# Shape Analysis with Anisotropic Windowed Fourier Transform

Simone Melzi
University of Verona
simone.melzi@univr.it

Emanuele Rodolà
USI Lugano
emanuele.rodola@usi.ch

Umberto Castellani
University of Verona
umberto.castellani@univr.it

Michael M. Bronstein
USI Lugano / Tel Aviv University / Intel
michael.bronstein@usi.ch

## Abstract

*We propose Anisotropic Windowed Fourier Transform (AWFT), a framework for localized space-frequency analysis of deformable 3D shapes. With AWFT, we are able to extract meaningful intrinsic localized orientation-sensitive structures on surfaces, and use them in applications such as shape segmentation, salient point detection, feature point description, and matching. Our method outperforms previous approaches in the considered applications.*

## 1. Introduction

Fourier analysis is a tool ubiquitously used in a wide range of problems in mathematics and engineering and is the pillar of classical signal processing [18]. The Fourier basis functions used to compute frequency-domain representations, are in fact eigenfunctions of the Laplace operator. This interpretation allows to naturally generalize Fourier analysis to non-Euclidean domains (manifolds or graphs) by considering the respective Laplacian of these domains [15].

Signal processing on geometric data is an active field of research since the seminal work of Taubin [29], introducing a Laplacian-based approach for mesh smoothing. His approach found extensive application in several other mesh processing tasks, including mesh parametrization [19] and compression [13] to name a few (we refer to [15] for an excellent survey). In the area of shape analysis, recent efforts concentrated on the definition of spectral isometry-invariant shape descriptors for matching and retrieval purposes [22, 26, 6]. Of particular interest is the family of spectral methods that build upon the ideas of diffusion geometry [9] in order to encode local geometric structures. Examples of such methods include the heat kernel signature (HKS) [28, 11] and the wave kernel signature (WKS) [2]. A recent promising trend is based on the use of machine learning methods to learn optimal task-specific feature descriptors from examples [17, 16, 25, 30, 4, 5]. The main advantage of spectral methods consists of their ability to encode *intrinsic* structure of the manifold, thus automatically endowing the respective methods with deformation invariance.

It is worth noting that the majority of spectral methods rely on a *global* frequency analysis, considering the Laplacian eigenfunctions as the basis. *Localized* Fourier analysis (a standard construction in signal processing known as the short-time or windowed Fourier transform) has recently been done on graphs [27] and manifolds [4].

**Contributions.** In this paper, we propose the Anisotropic Windowed Fourier Transform (AWFT) for localized and orientation-sensitive analysis on manifolds. Our work extends the previous works of [27] and [4] by adding directional information, resulting in a construction similar in its spirit to the classical Gabor transform. We show that in many cases the additional information captured by the AWFT is very beneficial. We demonstrate the utility of the proposed method on three common shape analysis tasks: segmentation, salient point detection, and the construction of intrinsic feature descriptors. Our results in these applications compare favorably to the state of the art on standard benchmarks.

## 2. Background.

**Manifolds.** We model a 3D shape as a connected smooth compact two-dimensional manifold (surface) $X$ embedded into $\mathbb{R}^3$. Locally around each point $x \in X$, the manifold is homeomorphic to the *tangent plane* $T_x X$. Every element in $T_x X$ is called *tangent vector* at $x$. The *tangent bundle* $T X$ can be seen as the disjoint union of all the tangent planes of $X$. A *(tangent) vector field* on the manifold is a smooth function $\mathbf{v} : X \to TX$. A *Riemannian metric* is an inner product $\langle \cdot, \cdot \rangle_{T_x X} : T_x X \times T_x X \to \mathbb{R}$ on the tangent plane depending smoothly on $x$. Given a Riemannian metric, it is possible to define the *second fundamental form*, a $2 \times 2$ ma-

trix, which describes how the manifold locally differs from a plane. The eigenvalues $\kappa_M, \kappa_m$ of the second fundamental form are called the *principal curvatures* and their corresponding eigenvectors $\mathbf{v}_M, \mathbf{v}_m$ are called the *principal curvature directions*. The pair $\mathbf{v}_M(x), \mathbf{v}_m(x)$ constitutes an orthonormal basis on the tangent plane $T_x X$.

**Laplacian.** We denote the space of square-integrable real functions (scalar fields) on the manifold by $L^2(X) = \{f : X \to \mathbb{R} : \int_X f(x)^2 dx < \infty\}$, where $dx$ is the area element induced by the Riemannian metric. The *intrinsic gradient* $\nabla_X f$ of a smooth scalar field $f \in L^2(X)$ is defined as

$$\langle \nabla_X f(x), \mathbf{v} \rangle_{T_x X} = \partial_{\mathbf{v}} f(x) \,,$$

where $\mathbf{v} \in T_x X$ is a tangent vector, and $\partial_{\mathbf{v}} f$ denotes the directional (covariant) derivative. Given a smooth vector field $\mathbf{v}$, the *intrinsic divergence* can be defined as its negative formal adjoint

$$\langle \nabla_X f(x), \mathbf{v}(x) \rangle_{T_x X} = -\langle f(x), \mathrm{div}_X \mathbf{v}(x) \rangle_X. \quad (1)$$

The *Laplace-Beltrami operator* (or *Laplacian*) is defined as

$$\Delta_X f(x) = -\mathrm{div}_X(\nabla_X f(x)). \quad (2)$$

**Fourier analysis on manifolds.** The Laplacian admits an eigendecomposition

$$\Delta_X \phi_i(x) = \lambda_i \phi_i(x)$$

with orthonormal eigenfunctions $\phi_0, \phi_1, \ldots$ (generalizing the standard Fourier basis) and corresponding non-negative eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \ldots$ (which can be interpreted as frequencies). The first eigenvector is constant $\phi_0(x) = \frac{1}{(\mathrm{area}(X))^{1/2}}$, and plays the role of DC in classical signal processing.

A function $f \in L^2(X)$ can be expressed as a *Fourier series*

$$f(x) = \sum_{i \geq 0} \langle f, \phi_i \rangle_{L^2(X)} \phi_i(x) \,,$$

where $\hat{f}_i = \langle f, \phi_i \rangle_{L^2(X)}$ are the *Fourier coefficients*. The *generalized convolution* of $f$ and $g$ on the manifold can be defined by the analogy to the classical case as

$$
\begin{aligned}
(f \star g)(x) &= \sum_{i \geq 0} \langle f, \phi_i \rangle_{L^2(X)} \langle g, \phi_i \rangle_{L^2(X)} \phi_i(x) \quad (3) \\
&= \sum_{i \geq 0} \hat{f}_i \hat{g}_i \phi_i(x).
\end{aligned}
$$

In the Euclidean case, this relation is known as the convolution theorem. In the non-Euclidean case, Eq. (3) can be regarded as a definition of a non-shift-invariant convolution.

