Optimal Intrinsic Descriptors for Non-Rigid Shape Analysis

Thomas Windheuser
Matthias Vestner
Emanuele Rodolà
Rudolph Triebel
Daniel Cremers

Computer Vision Group,
Department of Computer Science,
Technische Universität München

Abstract

We propose novel point descriptors for 3D shapes with the potential to match two shapes representing the same object undergoing natural deformations. These deformations are more general than the often assumed isometries, and we use labeled training data to learn optimal descriptors for such cases. Furthermore, instead of explicitly defining the descriptor, we introduce new Mercer kernels, for which we formally show that their corresponding feature space mapping is a generalization of either the Heat Kernel Signature or the Wave Kernel Signature. I.e. the proposed descriptors are guaranteed to be at least as precise as any Heat Kernel Signature or Wave Kernel Signature of any parameterisation. In experiments, we show that our implicitly defined, infinite-dimensional descriptors can better deal with non-isometric deformations than state-of-the-art methods.

1 Introduction

One of the key tasks in computer vision is the definition and determination of similarities between elements of a given domain, e.g. image key points, 2D contours or 3D shapes. Common to all problems of finding similarities is their need to first find an appropriate representation of these elements, and then to find a distance function that assigns small values to pairs of intuitively similar elements, and large values to dissimilar elements. Finding such a distance function becomes easy when the representation of the elements already incorporates this notion of similarity. Therefore, research in this field mainly concentrates in defining good *descriptors*, i.e. mappings into a metric space called the feature space in which the elements can be compared using standard distance functions. In this paper, we concretely investigate descriptors for points on 3D shapes, which are particularly useful to identify similar points on different shapes. This provides the potential to match shapes that represent the same object, but under different deformations. Therefore, we particularly look at descriptors that are invariant under such deformations.

So far, the most widely used kinds of descriptors for this task are the heat kernel signature (HKS) [14] and the wave kernel signature (WKS) [1]. They are based on the spectrum of the Laplace-Beltrami operator of a shape, which only depends on intrinsic properties of the shape. As a result, the HKS and the WKS are invariant under isometric deformations.

However, the class of *natural deformations* does not only consist of isometries, and it can be different for different applications. This problem is often tackled by tuning parameters according to the application. In case of the HKS and the WKS, these parameters are related to the sensitivity of the signature to local and global properties around the point of interest, but not directly to the accuracy of the descriptor. Since the user is usually interested in accuracy there is no straightforward solution to obtain an optimal parameterisation.

Our approach is to learn the optimal parameters of a signature from a set of training shapes that represent the class of deformations to be expected. Furthermore, instead of explicitly defining a descriptor, we introduce similarity measures based on a Mercer kernel, and show that the corresponding, infinite-dimensional descriptors can be regarded as generalizations of the HKS and the WKS. As a result, our trained, generalized descriptor yields a better performance in terms of identifying similar points on shapes with natural deformations.

1.1 Contribution and Related Work

The spectrum of the Laplace-Beltrami (LB) operator underlying our approach has been widely used in shape analysis. Reuter *et al.* [9] proposed to adopt the (truncated) spectrum as a *global* shape descriptor. Assigning each shape a "fingerprint", it was dubbed Shape DNA. Following similar ideas, Rustamov *et al.* [11] introduced the Global Point Signature (GPS). This *local* feature descriptor corresponds at each point of the shape to the values of the LB's eigenfunctions. Similarly to the Shape DNA, GPS is invariant under isometric deformations, but it suffers from well-known phenomena such as sign flips and order changes of the basis functions. The *heat kernel signature* (HKS) was proposed by Sun *et al.* [14] as a more robust alternative. It is a local point signature based on the fundamental solutions of the heat equation on Riemannian manifolds. It is efficient to compute and robust to small perturbations and nearly-isometric deformations of the shape. However, its practical performance directly depends on the parameterisation and, by construction, it is severely limited by its inability to precisely localize features. A direct extension of the HKS to a volumetric counterpart was proposed in [8]. In an attempt to rectify the poor feature localization of the HKS, Aubry *et al.* introduced the *wave kernel signature* (WKS) [1]. In their work, the authors proposed to replace heat diffusion by a different physical model, and analysed the behaviour of a quantum particle over the shape surface as described by the Schrödinger equation on the underlying manifold.

A different perspective was recently taken in [7]. Regarding both the HKS and WKS as *filters* in the Laplace-Beltrami spectral domain, the authors asserted that natural deformations affect the different "shape frequencies" in a different way. In this view, they proposed a learning approach that attempts at separating the information-carrying from the noise-carrying frequencies as measured on a representative dataset for which point-to-point correspondences are known in advance. To our knowledge, [7] currently represents the only effort at applying machine learning techniques for the purpose of feature detection and description in the area of deformable shape analysis. Taking a similar perspective that in principle arbitrary filters can be applied to the LB spectrum, Li and Ben Hamza [6] more recently proposed a *multi-resolution* signature based on a cubic spline wavelet generating kernel. More recent advances include the approach of Rodolà *et al.* [10], who propose the adoption of random forests in order to directly establish *correspondences* among two given shapes, rather than providing an optimal descriptor to employ in subsequent steps of the matching pipeline.

