödinger equation can then be expressed in the spectral domain as [20]

$$\psi(x,t) = \sum_{k\geq 1} \exp\left(i\nu_k t\right) f(\nu_k)\phi_k(x) \tag{9}$$

(note the complex unity in the exponential!). The probability to measure the particle at a point x at time t is given by $|\psi(x,t)|^2$. By integrating over all times, the average probability

$$p(x) = \lim_{T\to\infty} \frac{1}{T}\int_0^T |\psi(x,t)|^2 dt = \sum_{k\geq 1} f^2(\nu_k)\phi_k^2(x) \tag{10}$$

to measure the particle at a point x is obtained. Note that the probability depends on the initial energy distribution f.

Aubry *et al.* considered a family of log-normal energy distributions

$$f_e(\nu) \propto \exp\left(-\frac{(\log e - \log\nu)^2}{2\sigma^2}\right) \tag{11}$$

centered around some mean log energy $\log e$ with variance σ^2 (again, we allow ourselves a certain abuse of the physics and treat energy and frequency as synonyms). This particular choice of distributions is motivated by a perturbation analysis of the Laplacian spectrum [20].

Fixing the family of energy distributions, each point on the surface is associated with a *wave kernel signature* of the form

$$\mathbf{p}(x) = \left(p_{e_1}(x), \ldots, p_{e_n}(x)\right)^{\mathrm{T}}, \tag{12}$$

where $p_e(x)$ is the probability to measure a quantum particle with the initial energy distribution $f_e(\nu)$ at point x. The authors use logarithmically sampled $e_1, \ldots, e_n$.

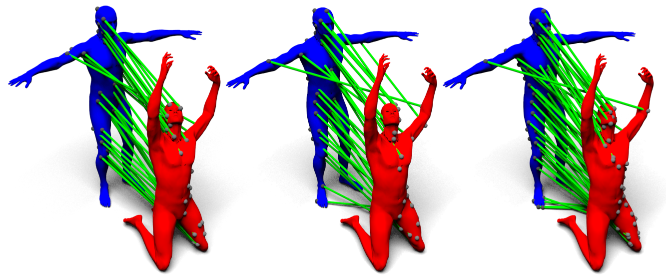

Fig. 3. Correspondences computed on TOSCA shapes using the spectral matching algorithmn [30]. Shown are the matches with geodesic distance distortion below 10% of the shape diameter, from left to right: HKS (34 matches), WKS (30 matches), and trained descriptor (54 matches).

The WKS descriptor resembles the HKS in the sense that it can also be thought of as an application of a set of filters with the frequency responses $f_e^2(\nu)$. However, unlike the HKS that uses low-pass filters, the responses of the WKS are *band-pass* (Figure 1, middle). This reduces the influence of the low frequencies and allows better separation of frequency bands across the descriptor dimensions. As the result, the wave kernel descriptor exhibits superior feature localization (Figure 2, middle).

3 Spectral descriptor learning

Despite their beautiful physical interpretation, both the heat and wave kernel descriptors suffer from several drawbacks.

The fact that the WKS deemphasizes large-scale features contributes to its higher *sensitivity* (i.e., the ability to identify positives). This property is crucial in matching problems, where a small set of candidate matches on one shape is found for a collection of reference points on the other. The ability to produce a correct match within a small set of best matches (high true positive rate at low false positive rate) greatly increases the performance of correspondence algorithms and allows to detect denser corresponding sets.

On the other hand, by emphasizing global features HKS has higher *specificity* (i.e., the ability to identify negatives). Without high specificity, many regions on the shape being matched may look similar to a query point, producing many false negative matches in geometrically inconsistent regions. This property is related (though indirectly) to *discriminativity*, that is, the ability of the descriptor to distinguish between a shape and other classes of distinct shapes. High discriminativity is important when the descriptor is used in retrieval applications, and the performance of the descriptor at low false negative rates has a big impact on retrieval algorithms based on it.

Sensitivity and specificity is visualized in the first two rows of Figure 2. The first row demonstrates the high specificity of HKS (each query point has few well-matching regions) as well as its relatively low sensitivity (the large extents and poor localization of the matching regions). The second row demonstrates the opposite behavior of WKS: each query point has many unrelated matching regions (low specificity), but the correctly matching region is well-localized (high sensitivity). While it is impossible to maximize both the sensitivity and the specificity, a good descriptor is expected to have both reasonably high.

Another drawback of both the heat and wave kernel descriptors is the fact that the frequency responses forming their elements have significant overlaps. As the result, the descriptor has redundant dimensions. Finally, both the heat and wave kernel signatures are only invariant to truly isometric deformations of the shape (and can be also made scale-invariant using the scheme proposed in [24]). Deformations that real shapes undergo frequently deviate from this model, and it is unclear how they influence the performance of the HKS and WKS.

We assert that many real-world deformations affect different frequencies differently. At the same time, the geometric features that allow to localize a point on a shape or to distinguish a shape from other shapes also depend differently on different frequencies. Emphasizing information-carrying frequencies while attenuating noise-carrying ones is a classical idea in signal and is the underlying principle of Wiener filtering [31].

3.1 Desired properties

This observation leads us to the main contribution of this paper: we propose to construct a collection of frequency responses forming an optimal spectral descriptor. In order to be useful, such a descriptor should satisfy the following properties:

1) *Sensitivity*: when a point on a shape is queried against another shape from the same class, a small set of best matches of the descriptor should contain a correct match with high probability (ideally, the first best match shall be correct). High sensitivity