**Anisotropic Laplacian.** Andreux et al. [1] considered an *anisotropic* Laplace-Beltrami operator of the form

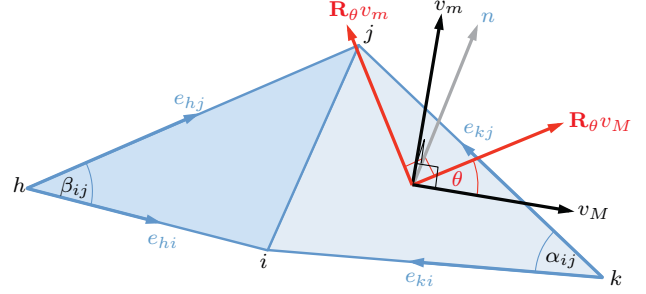$$\Delta_X f(x) = -\mathrm{div}_X(\mathbf{D}(x) \nabla_X f(x)) \,, \quad (4)$$



Figure 1. Triangular mesh discretization. The orthogonal basis vectors $v_M, v_m$, as well as their rotated counterparts (in red), lie in the plane of the respective triangle (reproduced from [5]).

where $\mathbf{D}(x)$ acts on the intrinsic gradient direction in the tangent space, represented in the orthogonal basis $\mathbf{v}_M(x), \mathbf{v}_m(x)$ of principal curvature directions. In particular, the authors considered anisotropy along the maximum curvature direction,

$$\mathbf{D}_\alpha(x) = \begin{bmatrix} \frac{1}{1+\alpha} & \\ & 1 \end{bmatrix}, \quad (5)$$

where parameter $\alpha > 0$ controls the level of anisotropy. Boscaini et al. [5] considered anisotropic Laplacians with anisotropy at angle $\theta$ w.r.t. the maximum curvature direction,

$$\mathbf{D}_{\alpha\theta}(x) = \mathbf{R}_\theta \mathbf{D}_\alpha(x) \mathbf{R}_\theta^\top, \quad (6)$$

where $\mathbf{R}_\theta$ is a rotation by $\theta$ in the tangent plane. The resulting Laplacian operator

$$\Delta_{\alpha\theta} f(x) = -\mathrm{div}_X(\mathbf{R}_\theta \mathbf{D}_\alpha(x) \mathbf{R}_\theta^\top \nabla_X f(x)) \quad (7)$$

is the centerpiece of the construction proposed in this paper.

## 3. Anisotropic Windowed Fourier Transform

The main drawback of standard Fourier analysis is that the basis functions are globally supported. As a result, it is practically impossible to localize a small spatial feature in the frequency domain. A common technique in signal processing, referred to as the Windowed Fourier Transform (WFT, also known as Short-Time Fourier Transform or spectrogram) is to *localize* frequency analysis to a window, considering the Fourier coefficients of a function for each window location. The result is a combined space-frequency representation.

**WFT on manifolds.** Shumann et al. [27] and Boscaini et al. [4] generalized this construction to graphs and manifolds, respectively. The window is defined in the frequency domain and its translation to point $\xi$ is expressed as a gen-

eralized convolution with a delta-function,

$$
\begin{aligned}
T_\xi g &= (g \star \delta_\xi)(x) = \sum_{i \geq 0} \hat{g}_i \langle \delta_\xi, \phi_i \rangle_{L^2(X)} \phi_i(x) \quad (8) \\
&= \sum_{i \geq 0} \hat{g}_i \phi_i(\xi) \phi_i(x) \, .
\end{aligned}
$$

The WFT is computed as

$$
\begin{aligned}
(Sf)_{\xi i} &= \langle f, \underbrace{\rho \phi_i \sum_{j \geq 0} \hat{g}_j \phi_j(\xi) \phi_j}_{g_{\xi i}(x)} \rangle_{L^2(X)} \quad (9) \\
&= \rho \sum_{j \geq 0} \hat{g}_j \phi_j(\xi) \langle f, \phi_i \phi_j \rangle_{L^2(X)} \, ,
\end{aligned}
$$

where $g_{\xi i}(x)$ is the window at position $\xi$ modulated with the $i$th frequency, referred to as an *atom*. We can think of it as a translated and modulated version of $g$,

$$
g_{\xi i}(x) = \rho M_i T_\xi g(x) = \rho \phi_i(x)(g \star \delta_\xi)(x) \, .
$$

$\rho = (\text{area}(X))^{1/2}$ is a normalization constant ensuring that the modulation with DC is norm-preserving,

$$
\rho M_i f(x) = \rho \phi_0(x) f(x) = \rho \frac{1}{\rho} f(x) = f(x).
$$

Note that $(Sf)_{\xi i}$ has two indices: spatial location $\xi$ and frequency $i$.

**AWFT.** While allowing for a localized frequency analysis, the WFT atoms are agnostic to directional information (see Figure 2, left). Such information can be introduced using the anisotropic Laplacian. Let

$$
\Delta_{\alpha\theta} \phi_{\alpha\theta,i}(x) = \lambda_{\alpha\theta,i} \phi_{\alpha\theta,i}(x)
$$

be the eigendecomposition of the anisotropic Laplacian with orthogonal eigenfunctions $\phi_{\alpha\theta,0}, \phi_{\alpha\theta,1}, \ldots$ and corresponding non-negative eigenvalues $0 = \lambda_{\alpha\theta,0} \leq \lambda_{\alpha\theta,1} \leq \ldots$. We define the *anisotropic WFT* as

$$
(Sf)_{\xi\alpha\theta i} = \sum_{j \geq 0} \hat{g}_{\alpha\theta,j} \phi_{\alpha\theta,j}(\xi) \langle f, \phi_{\alpha\theta,i} \phi_{\alpha\theta,j} \rangle_{L^2(X)}. \quad (10)
$$

The atoms $g_{\xi\theta i}(x)$ (Figure 2, right) are direction-aware. One of the important consequences of direction-awareness is the fact that AWFT is capable of disambiguating intrinsic reflection symmetries, as will be discussed in the following.