In our approach we use the LMNN method introduced by Weinberger and Saul [15] to find the best metric in the descriptor space. For a recent survey on this active field of research

called *metric learning* we refer to Bellet *et al.* [2]. In contrast to [7] we propose an infinite dimensional descriptor. The LMNN optimization in infinite dimensions is based on work done by Chatpatanasiri *et al.* [4] using the kernel PCA method [13]. We would like to point out the papers contributions, which distinguishes our proposed method especially from [7]:

- The method eliminates the need of tuning descriptor parameters. Neither does it have time parameters such as the HKS, nor do we need to choose the dimensionality of the descriptor as in [7]. In contrast, the adjustment of the descriptor is completely driven by the data, i. e. the shapes' deformations fed to the training process. The only two parameters of the objective function are directly related to the descriptor precision. Experiments suggest they can be fixed to constant values across applications, making the framework virtually parameter free.
- The method is a true generalization of the WKS and HKS and can potentially generalize other descriptors as well. Most importantly, we formally show that the proposed descriptors are guaranteed to be at least as accurate as WKS and HKS under any parameterisation with respect to the given shapes. Applications using WKS or HKS can avoid the parameter tuning problem by plugging in the proposed descriptor and are guaranteed to get optimal precision.

2 Intrinsic Geometry and Point Descriptors

We model a shape as a compact two-dimensional Riemannian manifold $\mathcal{M}$ without boundary. Our aim is to design a *pointwise descriptor*, i.e. a function ϕ that assigns to each point on the shape an element of a metric space, the *descriptor space*, such that corresponding points on different shapes are assigned similar descriptor values, while at the same time non-corresponding points are assigned dissimilar descriptors. The descriptor we propose is based on the spectrum of the *Laplace-Beltrami operator* $\Delta_{\mathcal{M}} = -\mathrm{div}_{\mathcal{M}}(\nabla_{\mathcal{M}})$. Being a compact, symmetric operator, the spectrum of $\Delta_{\mathcal{M}}$ consists of real eigenvalues $\lambda_1, \lambda_2, \ldots$ and the corresponding eigenfunctions $\gamma_1, \gamma_2, \ldots$ can be chosen to be orthonormal. Moreover, $\Delta_{\mathcal{M}}$ is a non-negative operator with a one-dimensional kernel, so we can order the eigenvalues $0 = \lambda_1 < \lambda_2 \leq \ldots$ and assign to each point $\mathbf{x} \in \mathcal{M}$ a vector $\mathbf{p} \in \mathbb{R}^{2K}$, $\mathbf{p} = (\lambda_1, \ldots, \lambda_K, \gamma_1(\mathbf{x}), \ldots, \gamma_K(\mathbf{x}))$. This vector is a representation of $\mathbf{x}$ that is invariant under isometric deformations of the shape. Since the class of isometric deformations includes reflections, any feature descriptor based on this representation will assign identical values to a point and its symmetric counterpart, whenever shapes exhibit bilateral intrinsic symmetries. In the following, we present two descriptors based on this representation that have been shown to perform well in the near isometric case. Later, we introduce descriptors that generalize both in order to tackle more general deformations.

2.1 Heat and Wave Kernel Signature

The *heat kernel signature* (HKS) [14] is defined as a mapping $\mathbf{H}_\theta : \mathbb{R}^{2K} \to \mathbb{R}^T$, parametrized by $\theta = \{t_1, \ldots, t_T\} \subset \mathbb{R}_+$. Each component $\mathbf{H}_\theta(\mathbf{p})_l = h_{t_l}(\mathbf{p})$ is a real-valued function $h_{t_l} : \mathbb{R}^{2K} \to \mathbb{R}$ defined by

$$h_{t_l}(\mathbf{p}) = \sum_{j=1}^{K} e^{-\lambda_j t_l} \gamma_j(\mathbf{x})^2. \tag{1}$$

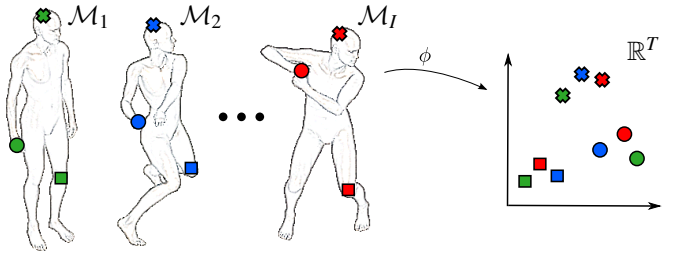


Figure 1: **Comparing points with a point descriptor:** The point descriptor $\phi : \mathcal{P} \to \mathbb{R}^T$ takes points from a set of shapes and maps them to the high-dimensional vector space $\mathbb{R}^T$. Ideally, the descriptors of points that are at corresponding locations on the shapes should have a small distance in the descriptor space $\mathbb{R}^T$. Points at distinct locations on the shapes are mapped to distinct locations in the descriptor space.

Physically, this can be interpreted as the amount of heat remaining at point $\mathbf{x}$ after time t_l when starting with a unit heat source δ_x at the very same point.