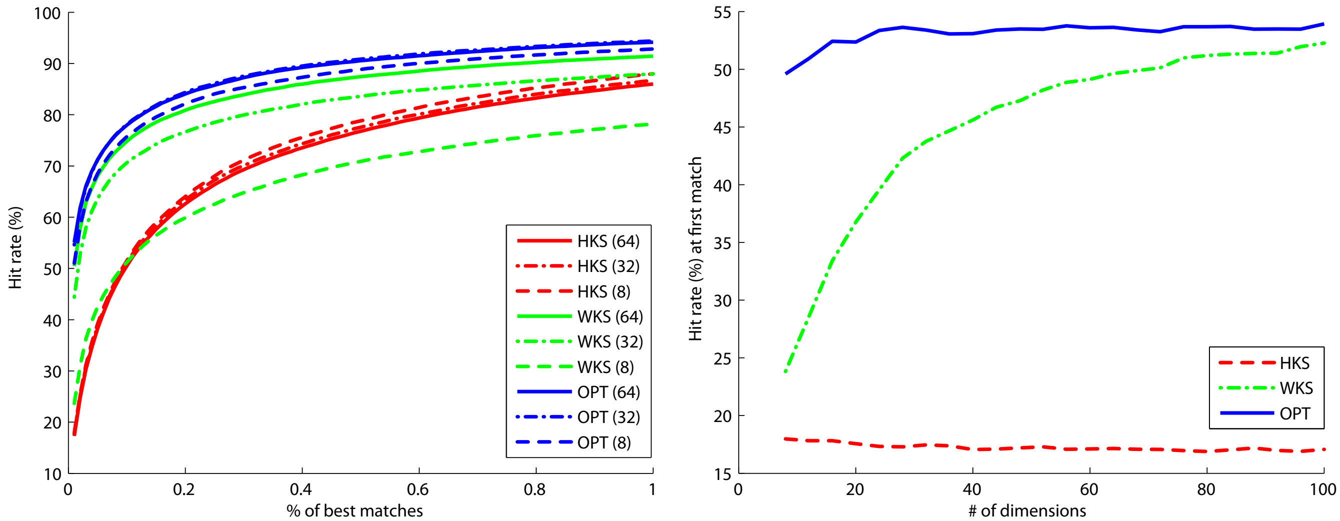

Fig. 4. Left: CMC curves of the HKS, WKS, and learned descriptors on the TOSCA shapes for different number of dimensions (shown in parenthesis). Right: hit rate of the first best match of the same descriptors as the function of the number of dimensions. The superior performance of the low-dimensional learned descriptor is a manifestation of its efficiency.

is akin to low number of false negatives. This property can be alternatively stated as *Localization*: a small displacement of a point on the manifold should greatly affect the descriptor computed at it.

2) *Specificity*: when a point on a shape is queried against another shape from the same class, the set of best matches of the descriptor should contain only points in the vicinity of the correct match with high probability. High specificity is akin to low number of false positives. This property can be alternatively stated as *Discriminativity*: the descriptor should be able to distinguish between the geometric content of a local region and that of other, possibly similarly looking, regions.
3) *Insensitivity to transformations*: the descriptor should be invariant or at least insensitive to a certain class of transformations that the shape may undergo.
4) *Efficiency*: the descriptor should capture as much information as possible within as little number of dimensions as possible.

The localization and sensitivity properties are important for matching tasks, while in order to be useful in shape retrieval tasks, the descriptor should have the discriminativity property. However, discriminativity is data-dependent: a descriptor can be discriminative on one corpus of data, while non-discriminative on another. While it is generally impractical to model classes of shapes axiomatically, machine learning offers an easy alternative of inferring them from training data.

By construction, spectral descriptors are isometry invariant. However, other invariance properties are usually hard to achieve and even harder to model for realistic transformations. We will therefore stick to learning in order to achieve invariance on examples of transformations the training shapes undergo.

3.2 Parametrization

We are interested in descriptors of the form

$$\mathbf{p}(x) = \sum_{k\geq 1} \mathbf{f}(\nu_k)\phi_k^2(x), \tag{13}$$

parameterized by a vector $\mathbf{f}(\nu) = (f_1(\nu), \ldots, f_n(\nu))^{\mathrm{T}}$ of frequency responses. Both the HKS and the WKS are particular cases of this general form. Unlike both heat and wave kernels that are strictly positive, we will allow $\mathbf{f}(\nu)$ assume negative values.

Since the responses $\mathbf{f}(\nu)$ are the design variables of the descriptor, they have to be parametrized with a finite set of parameters. The same parameters have to be compatible with any shape, even though different shapes differ in the set of eigenvalues $\{\nu_k\}$. In order to make the representation independent of a specific shape's eigenvalues, we fix a basis $\{b_1(\nu), \ldots, b_m(\nu)\}$, $m > n$, spanning a sufficiently wide interval $[0, \nu_{\max}]$

of frequencies. [1] This allows to express $\mathbf{f}(\nu)$ as

$$\mathbf{f}(\nu) = \mathbf{A}\mathbf{b}(\nu), \tag{14}$$

where $\mathbf{A}$ is the $n \times m$ matrix of coefficients representing the response using the basis functions $\mathbf{b}(\nu) = (b_1(\nu), \ldots, b_m(\nu))^{\mathrm{T}}$.

Since the eigenvalues ν_k form a growing progression, we can truncate the series (13) at $\nu_s \geq \nu_{\max}$. Substituting the representation (14), we obtain

$$\mathbf{p}(x) = \mathbf{A}(\mathbf{b}(\nu_1), \ldots, \mathbf{b}(\nu_s)) \begin{pmatrix} \phi_1^2(x) \\ \vdots \\ \phi_s^2(x) \end{pmatrix} = \mathbf{A}\mathbf{g}(x) \tag{15}$$

where the $m \times 1$ vector $\mathbf{g}(x)$ with the elements

$$g_j(x) = \sum_{k \geq 1} b_j(\nu_k)\phi_k^2(x) \tag{16}$$

captures all the shape-specific geometric information about the point x. For this reason, we refer to $\mathbf{g}$ as to the *geometry vector* of a point. Note that this representation no longer depends on a specific shape; the matrix of parameters $\mathbf{A}$ describes the same vector of frequency responses on any shape.