**Choice of the window.** The choice of the window allows for a tradeoff between spatial and frequency localization (by virtue of the uncertainty principle, it is impossible to achieve a perfect localization in both): a narrow window in the frequency domain (rapidly decaying Fourier coefficients

$\hat{g}_i$) results in a wide window in the spatial domain, and vice versa. We use a decaying window

$$
\hat{g}_{\alpha\theta,i} = e^{-\tau(\lambda_{\alpha\theta,i} - \lambda_{\alpha\theta,1})} \, , \quad (11)
$$

where the parameter $\tau$ controls the decay rate (larger values of $\tau$ produce windows with poorer spatial localization). In order to make this parameter scale-invariant, we recall that scaling the coordinates of the shape uniformly by a factor $\beta$ scales its area by $\beta^2$ and Laplacian eigenvalues by a factor of $\beta^{-2}$. We therefore multiply $\tau$ by $\text{area}(X)$.

**Total Weighted Power** Rabiei et al. [21] used the *total weighted power* (TWP) as an aggregate of all frequency information weighted by the normalized corresponding eigenvalue,

$$
(S_{\text{TWP}}f)_{\xi\alpha\theta} = \sum_{k \geq 0} \frac{\lambda_{\alpha\theta,k}^2}{\|\mathbf{\Lambda}_{\alpha\theta}\|_2^2} (Sf)_{\xi\alpha\theta k}^2, \quad (12)
$$

where $\|\mathbf{\Lambda}\|_2^2 = \sum_{k \geq 0} \lambda_{\alpha\theta,k}^2$ is the norm of the set of eigenvalues. Such an aggregation allows removing the potential ambiguities due to different signs and ordering of the eigenfunctions. Attributing greater value to high frequencies the *TWP* can be seen as a filter that emphasizes localized (high-frequency) properties.

# 4. Discretization

In the discrete setting, the surface $X$ is sampled at $N$ points $x_1, \ldots, x_N$, on which we construct a triangular mesh $T = (\{1, \ldots, N\}, E, F)$. A function on the surface is represented by an $N$-dimensional vector $\mathbf{f} = (f(x_1), \ldots, f(x_N))^\top$. The inner product is discretized as $\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}^\top \mathbf{A} \mathbf{g}$, where $\mathbf{A} = \text{diag}(a_1, \ldots, a_N)$ is the *mass matrix*, and $a_i$ denotes the local area element at vertex $i$. To each triangle $ijk$ of the mesh, we attach an orthonormal reference frame $\mathbf{U}_{ijk} = [\mathbf{v}_M, \mathbf{v}_m, \hat{\mathbf{n}}] \in \mathbb{R}^{3 \times 3}$, where $\hat{\mathbf{n}}$ is the unit normal vector to the triangle and $\mathbf{v}_m$ and $\mathbf{v}_M$ are the directions of principal curvature, computed using the method of [8]. The tensor $\mathbf{D}_\alpha$ for the triangle $ijk$ operating on tangent vectors is expressed w.r.t. $\mathbf{U}_{ijk}$ as:

$$
\mathbf{D}_\alpha = \begin{bmatrix} \frac{1}{1+\alpha} & & \\ & 1 & \\ & & 1 \end{bmatrix} .
$$

In the case $\theta = 0$, let $\mathbf{e}_{ij} \in \mathbb{R}^3$ denote the oriented edge pointing from vertex $i$ to vertex $j$, normalized to unit length, and consider the triangle $ijk$. We define the $\mathbf{H}$-weighted inner product between edges $\mathbf{e}_{kj}$ and $\mathbf{e}_{ki}$ as

$$
\langle \mathbf{e}_{kj}, \mathbf{e}_{ki} \rangle_{\mathbf{H}} = \mathbf{e}_{kj}^T \underbrace{\mathbf{U}_{ijk} \mathbf{D}_\alpha \mathbf{U}_{ijk}^T}_{\mathbf{H}} \mathbf{e}_{ki} \, , \quad (13)
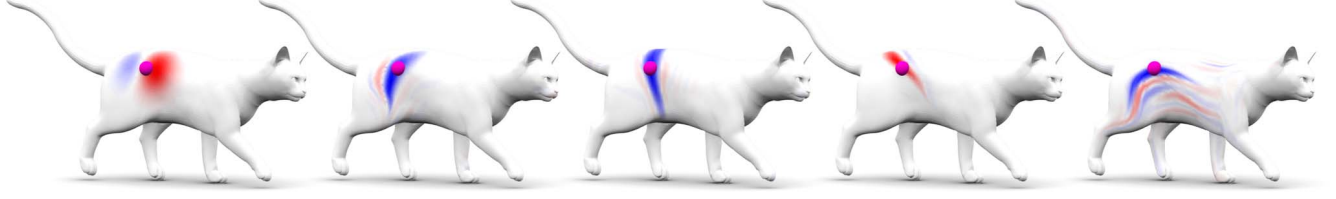$$

Figure 2. Some atoms of the AWFT on the point marked with small pink spheres on the cat shape. These are obtained with $\tau = 0.002$, the first on the left is the isotropic window while the others are obtained with $\alpha = 300$ and different $\theta$, from left to right 45, 90, 135, 180.

where the *shear matrix* $\mathbf{H}$ encodes the anisotropic scaling up to an orthogonal basis change. Note that in the isotropic case ($\alpha = 0$) we have $\mathbf{H} = \mathbf{I}$, such that the $\mathbf{H}$-weighted inner product simplifies to the standard inner product $\langle e_{kj}, e_{ki} \rangle_{\mathbf{H}} = \cos \alpha_{ij}$.

The discretization of the anisotropic Laplacian takes the form of an $n \times n$ sparse matrix $\boldsymbol{\Delta}_{\alpha\theta} = -\mathbf{A}^{-1}\mathbf{W}$. The *stiffness matrix* $\mathbf{W}$ is composed of weights

$$w_{ij} = \begin{cases} -\frac{1}{2}\left( \frac{\langle \mathbf{e}_{kj}, \mathbf{e}_{ki}\rangle_{\mathbf{H}}}{\sin \alpha_{ij}} + \frac{\langle \mathbf{e}_{hj}, \mathbf{e}_{hi}\rangle_{\mathbf{H}}}{\sin \beta_{ij}} \right) & (i,j) \in E \\ -\sum_{k \neq i} w_{ik} & i = j \\ 0 & \text{else} \end{cases} \quad (14)$$

where the notation is according to Figure 1. In the isotropic case, $\frac{\langle \mathbf{e}_{kj}, \mathbf{e}_{ki}\rangle_{\mathbf{H}}}{\sin \alpha_{ij}} = \frac{\cos \alpha_{ij}}{\sin \alpha_{ij}} = \cot \alpha_{ij}$, thus reducing equation (14) to the classical cotangent formula [20].