Similar to the HKS the *wave kernel signature* (WKS) [2] is a parametrized mapping $\mathbf{W}_\theta : \mathbb{R}^{2K} \to \mathbb{R}^T$. Here the parameters $\theta = \{(\mu_l, \sigma_l^2) | l = 1, \ldots, T\} \subset \mathbb{R}^2_+$ describe mean and variance of probability distributions given by their densities $f_{\mu,\sigma}$. Examples are

$$f^{\mathcal{N}}_{\mu,\sigma}(\lambda) \propto \exp(-\frac{(\lambda-\mu)^2}{2\sigma^2}) \quad \text{and} \quad f^{\mathcal{L}}_{\mu,\sigma}(\lambda) \propto \exp(-\frac{(\log(\lambda)-\log(\mu))^2}{2\sigma^2}). \tag{2}$$

The components of the wave kernel signature $\mathbf{W}_\theta(\mathbf{p})_l = w_{\mu_l,\sigma_l}(\mathbf{p})$ are then of the form

$$w_{\mu_l,\sigma_l} : \mathbb{R}^{2K} \to \mathbb{R}, w_{\mu_l,\sigma_l}(\mathbf{p}) = \sum_{j=1}^{K} f_{(\mu_l,\sigma_l)}(\lambda_j)\gamma_j(\mathbf{x})^2. \tag{3}$$

We write $\mathbf{W}^{\mathcal{N}}_\theta(\mathbf{p})_l = w^{\mathcal{N}}_{\mu_l,\sigma_l}(\mathbf{p})$ and $\mathbf{W}^{\mathcal{L}}_\theta(\mathbf{p})_l = w^{\mathcal{L}}_{\mu_l,\sigma_l}(\mathbf{p})$ to denote the choice of the probability distribution. This descriptor has a physical interpretation, namely the average probability (over time) to measure a particle in $\mathbf{x}$, with an initial energy distribution described by $f_{\mu,\sigma}$.

3 Proposed Method

As mentioned above, we aim at identifying corresponding ("similar") points on a given set of different shapes $\{\mathcal{M}_1, \ldots, \mathcal{M}_I\}$. We do this by finding a distance function $d : \mathcal{P} \times \mathcal{P} \to \mathbb{R}_{\geq 0}$, where $\mathcal{P} = \mathcal{M}_1 \cup \ldots \cup \mathcal{M}_I$, so that $d(\mathbf{p},\mathbf{q})$ is small if $\mathbf{p} \in \mathcal{P}$ and $\mathbf{q} \in \mathcal{P}$ [1] are similar and large if they are dissimilar. Given a set θ of T parameters the heat kernel signature $\mathbf{H}_\theta : \mathcal{P} \to \mathbb{R}^T$ maps any point $\mathbf{p}$ to a T-dimensional vector space. Using the ℓ_2-metric in $\mathbb{R}^T$, we can then define d as $d_{\mathbf{H}}(\mathbf{p},\mathbf{q}) := \|\mathbf{H}_\theta(\mathbf{p}) - \mathbf{H}_\theta(\mathbf{q})\|_2$. Similarly, we can define a distance $d_{\mathbf{W}}$ using the wave kernel signature, i.e. $d_{\mathbf{W}}(\mathbf{p},\mathbf{q}) := \|\mathbf{W}_\theta(\mathbf{p}) - \mathbf{W}_\theta(\mathbf{q})\|_2$. For a collection of shapes that arise from isometric transformations of one base shape, the distance functions $d_{\mathbf{H}}, d_{\mathbf{W}}$ fit very well. From the discussion of Section 2 it is easy to see that $d_{\mathbf{H}}(\mathbf{p},\mathbf{q}) = d_{\mathbf{W}}(\mathbf{p},\mathbf{q}) = 0$ if the isometric transformation $t : \mathcal{M}_1 \to \mathcal{M}_2$ deforming shape $\mathcal{M}_1$ into shape $\mathcal{M}_2$ maps $\mathbf{p} \in \mathcal{M}_1$ to $t(\mathbf{p}) = \mathbf{q} \in \mathcal{M}_2$. This concept is illustrated in Fig. 1 for a general descriptor $\phi : \mathcal{P} \to \mathbb{R}^T$.

[1] From here on we identify every point $\mathbf{x}$ directly with its representation $\mathbf{p} = (\lambda_1, \ldots, \lambda_K, \gamma_1(\mathbf{x}), \ldots, \gamma_K(\mathbf{x})) \in \mathbb{R}^{2K}$, such that we can say $\mathcal{P} \subset \mathbb{R}^{2K}$.

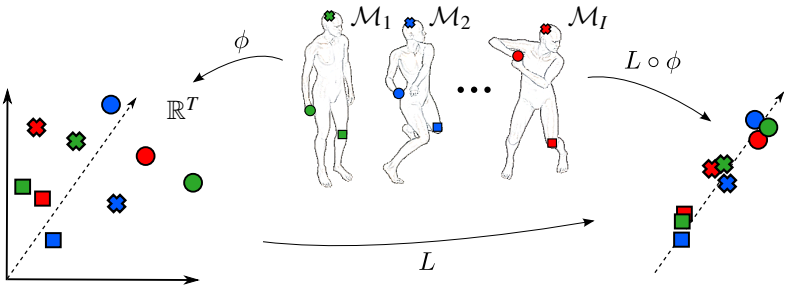


Figure 2: **Optimal distance in the descriptor space:** In general one cannot assume that a given descriptor ϕ groups similar points as well as depicted in Fig. 1. The proposed method optimizes for the distance d_M^ϕ in the descriptor space $\mathbb{R}^T$ such that the point descriptors are grouped as good as possible. Since the symmetric positive semi-definite matrix M has a decomposition $M = L^\top L$, d_M^ϕ is equivalent to transforming $\mathbb{R}^T$ by the linear transformation L and then taking the squared ℓ_2-distance. In the illustration we see that L projects the images of ϕ onto the dotted line resulting in the much better descriptor $L \circ \phi$.