3.3 Learning

Let $\mathbf{g} = \mathbf{g}(x)$ be the geometry vector representing some point x; let $\mathbf{g}_+ = \mathbf{g}(x_+)$ be another geometry vector representing a point that is knowingly similar to x (positive); and, finally, let $\mathbf{g}_- = \mathbf{g}(x_-)$ represent a knowingly dissimilar point (negative). We would like to select the matrix of parameters that maximizes the similarity of the descriptors $\mathbf{p} = \mathbf{A}\mathbf{g}$ and $\mathbf{p}_+ = \mathbf{A}\mathbf{g}_+$, and at the same time minimizes the similarity between $\mathbf{p}$ and $\mathbf{p}_- = \mathbf{A}\mathbf{g}_-$. Using the ℓ^2 norm as the similarity criterion, we obtain

$$\begin{aligned} d_\pm^2 &= \|\mathbf{p} - \mathbf{p}_\pm\|^2 = \|\mathbf{A}(\mathbf{g} - \mathbf{g}_\pm)\|^2 \\ &= (\mathbf{g} - \mathbf{g}_\pm)^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}(\mathbf{g} - \mathbf{g}_\pm). \end{aligned} \tag{17}$$

In other words, the Euclidean distance between the descriptors translates into a Mahalanobis distance between the corresponding geometry vectors. The problem of finding the best positive-definite matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ defining the Mahalanobis metric is known as *metric learning* and has been relatively well explored in the literature [32], [33], [34].

1. A finite basis cannot span the space of all responses $\mathbf{f}(\nu)$ even on a finite interval. This basis is merely ought to span a family of sufficiently smooth functions, and to be able to approximate the kernels of the WKS and HKS.

Here, we describe a simple yet efficient learning scheme based on [35], explicitly addressing the desired properties we required from a good spectral descriptor. We aim at finding a matrix $\mathbf{A}$ minimizing the Mahalanobis distance over the set of positive pairs, while maximizing it over the negative ones. Note that the distance depends only on the differences between positive and negative pairs of vectors. Taking expectation over all positive and negative pairs, we obtain [35]

$$\begin{aligned} \mathbb{E}(d_\pm^2) &= \mathbb{E}(\|\mathbf{p} - \mathbf{p}_\pm\|^2) = \mathbb{E}(\mathbf{e}_\pm^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{e}_\pm) \\ &= \mathrm{tr}\,(\mathbf{A}\mathbb{E}(\mathbf{e}_\pm\mathbf{e}_\pm^{\mathrm{T}})\mathbf{A}^{\mathrm{T}}) = \mathrm{tr}\,(\mathbf{A}\mathbf{C}_\pm\mathbf{A}^{\mathrm{T}}), \end{aligned} \tag{18}$$

where $\mathbf{e}_\pm = \mathbf{g} - \mathbf{g}_\pm$, and $\mathbf{C}_\pm$ stands for the second moment matrix of the differences of positive and negative pairs of geometry vectors. In practice, the expectations are replaced by averages over a representative set of difference vectors.

Our goal is to minimize $\mathbb{E}(d_-^2)$ simultaneously maximizing $\mathbb{E}(d_+^2)$. This can be achieved by minimizing the ratio $\mathbb{E}(d_-^2)/\mathbb{E}(d_+^2)$, which is solved by *linear discriminant analysis* (LDA). However, we unfavor this approach as it does not allow control over the tradeoff between sensitivity and specificity. Instead, we propose to minimize the difference

$$\begin{aligned} &(1-\alpha)\mathbb{E}(d_+^2) - \alpha\mathbb{E}(d_-^2) = \\ &\quad \mathrm{tr}\,(\mathbf{A}((1-\alpha)\mathbf{C}_+ - \alpha\mathbf{C}_-)\mathbf{A}^{\mathrm{T}}) = \mathrm{tr}\,(\mathbf{A}\mathbf{D}_\alpha\mathbf{A}^{\mathrm{T}}), \end{aligned} \tag{19}$$

where $0 \leq \alpha \leq 1$ controls the said tradeoff, and $\mathbf{D}_\alpha$ denotes the difference between the positive and the negative covariance matrices.

Note that since the scale of $\mathbf{A}$ is arbitrary, a trivial solution can be obtained. Even when fixing the scale, the solution will be a rank-1 matrix corresponding to the smallest eigenvector of $\mathbf{D}_\alpha$. While this can be avoided by arbitrarily demanding orthonormality of $\mathbf{A}$ (as done in [35]), such a remedy is completely artificial.

Instead, we remind that one of the desired properties of a descriptor was *efficiency*. In an efficient descriptor, each dimension should be statistically independent of the others. Replacing statistical independence by the more tractable lack of correlation, we demand

$$\mathbf{I} = \mathbb{E}(\mathbf{p}\mathbf{p}^{\mathrm{T}}) = \mathbf{A}\mathbb{E}(\mathbf{g}\mathbf{g}^{\mathrm{T}})\mathbf{A}^{\mathrm{T}} = \mathbf{A}\mathbf{C}\mathbf{A}^{\mathrm{T}} \tag{20}$$

where expectations are taken over all geometry vectors, and $\mathbf{C}$ denotes the covariance matrix of $\mathbf{g}$. A similar method was used in [36] for content-based image retrieval.

Combining (19) with (20), we obtain the following minimization problem

$$\min_{\mathbf{A}} \mathrm{tr}\left(\mathbf{A}\mathbf{D}_\alpha\mathbf{A}^\mathrm{T}\right) \quad \mathrm{s.t} \quad \mathbf{A}\mathbf{C}\mathbf{A}^\mathrm{T} = \mathbf{I}, \tag{21}$$

which we solve for an $n \times m$ matrix $\mathbf{A}$. The problem has a closed-form algebraic solution, which is easy to derive using variable substitution. Since $\mathbf{C}$ is a positive-definite matrix, we can substitute $\mathbf{B} = \mathbf{A}\mathbf{C}^{1/2}$, obtaining an equivalent minimization problem

$$\min_{\mathbf{B}} \mathrm{tr}\left(\mathbf{B}\mathbf{C}^{-1/2}\mathbf{D}_\alpha\mathbf{C}^{-1/2}\mathbf{B}^\mathrm{T}\right) \quad \mathrm{s.t} \quad \mathbf{B}\mathbf{B}^\mathrm{T} = \mathbf{I}, \tag{22}$$