To obtain the general case $\theta \neq 0$, it is sufficient to rotate the basis vectors $\mathbf{U}_{ijk}$ on each triangle around the respective normal $n$ by the angle $\theta$, equal for all triangles (see Figure 1, red). Denoting by $\mathbf{R}_\theta$ the corresponding $3 \times 3$ rotation matrix, this is equivalent to modifying the $\mathbf{H}$-weighted inner product with the directed shear matrix $\mathbf{H}_\theta = \mathbf{R}_\theta \mathbf{H} \mathbf{R}_\theta^T$. The resulting weights $w_{ij}$ in equation (14) are thus obtained by using the inner products $\langle \mathbf{e}_{kj}, \mathbf{e}_{ki} \rangle_{\mathbf{H}_\theta} = \mathbf{e}_{kj}^T \mathbf{H}_\theta \mathbf{e}_{ki}$.

The computation of the Laplacian eigenvectors is posed as a generalized eigenproblem

$$\mathbf{W}\boldsymbol{\Phi} = \mathbf{A}\boldsymbol{\Phi}\boldsymbol{\Lambda},$$

where $\boldsymbol{\Phi} = (\phi_1, \ldots, \phi_K)$ is an $N \times K$ matrix containing the first $K$ eigenvectors, and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_K)$ is a diagonal matrix containing the corresponding eigenvalues. For the anisotropic Laplacian, we compute a set of $K$ eigenvectors $\boldsymbol{\Phi}_{\alpha\theta}$ and eigenvalues $\boldsymbol{\Lambda}_{\alpha\theta}$ for each $\theta$ and $\alpha$. The AWFT is computed as

$$(Sf)_{\alpha\theta} = (\mathbf{f} \boxdot \boldsymbol{\Phi}_{\alpha\theta})^\top \mathbf{A} \boldsymbol{\Phi}_{\alpha\theta}(\hat{\mathbf{g}}_{\alpha\theta} \boxdot \boldsymbol{\Phi}_{\alpha\theta}^\top),$$

where the result is an $N \times K$-dimensional matrix for $\theta$ and $\alpha$; $\mathbf{f}$ is the input function represented as an $N$-dimensional vector, and $(\mathbf{a} \boxdot \mathbf{B})_{ij} = a_i b_{ij}$ denotes a $K \times N$ matrix obtained by element-wise multiplication of a $K$-dimensional column vector replicated $N$ times along the second dimension with a $K \times N$ matrix.

## 5. Applications and Results

In this section, we show the application of AWFT to three standard problems in geometry processing: *shape segmentation*, *salient point detection* and design of *local feature descriptors* for point to point matching. In each application, one has the freedom to define the four main ingredients of AWFT. First, the function $f$ to be analyzed, which represents the information we wish to encode. Second, the size of the window, determined by the parameter $\tau$. We typically use a set of different values $\tau_1, \ldots, \tau_t$ to perform a multi-scale analysis. Third, the window orientations $\theta_1 = 0, \theta_2 = \frac{\pi}{h}, \ldots, \theta_h = \frac{(h-1)\pi}{h}$. Fourth, the anisotropy parameter $\alpha$. We typically use a set of different values $\alpha_1, \ldots, \alpha_c$ to capture oriented structures of different width (see inset figure).
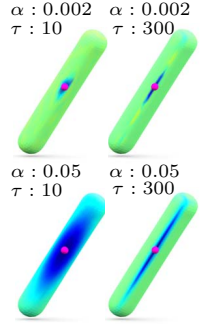


**Computation:** We used up to $K = 200$ Laplacian eigenvectors and eigenvalues computed using MATLAB *eigs* function. The computation of AWFT with the settings used in our experiments takes on average less than 5 seconds on a mesh with around 7000 vertices on a machine with 32GB of RAM and an Intel 3,6 GHz Core i7 cpu[1].

### 5.1. Shape Segmentation

Applying the AWFT to the constant function $f(x) = 1$ gives rise to a reweighted version of the GPS descriptor [26, 4] that is direction-aware. We apply on the squared AWFT features $(Sf)_{\xi\alpha\theta i}^2$ the segmentation method of Rodolà et al. [24] (which takes the classical GPS as input) using the settings proposed by the authors.

**Settings.** We use $f = 1$, single window with $\tau = 0.002$, two levels of anisotropy $\alpha_1 = 100, \alpha_2 = 300$, $\mathcal{F} = f$ the constant function, and orientations $\theta_1 = 0, \theta_2 = \frac{\pi}{4}, \ldots, \theta_4 = \frac{3\pi}{4}$. We use only the first $K = 50$ frequencies.

**Data and Evaluation.** We evaluate the segmentation results according to the Princeton Segmentation Benchmark

---

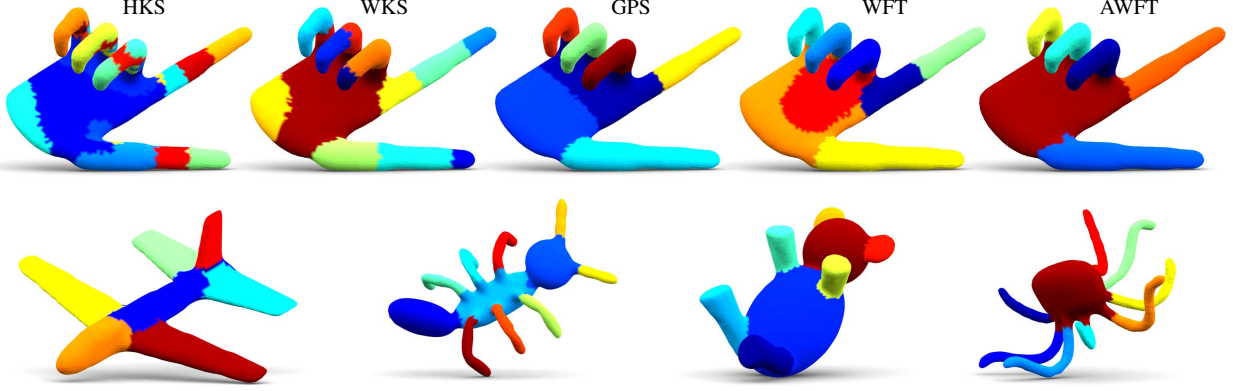[1]Code is available at: http://profs.sci.univr.it/~castella/awft.xhtml.