If the shapes in the collection differ by more than isometries there is no obvious solution. From the ill-posed nature of the problem (i.e. the ambiguity in the human notion of similarity) we can even argue that there is no point descriptor that will work out-of-the-box on every shape collection. For this reason many point descriptors have parameters that allow to adapt the descriptor to the shapes under consideration. Because this tuning of parameters can be very tedious, we propose a method that picks the optimal combination of parameters with respect to a given set of shapes.

The general idea of the proposed method is illustrated in Fig. 2. Given a point descriptor $\phi : \mathcal{P} \to \mathbb{R}^T$, we cannot in general assume that ϕ is very discriminative. But we can hope that ϕ is more discriminative along some dimensions of $\mathbb{R}^T$ than along other dimensions. The proposed method finds the linear combinations of dimensions that are as discriminative as possible with respect to the given set of shapes. More precisely the method finds the best distance $d_M^\phi(\mathbf{p},\mathbf{q}) := (\phi(\mathbf{p}) - \phi(\mathbf{q}))^\top M(\phi(\mathbf{p}) - \phi(\mathbf{q}))$, represented by the symmetric positive semi-definite (spsd) matrix M, such that similar points are close to each other and dissimilar points are well separated in the descriptor space. The optimal distance is defined as $\arg\min_{M \in \mathbb{S}^T} E(M)$, where $\mathbb{S}^T \subset \mathbb{R}^{T \times T}$ is the set of all spsd matrices and E is a suitable objective function. As depicted in Fig. 2, this is equivalent to finding the best linear transformation L of the descriptor space $\mathbb{R}^T$: Since M admits the decomposition $M = L^\top L$, we can rewrite $d_M^\phi(\mathbf{p},\mathbf{q}) = (L\phi(\mathbf{p}) - L\phi(\mathbf{q}))^\top (L\phi(\mathbf{p}) - L\phi(\mathbf{q})) = \|L\phi(\mathbf{p}) - L\phi(\mathbf{q})\|_2^2$. Whenever we speak of optimal distance d_M^ϕ, we can also say optimal descriptor $L \circ \phi$.

In the next section the *Large Margin Nearest Neighbour* (LMNN) function is introduced as the objective function E. This choice has two benefits: First, LMNN is a convex function such that we are able to obtain the globally optimal distance. Second, it admits $T = \infty$ using the *Kernel PCA* method [4, 13]. In other words, LMNN can choose from an infinite number of dimensions of the descriptor function ϕ. The main contribution of this paper takes advantage of this fact and is laid out in Section 5: We propose several descriptor functions $\phi : \mathcal{P} \to \mathbb{R}^\infty$ and show that the optimal metric learned with LMNN is as good as any possible HKS and WKS.

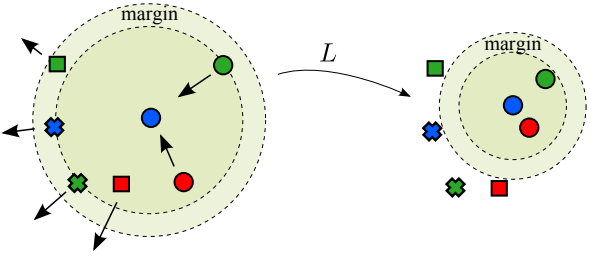


Figure 3: **Optimal LMNN distance:** The neighbourhood of an input sample (blue circle) changes as a result of the training process. In this example, the learned distance is such that the $k = 2$ nearest intra-class neighbours (green and red circles) lie within a smaller radius after application of the linear mapping L. Similarly, the extra-class neighbours are left outside this optimized neighbourhood by a fixed margin. Recall that the linear transformation L is an equivalent representation of distance matrix M.

4 Metric Optimization

In the optimization step the objective function E is optimized with respect to a given set of shapes $\mathcal{P} = \{\mathcal{M}_1, \ldots, \mathcal{M}_I\} \subset \mathbb{R}^{2K}$ with known correspondences among its points. The correspondences are expressed as relations: We write $\mathbf{p} \sim \mathbf{q}$ if the two points $\mathbf{p}, \mathbf{q} \in \mathcal{P}$ are in correspondence. Otherwise we write $\mathbf{p} \nsim \mathbf{q}$.