($\mathbf{C}$ is symmetric and so is its square-root $\mathbf{C}^{1/2}$; we therefore keep writing $\mathbf{C}^{-1/2}$ instead of its transpose). Let us denote by $\mathbf{C}^{-1/2}\mathbf{D}_\alpha\mathbf{C}^{-1/2} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\mathrm{T}$ the eigendecomposition of the scaled covariance difference, with the eigenvalues $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ sorted in ascending order, and the corresponding orthonormal eigenvectors $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$. The solution to (22) is given by the first n smallest eigenvectors, $\mathbf{B} = \mathbf{U}_n^\mathrm{T} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)^\mathrm{T}$. Note that one must ensure that all the eigenvectors correspond to negative eigenvalues; if this is not the case, n has to be reduced. Finally, the solution to our original problem (21) follows straightforwardly as

$$\mathbf{A} = \mathbf{U}_n^\mathrm{T}\mathbf{C}^{-1/2}. \tag{23}$$

3.4 Training set

So far, we have described a learning scheme allowing to construct efficient spectral descriptors with uncorrelated elements based on covariances of geometry vectors describing positive and negative pairs of points. Having no practical possibility to model the statistics of these vectors, their covariance matrices have to be computed empirically from a training set of positive and negative examples. The construction of such a set is therefore crucial for obtaining a good descriptor. In what follows, we describe how to construct the training set in order to achieve each of the desired properties mentioned before.

Localization. Let x be a point on a training shape X, and $B_r(x)$ is the geodesic metric ball of radius r centered at x. We fix a pair of radii $r < R$ and deem all points $x_+ \in B_r(x)$ positive, while deeming negative all $x_- \notin B_R(x)$. Points lying in the ring $B_R(x) \setminus B_r(x)$ are excluded from both sets. If the shape possesses an intrinsic symmetry $\varphi : X \to X$, then $B_r(\varphi(x))$ is also included in the positive set, while $B_R(\varphi(x))$ is excluded from the negative set.[2]

2. Our practice shows that when symmetries are ignored, the trained descriptors cede only a tiny amount of performance. We attribute this to the fact that the amount of incorrectly labeled positives is overwhelmed by the amount of correctly labeled positives and negatives.

The training set is created by sampling many reference points and corresponding positive and negative points on a collection of representative shapes. The selection of r and R gives explicit control over the localization capability of the descriptor.

Discriminativity. Let X and X_- be knowingly dissimilar shapes (i.e., belonging to different classes we would like to tell apart). A random point x on X and a random point x_- on X_- are deemed negative. The training set is created by sampling many random pairs of points on knowingly dissimilar pairs of shapes.

Insensitivty to transformations. Let X be a shape and X_+ its transformation belonging to a class of transformations insensitivity to which is desired. We further assume to be given a correspondence $\varphi : X \to X_+$ between the shapes. A random point x on X and the corresponding point $x_+ = \varphi(x)$ on X_+ are deemed positive. The training set is created by sampling many points on a collection of shapes, paired with corresponding points on the transformed versions of the same shape.

The combination of the positive and negative sets constructed this way allows to train for descriptor localization, discriminativity, and transformation insensitivity properties.

3.5 Sensitivity-Specificity tradeoff

The proposed learning scheme allows simple control over the tradeoff between the sensitivity and the specificity of the descriptor through the parameter α. The bigger is α, the bigger is the relative influence of $\mathbf{C}_-$ compared to $\mathbf{C}_+$. Therefore, for large values of α, the descriptor will emphasize producing large distances on the negative set (low false positive rate), while trying to keep small distances on the positive set (high true positive rate). As the result, high sensitivity is obtained. For small values of α, the converse is observed: the descriptor emphasizes performance on the positive set, resulting in higher specificity.

In order to select the optimal α for a highly-sensitive descriptor, we empirically compute the false negative rate at some small fixed false positive rate (e.g., 1% or 0.1%) and select the α for which it is minimized. For highly-specific descriptors, α is selected to minimize the false positive rate at some small false negative rate. The behavior of the error rates as a function of α is illustrated in Figure 5.

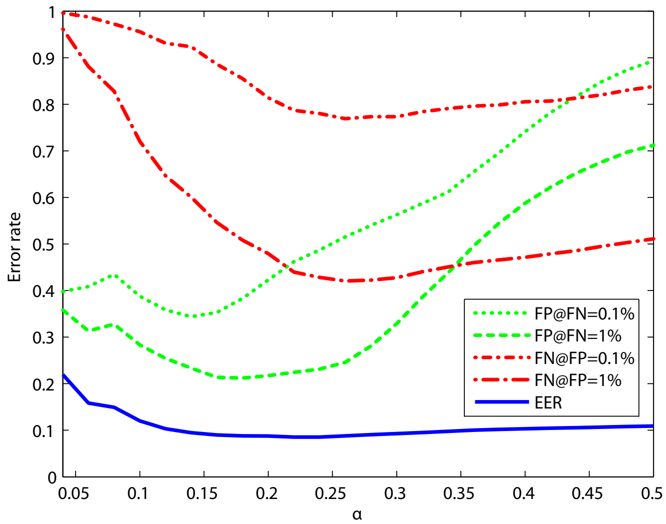

Fig. 5. Error rates on the TOSCA shapes as a function of the parameter α. Large values of α result in high sensitivity, while for small values high specificity is obtained. False positives (FP) increase with α, even though values are still low for the optimal false negatives (FN) at $\alpha \approx 0.25$ that was chosen for the rest of the experiments.

4 Experimental results

4.1 Datasets

The experiments reported in the sequel were performed on the TOSCA [37] and SCAPE [38] datasets. The TOSCA dataset comprises 7 shape classes (centaur, horse, two male figures, female, cat, and dog). In each such class, an extrinsically symmetric "null" shape underwent a few different near-isometric deformations. Typical vertex count ranges from $5,000$ to $50,000$. The SCAPE dataset contains a scanned human figure in about 70 different poses, each containing $12,500$ vertices. Both datasets include vertex-wise correspondences between all deformed instances of the shapes belonging to the same class, and have compatible triangulation. In order to reduce computational and storage complexity, shapes with over $10,000$ vertices were downsampled maintaining compatible triangulations and groundtruth correspondences. SCAPE shapes were scaled the have roughly the same size as TOSCA shapes.