Figure 3. First row: segmentation of the hand shape obtained using (left to right) HKS, WKS, GPS, WFT and AWFT. Second row: a few segmentation examples obtained with AWFT.

|          | HKS          | WKS          | GPS          | WFT          | AWFT             |
|----------|--------------|--------------|--------------|--------------|------------------|
| airplane | 0.41 / 0.82  | 0.38 / 0.84  | 0.47 / 0.80  | 0.42 / 0.85  | **0.48 / 0.91**  |
| ant      | 0.57 / **0.97** | 0.61 / 0.90 | **0.64** / 0.84 | 0.63 / 0.90 | 0.62 / 0.90      |
| bird     | 0.44 / 0.84  | 0.45 / **0.89** | **0.52** / 0.86 | 0.44 / 0.86 | **0.52 / 0.89**  |
| fish     | **0.62 / 0.95** | 0.43 / 0.92 | 0.43 / 0.86  | 0.37 / 0.86  | 0.39 / 0.91      |
| hand     | 0.52 / 0.91  | 0.55 / 0.92  | 0.61 / 0.88  | 0.61 / 0.88  | **0.75 / 0.94**  |
| octopus  | 0.35 / **0.98** | 0.35 / 0.95 | **0.44** / 0.87 | 0.32 / 0.90 | 0.40 / 0.95      |
| plier    | 0.35 / 0.90  | 0.36 / **0.92** | 0.43 / 0.83 | 0.39 / 0.88  | **0.51** / 0.91  |
| teddy    | 0.48 / 0.90  | 0.51 / 0.88  | 0.63 / 0.75  | 0.53 / 0.83  | **0.68 / 0.93**  |
| mean     | 0.47 / 0.91  | 0.46 / 0.90  | 0.52 / 0.84  | 0.46 / 0.87  | **0.54 / 0.92**  |

Table 1. Performance in *Fidelity* / *Goodness* on 8 categories of the Princeton Segmentation Benchmark. Best performance is in bold.

[7], consisting of 380 meshes from 19 object classes (20 shapes per class). A common meaningful segmentation is given as ground truth for each category. The ground truth assigns the same label to semantically similar segments (*e.g.*, the two arms of a man).

We use two different evaluation criteria, *Fidelity* and *Goodness*. For every segment in the ground-truth, Fidelity is the average ratio between the number of the maximum subset of points that is segmented together by the method and the number of points that compose that ground-truth segment. Goodness is in some sense the inverse. For every segment obtained by the method, it is defined as the average ratio between the maximum number of points that are segmented together in the ground truth and the number of points that compose that segment provided by the method.

**Results.** Table 1 summarizes the segmentation results in terms of Goodness and Fidelity. For comparison, we show segmentation results obtained with HKS, WKS, GPS and isotropic WFT features. Our AWFT produces the best score on average. Segmentation examples in Figure 3 show that AWFT is able to perform a segmentation that is both semantically meaningful and geometrically consistent.

## 5.2. Salient Point Detection

We use the AWFT to construct a saliency map that allows to detect key points on surfaces. We use the logarithm of the mean curvature as an input function (capturing large curvature variations) and compute the TWP of the AWFT. This way, we obtain a set of saliency maps for each value of $\alpha$, $\theta$ and $\tau$. For every such saliency map, we compute local maxima as follows: For each point $\xi$ we take the corresponding window $T_\xi g_{\alpha\theta}$ (computed using the same $\tau$ used for the map) and locate its maximum. We then perform a non-maximum suppression. Each map is then normalized as proposed in [12]. A single saliency map is obtained by summing up the maps for different $\alpha$, $\theta$ and $\tau$. Finally, the maximum detection and non-maximum suppression is performed again, producing the salient points.

**Settings.** We use $f = \log(H)$, various window sizes $\tau = 0.0002, 0.0007, 0.001, 0.0015, 0.0055$, a single level of anisotropy $\alpha = 300$, and angles $\theta_1 = 0, \theta_2 = \frac{\pi}{12}, \ldots, \theta_{12} = \frac{11\pi}{12}$.

**Data and Evaluation.** We follow the test proposed in [10] on two different datasets. Dataset A consists of 24 objects hand-marked by 23 human subjects. Dataset B contains 43 models marked by at least 16 subjects. The human annotations are used as the ground-truth for both datasets. The adopted evaluation criterion (WME) is based on importance of the selected points, where importance is based on these hand-marked selections.

**Results.** Figure 4 shows a few saliency maps and the corresponding selected salient points. Note how the saliency maps identify semantically coherent parts among the animal and cup shape classes. Figure 5 evaluates the performance of various feature detection methods. We use two non-maximum suppression settings (denoted AWFT and AWFT2) to tradeoff between false positives and negatives. Table 2 shows the average area-under-the-curve (AUC) for different methods. Our methods achieve performance in
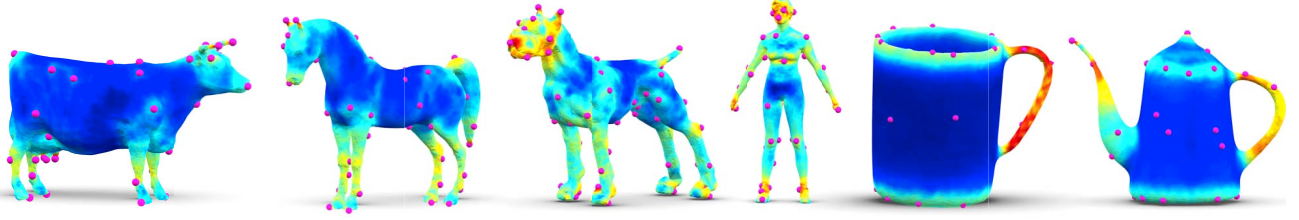
Figure 4. Saliency maps computed with AWFT on some shapes from the evaluation benchmark. The areas considered as more salient are in red while the less important ones tend to blue. The final selected points are highlighted by small pink spheres.
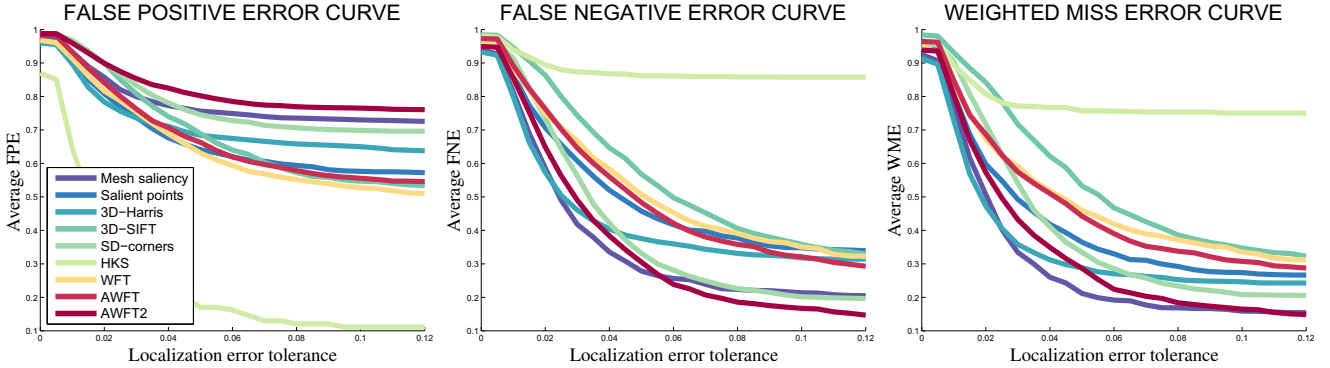