We choose the *large margin nearest-neighbour* function $E_{\text{LMNN}} : \mathbb{S}^T \to \mathbb{R}_{\geq 0}$, introduced by Weinberger *et al.* [15], to evaluate possible distance matrices $M \in \mathbb{S}^T$. It is defined by:

$$E^{\phi}_{\text{LMNN}}(M) = (1-c) \sum_{\substack{i,j: \\ \mathbf{p}_i \sim \mathbf{p}_j}} d^{\phi}_M(\mathbf{p}_i, \mathbf{p}_j) \quad + c \sum_{\substack{i,j,l: \\ \mathbf{p}_i \sim_k \mathbf{p}_j, \mathbf{p}_i \nsim \mathbf{p}_l}} [1 + d^{\phi}_M(\mathbf{p}_i, \mathbf{p}_j) - d^{\phi}_M(\mathbf{p}_i, \mathbf{p}_l)]_{\geq 0}. \tag{4}$$

$\phi : \mathbb{R}^{2K} \to \mathbb{R}^T$ is the descriptor and $c \in (0,1)$ is a real parameter that weights the two objective terms. The first term favours a small distance between descriptors of corresponding points, since we are looking for the minimum of E^{ϕ}_{LMNN}. The second term uses parameter $k \in \mathbb{N}$ and denotes the nearest k corresponding points $\mathbf{p}_j$ around point $\mathbf{p}_i$ by $\mathbf{p}_i \sim_k \mathbf{p}_j$. It draws a margin around these points and penalizes every non-corresponding point entering this margin. Here, $[\cdot]_{\geq 0}$ is the projection onto the non-negative reals. The concept is illustrated in Fig. 3.

For any given finite-dimensional point descriptor ϕ, we can now find an optimal distance by solving for $\arg\min_{M \in \mathbb{S}^T} E^{\phi}_{\text{LMNN}}(M)$ using convex optimization. However, as we will see in the next section, we can give performance guarantees if we allow the function ϕ to map into an *infinite-dimensional* descriptor space. This means optimizing E_{LMNN} over $\mathbb{S}^{\infty}$. While this looks at a first glance computationally intractable, we see that the descriptors $\{\phi(\mathbf{p})\}_{\mathbf{p} \in \mathcal{P}}$ lie in a finite subspace of $\mathbb{R}^{\infty}$. And indeed, the approach of Chatpatanasiri *et al.* [4] allows to optimize E^{ϕ}_{LMNN} globally for $\phi : \mathbb{R}^{2K} \to \mathbb{R}^{\infty}$.

The subspace spanned by descriptors $\Phi = \{\phi(\mathbf{p})\}_{\mathbf{p} \in \mathcal{P}}$ has dimension of at most $|\mathcal{P}|$. Thus, it exists an orthonormal basis $\Psi = \{\psi_j \in \mathbb{R}^{\infty}\}_{j=1}^{n}, n \leq |\mathcal{P}|$, of that subspace. The *kernel PCA method* [13] allows to compute the now finite-dimensional coordinates $\beta \in \mathbb{R}^n$ of $\phi(\mathbf{p})$ with respect to this basis. Using matrix notation it holds:

$$\phi(\mathbf{p}) = \sum_{j=1}^{n} \psi_j \beta_j = \Psi\beta \quad \Rightarrow \quad \beta = \Psi^{\top}\phi(\mathbf{p}) = \sum_{j=1}^{n} \langle \psi_j, \phi(\mathbf{p}) \rangle. \tag{5}$$

Note that, according to [13], we do not need to compute $\phi(\mathbf{p})$, but we only need to be able to compute the *kernel function* $\kappa : \mathbb{R}^{2K} \times \mathbb{R}^{2K} \to \mathbb{R}$ defined as $\kappa(\mathbf{p},\mathbf{q}) = \langle \phi(\mathbf{p}), \phi(\mathbf{q}) \rangle$. Kernel functions of this form are also known as *Mercer kernels* [12].

Now, the representer theorem of Chatpatanasiri *et al.* [4] (Theorem 3) states that the minimum of E_{LMNN} with respect to $\phi(\mathbf{p})$ is equivalent to the minimum with respect to the projections β. Moreover, we can compute the distance between two new descriptors $d_M^{\phi}(\mathbf{v},\mathbf{w})$, where $\mathbf{v},\mathbf{w} \in \mathbb{R}^{2K}$ possibly $\mathbf{v},\mathbf{w} \notin \mathcal{P}$, whenever we can compute the kernel function κ.

5 Generalized Intrinsic Point Descriptors

We propose point descriptors $\phi^{\mathcal{N}}, \phi^{\mathcal{L}}, \phi^{\mathcal{H}}$ and show that the WKS and HKS of arbitrary parameterisations can be expressed by a linear combination of the proposed descriptors.

Proposition 1 (WKS). [2] *Let the point descriptors $\phi^{\mathcal{N}}, \phi^{\mathcal{L}} : \mathbb{R}^{2K} \to \mathbb{R}^{\infty}$ be defined by*

$$\phi^{\mathcal{N}}_{i,j,k}(\mathbf{p}) := \frac{\sqrt{\binom{2j}{k}}\lambda_i^k}{\sqrt{j!}}\gamma_i^2 \,,\; \phi^{\mathcal{L}}_{i,j,k}(\mathbf{p}) := \frac{\sqrt{\binom{2j}{k}}\log(\lambda_i)^k}{\sqrt{j!}}\gamma_i^2 \,, \text{ where } \begin{smallmatrix} i\in\{1,\dots,K\} \\ j\in\{0,\dots,\infty\} \\ k\in\{0,\dots,2j\} \end{smallmatrix}. \tag{6}$$

For any $\mu,\sigma \in \mathbb{R}_+$ there exist vectors $\mathbf{a}^{\mathcal{N}}, \mathbf{a}^{\mathcal{L}} \in \mathbb{R}^{\infty}$, such that $w^{\mathcal{N}}_{\mu,\sigma}(\mathbf{p}) = \langle \mathbf{a}^{\mathcal{N}}, \phi^{\mathcal{N}}(\mathbf{p}) \rangle$ and $w^{\mathcal{L}}_{\mu,\sigma}(\mathbf{p}) = \langle \mathbf{a}^{\mathcal{L}}, \phi^{\mathcal{L}}(\mathbf{p}) \rangle$ for all $\mathbf{p} \in \mathbb{R}^{2K}$.