We used the finite elements scheme [27] to compute the first 300 eigenvalues and eigenvectors of the Laplace-Beltrami operator on each shape. Neumann boundary conditions were used. The range of frequencies $\nu_{\max}$ was set to the maximal value of ν_{300} over the entire set of training shapes. The interval was evenly divided into $m = 150$ segments and the cubic spline basis was used as $\{b_j(\nu)\}$. The resulting geometric-vectors $\mathbf{g}$ were normalized to have unit ℓ^2 length. Unless specified otherwise, 16-dimensional descriptors were computed.

For comparison, we also evaluated the HKS and WKS descriptors. The HKS time scales were optimized according to [4]. The WKS energy levels and the variance σ^2 were set as described in [20]. For the fairness of comparison, Euclidean distance was used for all descriptors.

Training sets. The training sets were built from 150-dimensional geometric vector triplets of the form $(\mathbf{g}, \mathbf{g}_+, \mathbf{g}_-)$ as described Section 3.4. We used the farthest point sampling (FPS) strategy [39] with the geodesic distances to select 1000 uniformly placed points on the shape. To each such point, we paired 50 positive points sampled at random from the ball of radius r, and 50 negative points. Half of the negatives were selected from the "near" ring lying between the radii R and $4R$ around the central point; another half was filled with points farther than $4R$. We found that this sampling strategy emphasizes the locality of the descriptor. The radii r and R were set to about 1% and 2% of the average intrinsic shape diameter, respectively. A total of $98,750$ triplets was generated on the TOSCA set and $99,550$ on SCAPE. On the TOSCA set, we used the female and one of the male shapes (David) for training. On the SCAPE set, we used shapes $20-29$ and $50-70$ – an arbitrary choice motivated mainly by visual considerations.

Test sets. Quantitative descriptor performance evaluation was performed on a selection of 1000 points from the shapes selected using FPS in the descriptor space, similar to the experiment reported in [20]. For the fairness of comparison, the latter selection process was done for all the descriptors under test, and the union was used in the evaluation. The second male shape (Michael) was used for test on the TOSCA set. On SCAPE, the remaining shapes not used for training were used for test.

4.2 Evaluation methodology

We use two quantitative and two qualitative criteria to evaluate the performance of the learned descriptors.

Receiver operating characteristic (ROC). For each positive and negative pair of descriptors $(\mathbf{p}, \mathbf{p}_+)$ and $(\mathbf{p}, \mathbf{p}_-)$ from the test set, we measure the corresponding ℓ^2 distances, $d_\pm = \|\mathbf{p} - \mathbf{p}_\pm\|_2$. Deeming as "positive" all pairs with the distance below a threshold τ, and "negative" otherwise produces a measure of the incorrectly classified negatives, the *false positive rate* (FPR) defined as

$$\mathrm{FPR}(\tau) = \frac{|\{d_- < \tau\}|}{|\{d_-\}|}, \quad (24)$$

Similarly, the *false negative rate* (FNR) is defined as

$$\mathrm{FNR}(\tau) = \frac{|\{d_+ \geq \tau\}|}{|\{d_+\}|}. \quad (25)$$

The complementary *true positive* and *true negative* rates are computed as $\mathrm{TPR} = 1 - \mathrm{FNR}$ and $\mathrm{TNR} = 1 - \mathrm{FPR}$, respectively. The ROC curve is defined as $(\mathrm{FPR}(\tau), \mathrm{TPR}(\tau))$, varying the value of the threshold τ. To define the descriptor performance by a single number, it is customary to evaluate the FPR at some low FNR (usually, 1% or 0.1%), and vice versa, the FNR at some low FPR. The term *equal error rate* (EER) refers to the point on ROC curve at which FPR equals FNR.

Cumulative match characteristic (CMC). The CMC curve evaluates the probability of finding the correct match within the first k best matches. The *hit rate* at k is calculated by sorting all the distances $\{d_+\} \cup \{d_-\}$ in ascending order, and evaluating the percentage of positives in the first k distances. The CMC is a monotonically increasing curve of the hit rate as a function of k.

Similarity map. We used similarity maps to visualize and qualitatively assess the localization capabilities of different descriptors. One point on a shape is selected as reference, and the remaining points on the shape are colored according to the distance between their descriptor to the descriptor at the reference point. We also show several other shapes colored according to the distance in the descriptor space from each point on the shape to the reference point on the first shape. Since the range of distances can be greatly affected by a few high outliers, the color map is saturated at the median distance.

Spectral matching. We put the descriptor to the actual test of generating correspondences using a method similar to [30]. We stop generating matches from the latter method when the maximal geodesic distortion gets higher than a certain threshold.

4.3 Experiments

To assess the influence of the parameter α, we measured the error rates of the descriptors learned with different values of the parameter. The TOSCA dataset was used for training and testing. The results are summarized in Figure 5. From this experiment, we selected the value of $\alpha = 0.25$ giving the lowest FNR at 1% and 0.1% FPR.

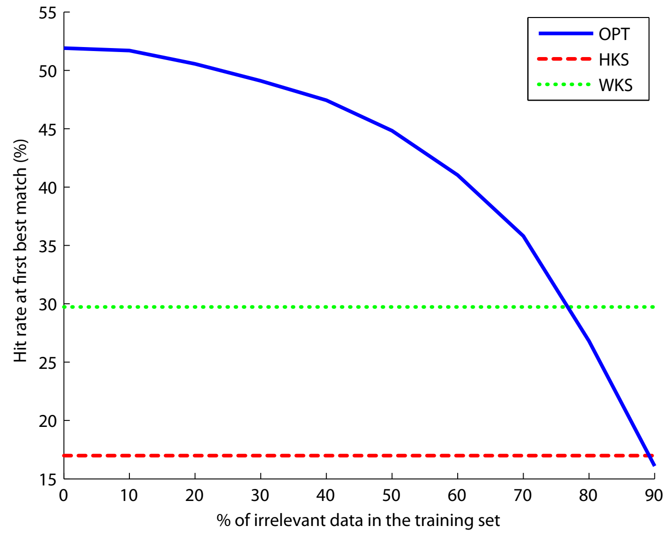

Fig. 6. Hit rate at the first best match on the TOSCA shapes for the optimal descriptor learned on a training set contaminated by a different amount of irrelevant shapes. The HKS and WKS descriptors are shown for reference.