Figure 5. The False Negative, False Positive and Weighted Miss error curves for the proposed methods (in warm colors) and the methods presented in the used benchmark (in cold colors).

| Mean | Mesh saliency | Salient points | 3D-Harris | 3D-SIFT | SD-corners | WFT | AWFT | AWFT2 |
|------|------|------|------|------|------|------|------|------|
| **0.59** | 0.57 | 0.59 | 0.57 | 0.63 | 0.59 | 0.59 | 0.59 | 0.59 |

Table 2. The mean AUC computed on the False Negative and the False Positive error curves for all the comparable methods tested in these curves (the proposed ones and all the others presented in the benchmark, not HKS that is not compatible with other methods).

line with the state of the art.

## 5.3. Descriptors for Point to Point Matching

In order to obtain a concise and informative descriptor for every point on the surface, we use the TWP to reduce the dimension of the output of AWFT on some geometrically meaningful functions. One of the key deficencies of standard descriptors such as HKS, WKS or WFT is their invariance to intrinsic symmetries. This is visualized in Figure 6, where we show a bilaterally intrinsically symmetric shape (cat). The WFT at symmetric points (leftmost plot in solid and dotted) are nearly identical. On the other hand, orientation is not preserved by the intrinsic symmetry (in fact, if $\sigma : X \to X$ is a bilateral symmetry, then $(Sf)_{\xi\alpha\theta i} = (Sf)_{\sigma(\xi)\alpha, -\theta, i}$). As a result, the AWFT descriptor can distinguish intrinsically symmetric points, as we experimentally show in the following.

**Settings.** We used five input features: $f_1, f_2$ are the third and fourth *geometry vectors* [17] encoding some of the spectral geometry of the shape; $f_3 = 1$; $f_4 = \phi_1$ is the (isotropic) Fiedler vector providing a consistent ordering on the mesh vertices [15]; and $f_5$ is the ShapeIndex [14] encoding the curvatures of the surface. We used $\tau = 0.002, 0.05$, $\alpha_1 = 100, \alpha_2 = 300$, and $\theta_1 = 0, \theta_2 = \frac{\pi}{4}, \ldots, \theta_4 = \frac{3\pi}{4}$.

**Data and Evaluation.** We used two public-domain datasets of scanned human shapes in different poses: FAUST [3] and CAESAR [23]. These datasets are considerably challenging due to the presence of non-isometric deformations as well as significant variability between different human subjects. FAUST is composed of 10 poses of 10 subjects (100 shapes in total), each having nearly 7K vertices. We select a random set of 60 shapes from the *fitted-meshes* subset of CAESAR, where each shape has ∼6K vertices. Ground truth point-wise correspondence is available in both datasets. We refer the reader to the supplementary material for tests on larger meshes (∼28K vertices).

We evaluate descriptor performance using *cumulative match characteristic* (CMC), *receiver operator characteristic* (ROC) and *correspondence quality characteristic* (CQC). The CMC estimates the probability of a correct correspondence to be among the $k$ nearest neighbors in descriptor space. ROC measures the percentage of positives and negatives pairs falling below various thresholds of their distance in descriptor space (*true positive* and *negative rates*, respectively). CQC exhibits the percentage of nearest-neighbor matches that are at most $r$-geodesically
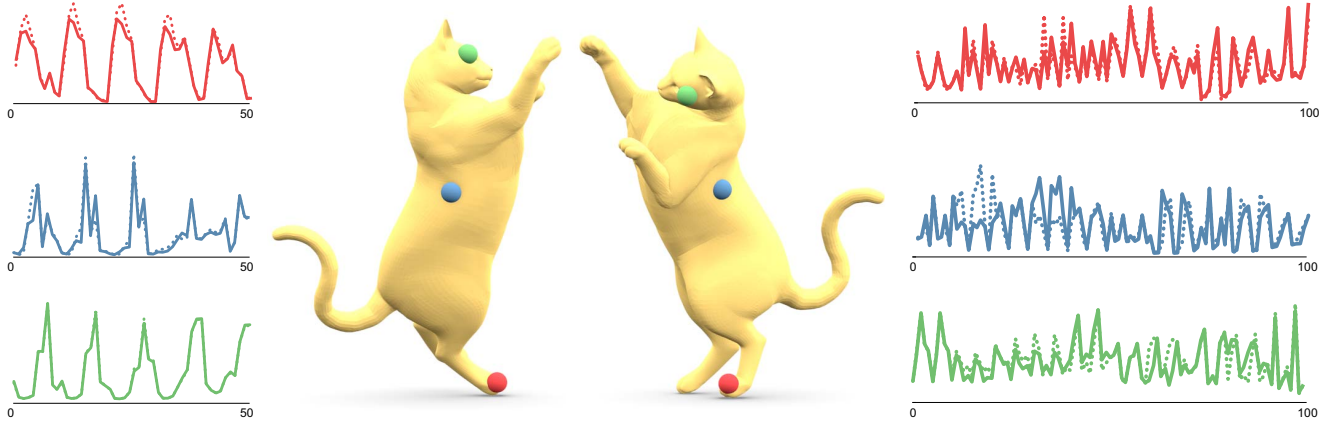
Figure 6. Comparison between WFT (left) and AWFT (right) descriptor computed at three symmetric points of the cat shape (points and corresponding curves are color-coded). Solid and dotted curves represent descriptors from the left and right side of the cat, respectively.
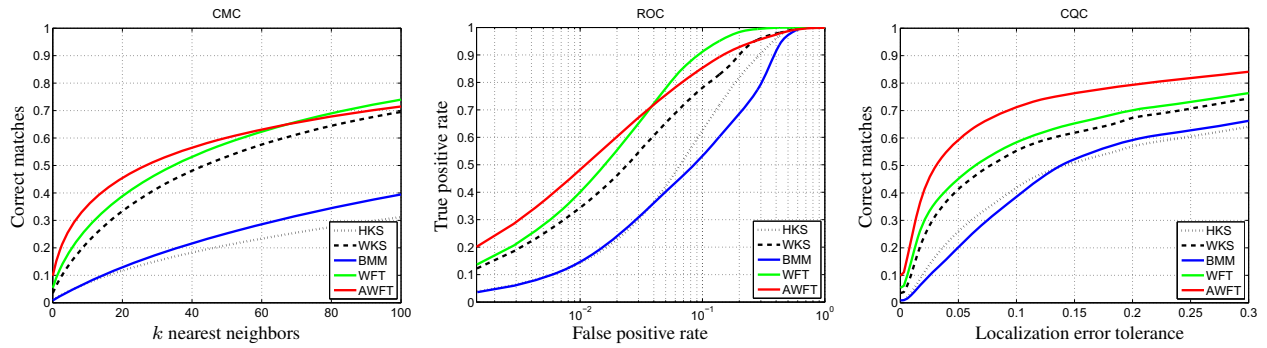