Proposition 2 (HKS). *Let the point descriptor $\phi^{\mathcal{H}} : \mathbb{R}^{2K} \to \mathbb{R}^{\infty}$ be defined by*

$$\phi^{\mathcal{H}}_{i,j}(\mathbf{p}) := \frac{\lambda_i^j}{\sqrt{j!}}\gamma_i^2 \,, \text{ where } \begin{smallmatrix} i\in\{1,\dots,K\} \\ j\in\{0,\dots,\infty\} \end{smallmatrix}. \tag{7}$$

For any $t \in \mathbb{R}_+$ there exists a vector $\mathbf{a} \in \mathbb{R}^{\infty}$, such that $h_t(\mathbf{p}) = \langle \mathbf{a}, \phi^{\mathcal{H}}(\mathbf{p}) \rangle$ for all $\mathbf{p} \in \mathbb{R}^{2K}$.

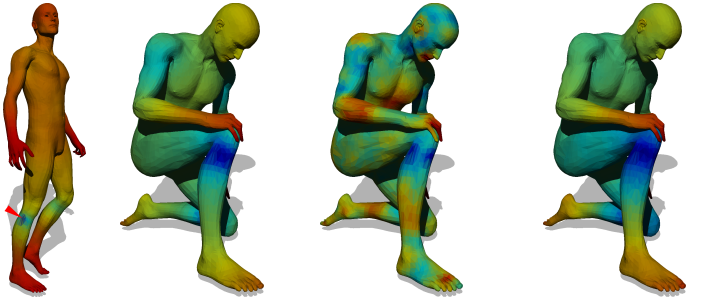
Consequently all $w^{\mathcal{N}}_{\mu,\sigma}$ lie in a vector space of real-valued functions which is spanned by $\{\phi^{\mathcal{N}}_{i,j,k}\}_{i\in\{1,\dots,K\},j\in\{0,\dots,\infty\},k\in\{0,\dots,2j\}}$. Let θ be fixed parameters of WKS $\mathbf{W}^{\mathcal{N}}_{\theta}$ and let $\mathbf{p},\mathbf{p}' \in \mathcal{P}$ be two points. As a result of Proposition 1 there exists a matrix M, such that taking the squared ℓ_2-metric between the descriptors $\mathbf{W}^{\mathcal{N}}_{\theta}(\mathbf{p}), \mathbf{W}^{\mathcal{N}}_{\theta}(\mathbf{p}')$ is identical to taking the distance between $\phi^{\mathcal{N}}(\mathbf{p}), \phi^{\mathcal{N}}(\mathbf{p}')$ with respect to M, i.e. $\|\mathbf{W}^{\mathcal{N}}_{\theta}(\mathbf{p}) - \mathbf{W}^{\mathcal{N}}_{\theta}(\mathbf{p}')\|_2^2 = d_M^{\phi^{\mathcal{N}}}(\phi^{\mathcal{N}}(\mathbf{p}), \phi^{\mathcal{N}}(\mathbf{p}'))$. Since the proposed framework finds the optimum over all possible M, it also optimizes over all possible WKS $\mathbf{W}^{\mathcal{N}}_{\theta}$. The main result follows directly from this observation:

Corollary 3. *Let $M^* = \arg\min_{M\in\mathbb{S}} E^{\phi^{\mathcal{N}}}_{\text{LMNN}}(M)$ be the minimizer of function (4) for a given set of shapes. It holds for all $\mathbf{W}^{\mathcal{N}}_{\theta}$ with any parameterisation θ that $E^{\phi^{\mathcal{N}}}_{\text{LMNN}}(M^*) \leq E^{\mathbf{W}^{\mathcal{N}}_{\theta}}_{\text{LMNN}}(M')$ for all $M' \in \mathbb{S}^T$. Analogously, the same holds for $\phi^{\mathcal{L}}$ and arbitrary $\mathbf{W}^{\mathcal{L}}_{\theta}$ and $\phi^{\mathcal{H}}$ and any $\mathbf{H}_{\theta}$.*

Note that $E^{\mathbf{W}^{\mathcal{N}}_{\theta}}_{\text{LMNN}}(\text{Id})$ measures the performance of the WKS with respect to the squared ℓ_2-metric. In other words: We cannot choose parameters for the wave kernel signature such that it is better than the proposed optimal descriptor on the given set of shapes.

It remains to show that the infinite dimensionality of $\phi^{\mathcal{N}}$'s images is tractable. Remember, it suffices for the kernelized LMNN method to compute the kernel function κ of two points. The next proposition states that this is possible in finite time and space.