Figure 4 depicts the performance of the learned descriptors as well as of the HKS and WKS for different number of dimensions. We observed that the learned descriptor gives excellent performance (over 50% hit rate at first match) even for as little as 16 dimensions, while the HKS and the WKS perform significantly worse (lower than 25% hit rate). Increasing the number of dimensions improves the performance of the WKS, while the HKS shows no sign of performance, and even a slight degradation. For dimensionality above 100, the WKS approaches the learned descriptor, a phenomenon deserving further investigation.

The CMC and ROC curves of the learned 16-dimensional descriptors are compared to those of the HKS and WKS in Figure 8 on the TOSCA data, and Figure 9 on the SCAPE data. In the latter figure, we also show the performance of the descriptor learned on TOSCA and tested on SCAPE. Such a transfer of the learned descriptor is possible with a negligible drop in the CMC and only a small degradation of the ROC. This given an experimental evidence of the generalization power of the descriptors. For a qualitative assessment, we show similarity maps of different descriptors in Figure 2

To study the influence of the content of the training set on the performance of the descriptor, we trained a 16-dimensional descriptor on training sets containing a part of the TOSCA training set, and the remaining part filled with geometric vectors from irrelevant shapes drawn from the Princeton shape benchmark [40] and belonging to hundreds of non-human classes. The performance evaluated in terms of hit rate at first match is depicted as a function of percentage of irrelevant shapes in the training set in Figure 6. The performance drops with the increase of the "contamination level". Still, even when the training set is contaminated by 50% of irrelevant shapes, the learned descriptors significantly outperform the WKS and the HKS.

To study the influence of missing data on the performance of the descriptor, tested the descriptors from the previous experiment on a corpus of TOSCA shapes with removed parts. The CMC curve in Figure 7 shows that the optimal descriptor is affected less by the missing parts, while the HKS is affected the most due to its lack of locality.

Finally, in order to test the performance of different descriptors in a correspondence task, we performed a simple shape matching experiment on TOSCA shapes. 100 points were sampled on one of the shapes, and were each paired with the best 20 matches on the other shape using the ℓ^2 distances between the descriptors. The latter list of possible matches was used to construct pairwise affinity matrix. We then invoked our variant of the spectral correspondence algorithm [30] to compute the point-to-point correspondence between the shapes. Matches resulting in geodesic distortion higher than 10% of the shape diameter were rejected. A comparison between the correspondences produced by the learned descriptors, the HKS and the WKS is shown in Figure 3. Our experiments show that the learned descriptors consistently produce more correct matches.

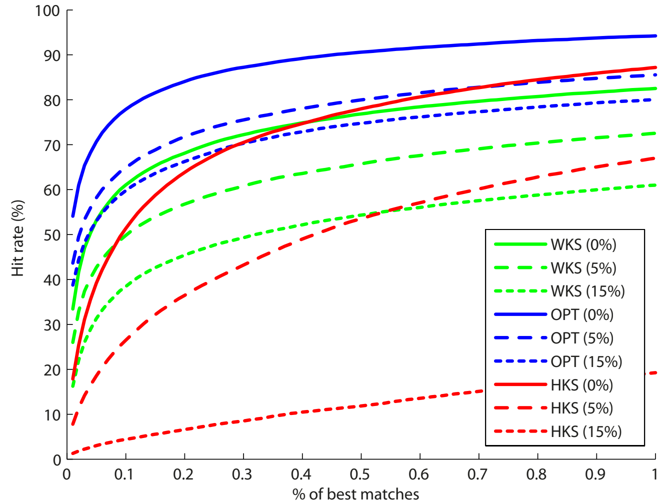

Fig. 7. CMC curves of the 16-dimensional HKS, WKS and optimal descriptors on the TOSCA shapes with partially missing data. The amount of missing data in percent is specified in parentheses.

5 Conclusion

We presented a generic framework for the construction of feature descriptors for deformable shapes based on their spectral properties. The proposed descriptor is computed by applying a bank of "filters" to the shape's geometric features at different "frequencies", and it generalizes the heat and wave kernel signatures. We also showed a learning approach allowing to construct filters for optimized specific shape analysis tasks, resembling in its spirit optimal signal filtering by means of a Wiener filter.

We formulated the learning approach in terms of the ℓ^2 distance and related it to Mahalanobis metric learning. While the adopted algebraic solution gave good results, other Mahalanobis metric learning approaches, such as the maximum-margin learning [33] can be readily used. Some of these metric learning approaches were designed with a specific task in mind (e.g., ranking), and might be beneficial for the construction of spectral descriptors in some applications. Evidence shows that distances other than the Euclidean one (e.g., the ℓ^1 distance) improve the performance of spectral descriptors. Also, applications where compact and easily searchable descriptors are of importance may benefit from hash learning techniques [41], essentially based on the Hamming distance. We intend to explore alternative learning frameworks and different distances in follow-up studies.

While the main focus of this paper was the construction of the descriptor itself, in future studies we are going to explore its performance in real shape retrieval and matching tasks. Particularly, in retrieval tasks spectral feature descriptors are used to generate global shape descriptors by means of vector quantization or sparse coding, a growingly popular alternative in the computer vision community. Taking this highly non-linear process into account when constructing the feature descriptor will also be a subject of our future research.