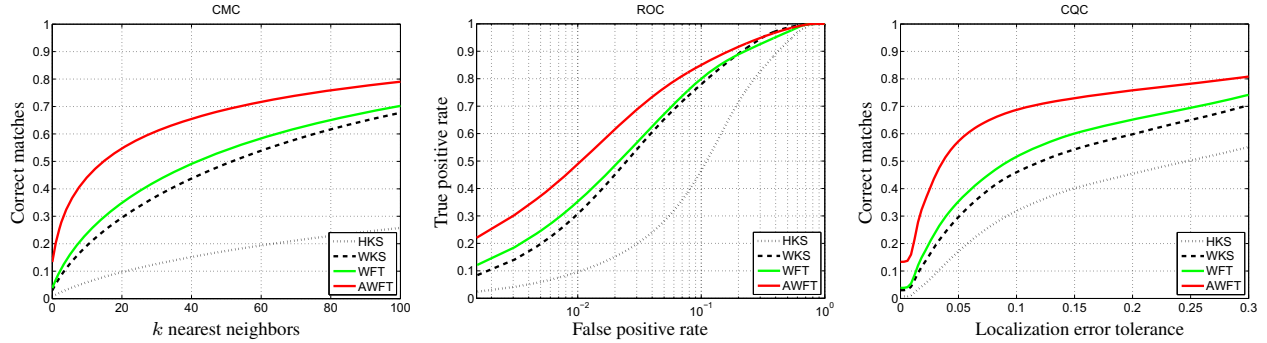Figure 7. Performance evaluation on FAUST dataset.



Figure 8. Performance evaluation on CAESAR dataset.

distant from the ground truth correspondence.

**Results.** The evaluation results are summarized in Figures 7–8, which clearly demonstrate that AWFT outperforms the classical methods. In Figure 7, we also compare to the descriptor proposed in [4] without the learning step (referred to as BMM). We use the first 20 coefficients of the WFT computed on 5 geometry vectors as in [4], with the same two $\tau$ used in AWFT. Despite the double dimensionality BMM does not offer comparable performance without the class-specific learning step. In Figure 9, we show the distance between the descriptor at a point (indicated with

the white sphere on the first shape on the left) and the rest of the points on the same shape as well as other shapes. AWFT descriptors are the most localized and discriminative, and correctly disambiguate symmetries.

## 6. Conclusions

In this paper, we proposed an anisotropic windowed Fourier transform on manifolds. The proposed tool enables to perform local directional frequency analysis and improve the study of geometry in several applications in shape analysis. Its power and versatility were tested on
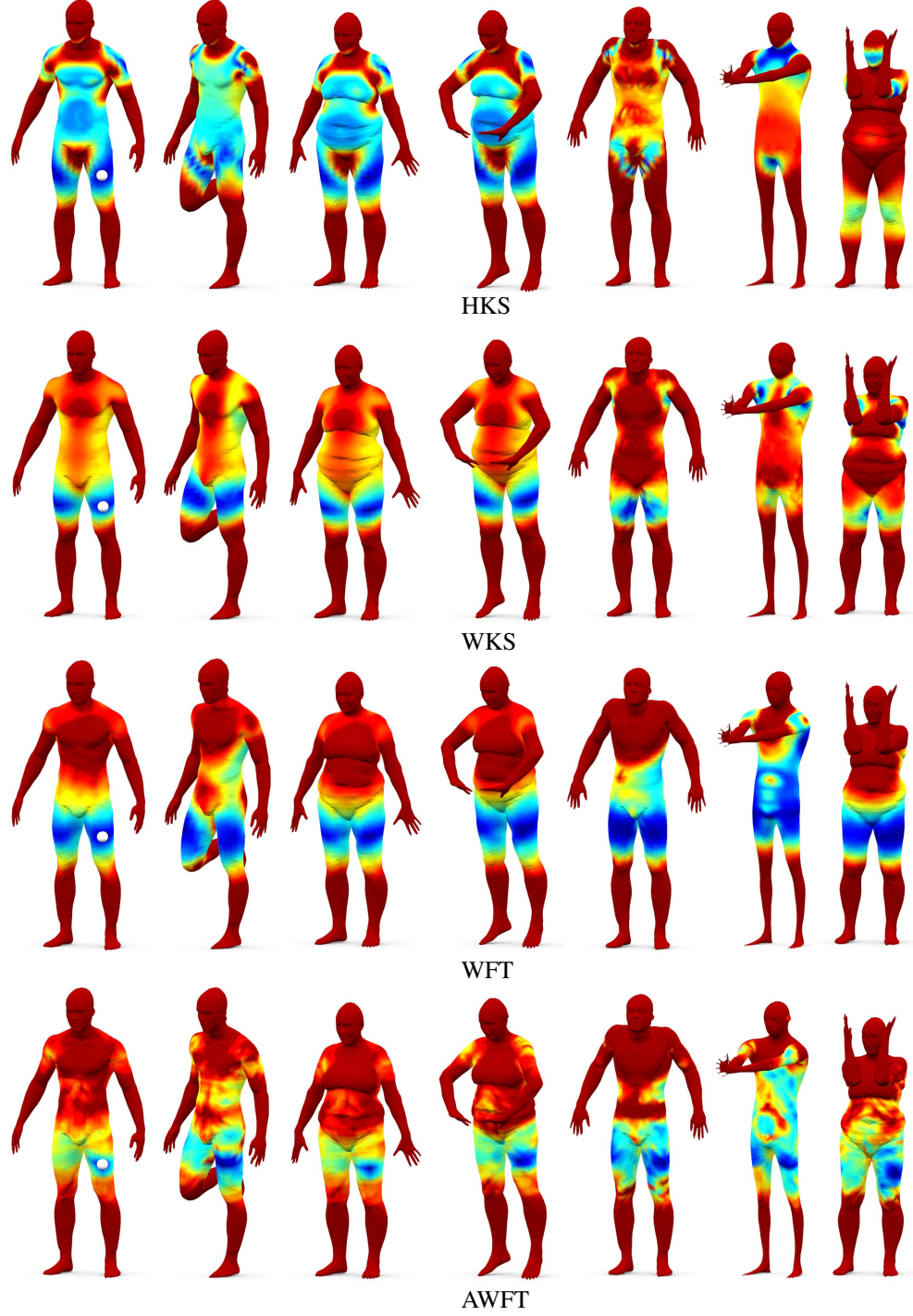
Figure 9. Normalized Euclidean distance between the descriptor at a reference point on the leg (white sphere) and the descriptors computed at the rest of the points for different transformations (from left to right: near isometric deformations, non-isometric deformations, subsampling, smoothing and remeshing). Cold and hot colors represent small and large distances, respectively. For visualization clarity, distances are saturated at $30\%$ of the maximum.