[2] All proofs can be found in the supplementary material. Moreover we shorten $\gamma_i(\mathbf{x})$ to γ_i for better readability.



reference shape | wave kernel signature [1] | learned cubic spline descriptor [7] | proposed descriptor with kernel $\kappa^{\mathcal{L}}$

Figure 4: **Qualitative comparison of descriptors:** Distance maps between the descriptor at a reference point (indicated by a red arrow) and the descriptors computed on the shape after deformation. Colours range from blue (small distance) to red (large distance). Qualitatively, WKS and the proposed method do very well at indicating the right location while the cubic spline descriptor exhibits several local minima across the shape. Both test shapes are from the class *michael* (TOSCA), whereas the proposed descriptor and the spline descriptor were trained on the class *david*. The distance map on the reference shape is generated by the proposed method. Colours are in log scale and clipped to the ranges, thus they are not comparable across shapes.

Proposition 4. *Let the kernel function* $\kappa^{\mathcal{N}} : \mathbb{R}^{2K} \times \mathbb{R}^{2K} \to \mathbb{R}$ *express the dot product between any two pairs* $\mathbf{p}, \mathbf{p}' \in \mathbb{R}^{2K}$ *in the descriptor space, i. e.* $\kappa^{\mathcal{N}}(\mathbf{p}, \mathbf{p}') := \langle \phi^{\mathcal{N}}(\mathbf{p}), \phi^{\mathcal{N}}(\mathbf{p}') \rangle$. $\kappa^{\mathcal{N}}$ *can be computed by the expression*

$$\kappa^{\mathcal{N}}(\mathbf{p}, \mathbf{p}') = \sum_{i=1}^{K} \exp((1 + \lambda_i \lambda_i')^2) \gamma_i^2 \gamma_i'^2. \tag{8}$$

Similarly, the kernel functions $\kappa^{\mathcal{L}}(\mathbf{p}, \mathbf{p}') := \langle \phi^{\mathcal{L}}(\mathbf{p}), \phi^{\mathcal{L}}(\mathbf{p}') \rangle$ *and* $\kappa^{\mathcal{H}}(\mathbf{p}, \mathbf{p}') := \langle \phi^{\mathcal{H}}(\mathbf{p}), \phi^{\mathcal{H}}(\mathbf{p}') \rangle$ *can be computed via the identities*

$$\kappa^{\mathcal{L}}(\mathbf{p}, \mathbf{p}') = \sum_{i=1}^{K} \exp((1 + log(\lambda_i) log(\lambda_i'))^2) \gamma_i^2 \gamma_i'^2 \quad \text{and} \quad \kappa^{\mathcal{H}}(\mathbf{p}, \mathbf{p}') = \sum_{i=1}^{K} \exp(\lambda_i \lambda_i') \gamma_i^2 \gamma_i'^2. \tag{9}$$

6 Experiments

In order to evaluate the performance of our method, we conduct two kinds of experiments. The first kind of experiments is aimed at evaluating our approach against several comparable state-of-the-art point descriptors in a deformable shape matching scenario. Specifically, we evaluate the robustness of the descriptors when the underlying shapes are allowed to undergo non-rigid deformations. An example of such deformations is shown in Fig. 4. For this experiment we used shapes from the TOSCA dataset [3], for which ground-truth point-to-point correspondences are available. We built the training set by using only the 7 shapes from the class *david*, restricted to 600 farthest point samples [5] (under the geodesic metric) on each shape. For the test set we used the first 10 shapes from the class *michael*.

The kernelized LMNN algorithm was fed with these samples and the corresponding ground-truth labels; the parameters in Eq. (4) were set to $c = 0.5$ (margin violation trade-off) and $k = 10$ (nearest intra-class neighbours). We used the kernels $\kappa^{\mathcal{L}}$ and $\kappa^{\mathcal{H}}$. In all the

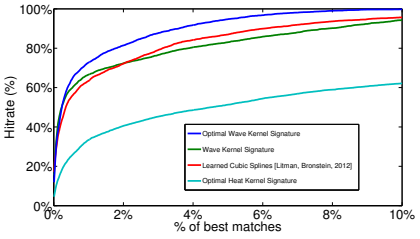


Figure 5: **Descriptor comparison** CMC curves of different descriptors on the TOSCA dataset; the method proposed in this paper with kernel $\kappa^{\mathcal{L}}$ is plotted as a blue curve.

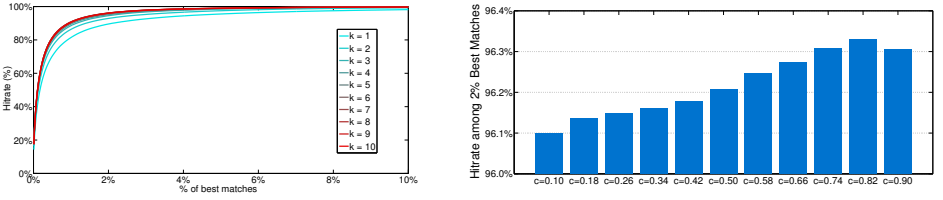


Figure 6: **Parameter sensitivity analysis** **Left:** Precision of the learned descriptor on the test set *michael* with kernel $\kappa^{\mathcal{L}}$, fixed $c = 0.5$ depending on different values for learning parameter k. Precision directly increases with higher values of k. **Right:** Same experimental setup as on the left, with fixed $k = 7$ and different values of c. The plot shows the hitrate when looking at the 2% best matches. The difference in precision among different values of c is only visible in this close-up.