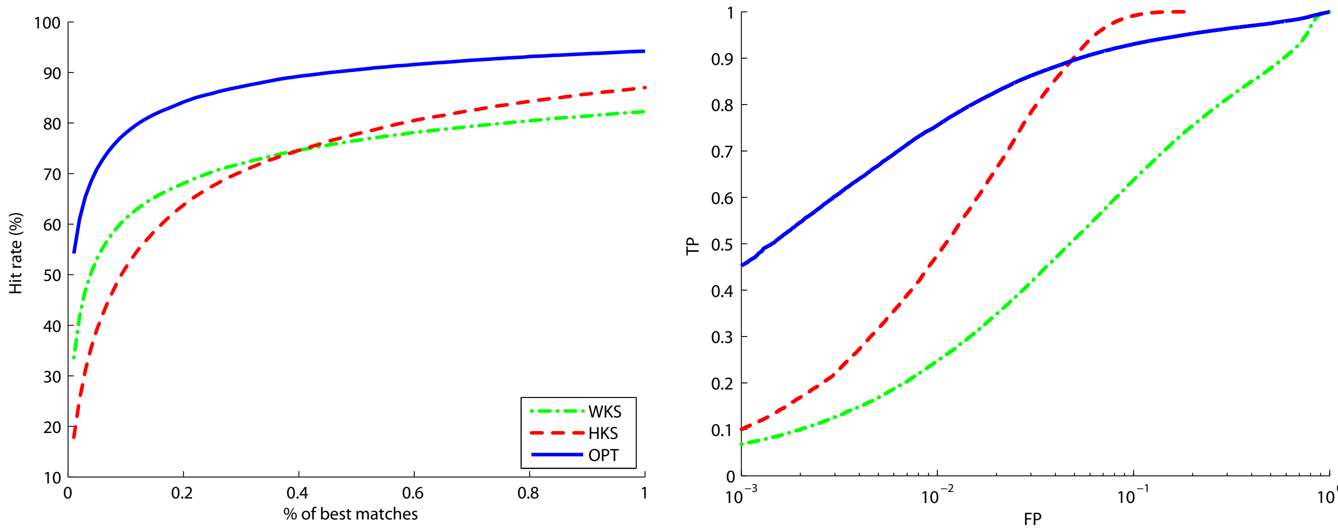

Fig. 8. CMC (left) and ROC (right) curves of the 16-dimensional HKS, WKS and optimal descriptors on the TOSCA shapes.

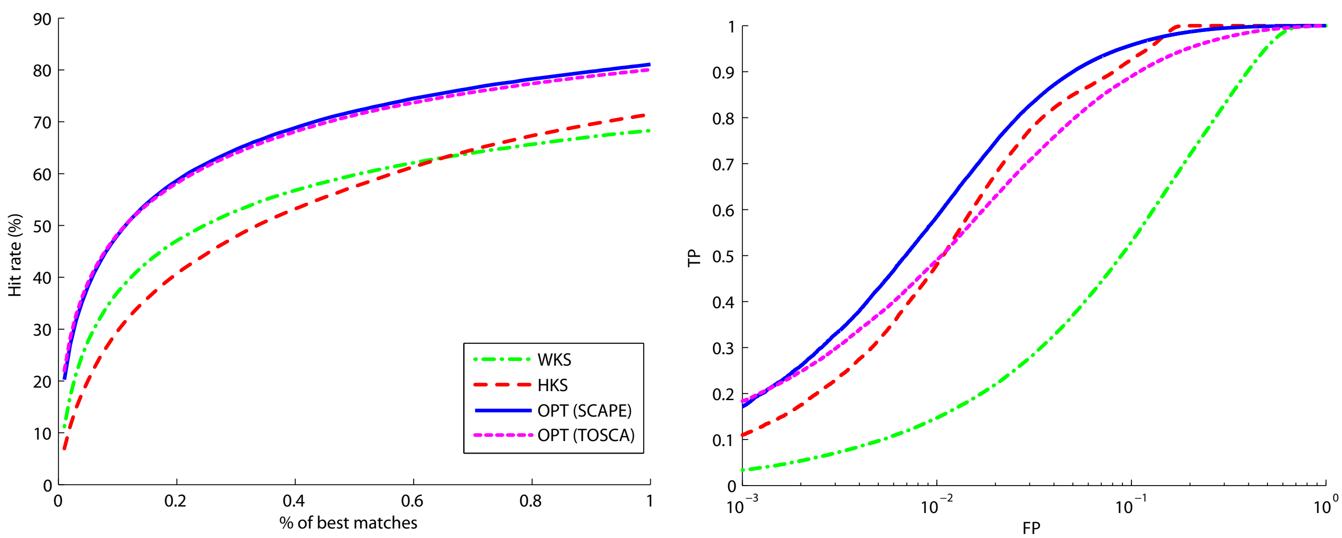

Fig. 9. CMC (left) and ROC (right) curves of the 16-dimensional HKS, WKS and optimal descriptors on the SCAPE shapes. Observe that descriptors trained on the TOSCA set have negligibly lower hit rate compared to the ones trained on the SCAPE data.

REFERENCES

[1] N. Gelfand, N. J. Mitra, L. J. Guibas, and H. Pottmann, "Robust global registration," in *Proceedings of the third Eurographics symposium on Geometry processing*, 2005, pp. 197–206.

[2] C. Wang, A. M. Bronstein, M. M. Bronstein, and N. Paragios, "Discrete minimum distortion correspondence problems for non-rigid shape matching," in *Proc. Scale Space and Variational Methods (SSVM)*, vol. 6667. Springer Berlin / Heidelberg, 2011, pp. 580–591.

[3] N. J. Mitra, L. Guibas, J. Giesen, and M. Pauly, "Probabilistic fingerprints for shapes," in *ACM International Conference Proceeding Series*, vol. 256. Citeseer, 2006, pp. 121–130.

[4] A. Bronstein, M. Bronstein, L. Guibas, and M. Ovsjanikov, "Shape google: geometric words and expressions for invariant shape retrieval," *ACM Transactions on Graphics (TOG)*, vol. 30, no. 1, p. 1, 2011.

[5] P. Skraba, M. Ovsjanikov, F. Chazal, and L. Guibas, "Persistence-based segmentation of deformable shapes," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010, pp. 45–52.

[6] S. Belongie, J. Malik, and J. Puzicha, "Shape context: A new descriptor for shape matching and object recognition," *Advances*

in neural information processing systems, pp. 831–837, 2001.

[7] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.

[8] S. Manay, B. Hong, A. Yezzi, and S. Soatto, "Integral invariant signatures," *Lecture Notes in Computer Science*, pp. 87–99, 2004.

[9] M. Pauly, R. Keiser, and M. Gross, "Multi-scale feature extraction on point-sampled surfaces," in *Computer Graphics Forum*, vol. 22, no. 3, 2003, pp. 281–289.

[10] A. Hamza and H. Krim, "Geodesic object representation and recognition," in *Discrete Geometry for Computer Imagery*, 2003, pp. 378–387.

[11] A. Elad and R. Kimmel, "On bending invariant signatures for surfaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 1285–1311, 2003.

[12] Y. Lipman and T. Funkhouser, "Möbius voting for surface correspondence," in *ACM Transactions on Graphics (TOG)*, vol. 28, no. 3, 2009, p. 72.

[13] P. Bérard, G. Besson, and S. Gallot, "Embedding Riemannian manifolds by their heat kernel," *Geometric and Functional Analysis*, vol. 4, no. 4, pp. 373–398, 1994.

[14] R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006.

[15] F. Mémoli, "Spectral gromov-wasserstein distances for shape matching," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 256–263.

[16] A. M. Bronstein, M. M. Bronstein, R. Kimmel, M. Mahmoudi, and G. Sapiro, "A gromov-hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching," *International Journal of Computer Vision*, vol. 89, no. 2-3, pp. 266–286, 2010.

[17] B. Lévy, "Laplace-beltrami eigenfunctions towards an algorithm that," in *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*. IEEE, 2006, pp. 13–13.

[18] R. Rustamov, "Laplace-Beltrami eigenfunctions for deformation invariant shape representation," in *Proc. Symp. on Geometry Processing (SGP)*, 2007, pp. 225–233.

[19] J. Sun, M. Ovsjanikov, and L. Guibas, "A Concise and Provably Informative Multi-Scale Signature Based on Heat Diffusion," in *Computer Graphics Forum*, vol. 28, no. 5, 2009, pp. 1383–1392.

[20] M. Aubry, U. Schlickewei, and D. Cremers, "The wave kernel signature: A quantum mechanical approach to shape analysis," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1626–1633.

[21] A. Bronstein, M. Bronstein, U. Castellani, A. Dubrovina, L. Guibas, R. Horaud, R. Kimmel, D. Knossow, E. von Lavante, D. Mateus *et al.*, "SHREC 2010: robust correspondence benchmark," in *Eurographics Workshop on 3D Object Retrieval (3DOR'10)*, 2010, pp. 87–91.

[22] A. Bronstein, M. Bronstein, U. Castellani, B. Falcidieno, A. Fusiello, A. Godil, L. Guibas, I. Kokkinos, Z. Lian, M. Ovsjanikov *et al.*, "SHREC 2010: robust large-scale shape retrieval benchmark," in *Eurographics Workshop on 3D Object Retrieval (3DOR'10)*, 2010, pp. 71–78.

[23] T. Dey, K. Li, C. Luo, P. Ranjan, I. Safa, and Y. Wang, "Persistent heat signature for pose-oblivious matching of incomplete models," in *Computer Graphics Forum*, vol. 29, no. 5, 2010, pp. 1545–1554.

[24] M. M. Bronstein and I. Kokkinos, "Scale-invariant heat kernel signatures for non-rigid shape recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1704–1711.

[25] J. Aflalo, A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Deformable shape retrieval by learning diffusion kernels," in *Proc. Scale Space and Variational Methods (SSVM)*, 2011.

[26] M. Kac, "Can one hear the shape of a drum?" *The American Mathematical Monthly*, vol. 73, no. 4, pp. 1–23, 1966.

[27] M. Reuter, F. Wolter, and N. Peinecke, "Laplace-Beltrami spectra as "Shape-DNA" of surfaces and solids," *Computer-Aided Design*, vol. 38, no. 4, pp. 342–366, 2006.

[28] A. Sharma and R. Horaud, "Shape matching based on diffusion embedding and on mutual isometric consistency," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 2010, pp. 29–36.

[29] D. Raviv, M. M. Bronstein, A. M. Bronstein, and R. Kimmel, "Volumetric heat kernel signatures," in *Proceedings of the ACM workshop on 3D object retrieval*. ACM, 2010, pp. 39–44.

[30] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *International Conference of Computer Vision (ICCV)*, vol. 2, 2005, pp. 1482 – 1489.

[31] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series, with engineering applications*. MIT Press, 1949.

[32] L. Yang and R. Jin, "Distance metric learning: A comprehensive survey," *Michigan State Universiy*, pp. 1–51, 2006.

[33] K. Weinberger, J. Blitzer, and L. Saul, "Distance metric learning for large margin nearest neighbor classification," in *Advances in Neural Information Processing Systems 18*. MIT Press, 2005, pp. 1473–1480.

[34] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-theoretic metric learning," in *Proceedings of the 24th international conference on Machine learning*. ACM, 2007, pp. 209–216.

[35] C. Strecha, A. Bronstein, M. Bronstein, and P. Fua, "LDAHash: Improved matching with smaller descriptors," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 66 –78, 2012.

[36] X. He, D. Cai, and J. Han, "Learning a maximum margin subspace for image retrieval," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 20, no. 2, pp. 189 –201, feb. 2008.

[37] A. Bronstein, M. Bronstein, and R. Kimmel, *Numerical geometry of non-rigid shapes*. Springer, 2008.

[38] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis, "Scape: shape completion and animation of people," in *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3. ACM, 2005, pp. 408–416.

[39] Y. Eldar, M. Lindenbaum, M. Porat, and Y. Y. Zeevi, "The farthest point strategy for progressive image sampling," *Image Processing, IEEE Transactions on*, vol. 6, no. 9, pp. 1305–1315, 1997.

[40] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The princeton shape benchmark," in *Shape Modeling Applications, 2004. Proceedings*. IEEE, 2004, pp. 167–178.

[41] Y. Weiss, A. Torralba, and R. Fergus, "Spectral hashing." Advances in Neural Information Processing Systems, 2008.