experiments we compare with the method from [7] using the downloadable software made available by the authors. For this method we used the same training set described above, with a cubic spline basis and a descriptor size of 64. In Fig. 5 we report the quantitative performance of the standard WKS, as well as the descriptors learned using the approach of [7] and our approach. Following the experimental setup of [7], we plot for each method the corresponding *cumulative match characteristic* curve (CMC). This curve evaluates the relative frequency of finding the correct match within the first k best matches (expressed in percentage and called *hitrate* in the plot) with respect to the appropriate metric. Note that, in this experiment, only the *exact* matches (according to the ground-truth) are considered correct; this is different from the evaluation appearing in [7], where a candidate match is deemed as "positive" if the distance of the matching descriptors is below a fixed threshold. From the plots we can observe that, while both methods give excellent performance, our learned descriptor with kernel $\kappa^{\mathcal{L}}$ has almost a 10% improvement over the cubic spline approach. In Fig. 4 we show qualitative examples of the similarity maps produced by each method.

In the second type of experiments we evaluate the influence of learning parameters c and k on the descriptor accuracy. We set up the training set as described before and only used kernel $\kappa^{\mathcal{L}}$. In the first round of experiments we keep $c = 0.5$ and train the descriptor with different values $k \in \{1, \dots, 10\}$. The quantitative evaluation on class *michael*, this time evaluated on all points of the shapes, is plotted on the left of Fig. 6. The plot shows how the accuracy of the descriptor increases consistently with higher values of k. Next, we repeat the experiment with fixed $k = 7$ and linearly increasing values of c (Fig. 6, right). We can observe in these plots that the proposed method is largely insensitive to the specific choice of this parameter, whereas larger values for k directly increase the descriptor's accuracy. This observation suggests that we can even eliminate parameter k completely by choosing k always large enough, sucht that $\mathbf{p} \sim \mathbf{q} \Leftrightarrow \mathbf{p} \sim_k \mathbf{q}$ for all $\mathbf{p}, \mathbf{q} \in \mathcal{P}$.

7 Conclusions

We presented a learning-based framework for the construction of optimal intrinsic descriptors on deformable shapes. The proposed approach is based on the intuition that typical feature descriptors carry larger discriminativity along certain dimensions, rather than across the whole space. In order to elicit this discriminative power, we formulated the problem as one of metric learning in an infinite-dimensional descriptor space. The resulting optimization problem is solved optimally via a kernelized large-margin method. The learned descriptors generalize previous proposals, and are *optimal* in the sense that one cannot find a parameterisation of the baseline descriptors that perform better on the given training set. Our descriptors compare favourably with respect to the state of the art on a standard dataset.

References

[1] Mathieu Aubry, Ulrich Schlickewei, and Daniel Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops)*, pages 1626–1633. IEEE, 2011.

[2] Aurélien Bellet, Amaury Habrard, and Marc Sebban. A survey on metric learning for feature vectors and structured data. Technical report, http://arxiv.org/pdf/1306.6709.pdf, 2013.

[3] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel. *Numerical geometry of non-rigid shapes*. Springer, 2008.

[4] Ratthachat Chatpatanasiri, Teesid Korsrilabutr, Pasakorn Tangchanachaianan, and Boonserm Kijsirikul. A new kernelization framework for mahalanobis distance learning algorithms. *Neurocomputing*, 73(10):1570–1579, 2010.

[5] Yuval Eldar, Michael Lindenbaum, Moshe Porat, Senior Member, and Yehoshua Y. Zeevi. The farthest point strategy for progressive image sampling. *IEEE Trans. on Image Processing*, page 1315, 1997.

[6] Chunyuan Li and A. Ben Hamza. A multiresolution descriptor for deformable 3d shape retrieval. *Vis. Comput.*, 29(6-8):513–524, June 2013. ISSN 0178-2789.

[7] Roee Litman and Alexander M Bronstein. Learning spectral descriptors for deformable shape correspondence. *Transactions on Pattern Analysis and Machine Intelligence*, 1 (1), 2013.

[8] Dan Raviv, Michael M. Bronstein, Alexander M. Bronstein, and Ron Kimmel. Volumetric heat kernel signatures. In *Proceedings of the ACM Workshop on 3D Object Retrieval*, 3DOR, pages 39–44, 2010.

[9] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace-Beltrami spectra as 'Shape-DNA' of surfaces and solids. *Comput. Aided Des.*, 38(4):342–366, April 2006.

[10] Emanuele Rodolà, Samuel Rota Bulò, Thomas Windheuser, Matthias Vestner, and Daniel Cremers. Dense non-rigid shape correspondence using random forests. In *IEEE Intl. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[11] Raif M Rustamov. Laplace-beltrami eigenfunctions for deformation invariant shape representation. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, pages 225–233. Eurographics Association, 2007.

[12] Bernhard Schölkopf and Alexander J Smola. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT Press, 2002.

[13] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5):1299–1319, 1998.

[14] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer Graphics Forum*, volume 28, pages 1383–1392. Wiley Online Library, 2009.

[15] Kilian Q Weinberger and Lawrence K Saul. Distance metric learning for large margin nearest neighbor classification. *The Journal of Machine Learning Research*, 10:207–244, 2009.