classical datasets for different tasks as shape segmentation, salient points detection and point to point matching. In future work, we will study additional applications of AWFT, as well as its computation on other types of data such as point clouds.

# References

[1] M. Andreux, E. Rodolà, M. Aubry, and D. Cremers. Anisotropic laplace-beltrami operators for shape analysis. In *Proc. NORDIA*, pages 299–312. Springer, 2014. 2

[2] M. Aubry, U. Schlickewei, and D. Cremers. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Proc. ICCV Workshops*, 2011. 1

[3] F. Bogo, J. Romero, M. Loper, and M. J. Black. FAUST: Dataset and evaluation for 3D mesh registration. In *Proc. CVPR*, 2014. 6

[4] D. Boscaini, J. Masci, S. Melzi, M. M. Bronstein, U. Castellani, and P. Vandergheynst. Learning class-specific descriptors for deformable shapes using localized spectral convolutional networks. In *Computer Graphics Forum*, volume 34, pages 13–23. Wiley, 2015. 1, 2, 4, 7

[5] D. Boscaini, J. Masci, E. Rodolà, M. M. Bronstein, and D. Cremers. Anisotropic diffusion descriptors. *Computer Graphics Forum*, 35(2), 2016. 1, 2

[6] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics (TOG)*, 30(1):1, 2011. 1

[7] X. Chen, A. Golovinskiy, and T. Funkhouser. A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3), Aug. 2009. 5

[8] D. Cohen-Steiner and J.-M. Morvan. Restricted delaunay triangulations and normal cycle. In *Proceedings of symposium on Computational geometry*, pages 312–321. ACM, 2003. 3

[9] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006. 1

[10] H. Dutagaci, C. Cheung, and A. Godil. Evaluation of 3d interest point detection techniques via human-generated ground truth. *The Visual Computer*, 28:901–917, 2012. 5

[11] K. Gebal, J. A. Bærentzen, H. Anæs, and R. Larsen. Shape analysis using the auto diffusion function. *CGF*, 28(5):1405–1413, 2009. 1

[12] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11):1254–1259, 1998. 5

[13] Z. Karni and C. Gotsman. Spectral compression of mesh geometry. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 279–286. ACM Press/Addison-Wesley Publishing Co., 2000. 1

[14] J. Koenderink and A. van Doorn. Surface shape and curvature scales. *Image and Visual Computing*, 10:557–565, 1992. 6

[15] B. Lévy and H. R. Zhang. Spectral mesh processing. In *ACM SIGGRAPH 2010 Courses*, page 8. ACM, 2010. 1, 6

[16] R. Litman, A. Bronstein, M. Bronstein, and U. Castellani. Supervised learning of bag-of-features shape descriptors using sparse coding. *CGF*, 33(5):127–136, 2014. 1

[17] R. Litman and A. M. Bronstein. Learning spectral descriptors for deformable shape correspondence. *TPAMI*, 36(1):170–180, 2014. 1, 6

[18] S. Mallat. *A wavelet tour of signal processing*. Academic press, 1999. 1

[19] P. Mullen, Y. Tong, P. Alliez, and M. Desbrun. Spectral conformal parameterization. In *Computer Graphics Forum*, volume 27, pages 1487–1494. Wiley Online Library, 2008. 1

[20] U. Pinkall and K. Polthier. Computing discrete minimal surfaces and their conjugates. *Experimental Mathematics*, 2(1):15–36, 1993. 4

[21] H. Rabiei, F. Richard, M. Roth, J.-L. Anton, O. Coulon, and J. Lefèvre. The graph windowed fourier transform: a tool to quantify the gyrification of the cerebral cortex. In *Workshop on Spectral Analysis in Medical Imaging (SAMI)*, 2015. 3

[22] M. Reuter, F.-E. Wolter, and N. Peinecke. Laplace–beltrami spectra as shape-dnaof surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006. 1

[23] K. M. Robinette, H. Daanen, and E. Paquet. The caesar project: a 3-d surface anthropometry survey. In *Proc. 3DIM*, pages 380–386, 1999. 6

[24] E. Rodolà, S. Rota Bulò, and D. Cremers. Robust region detection via consensus segmentation of deformable shapes. In *Computer Graphics Forum*, volume 33, pages 97–106. Wiley, 2014. 4

[25] E. Rodolà, S. Rota Bulò, T. Windheuser, M. Vestner, and D. Cremers. Dense non-rigid shape correspondence using random forests. In *Proc. CVPR*, 2014. 1

[26] R. M. Rustamov. Laplace-Beltrami eigenfunctions for deformation invariant shape representation. In *Proc. SGP*, 2007. 1, 4

[27] D. I. Shuman, B. Ricaud, and P. Vandergheynst. Vertex-frequency analysis on graphs. *arXiv:1307.5708*, 2013. 1, 2

[28] J. Sun, M. Ovsjanikov, and L. J. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *CGF*, 28(5):1383–1392, 2009. 1

[29] G. Taubin. A signal processing approach to fair surface design. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 351–358. ACM, 1995. 1

[30] T. Windheuser, M. Vestner, E. Rodolà, R. Triebel, and D. Cremers. Optimal intrinsic descriptors for non-rigid shape analysis. In *Proc. BMVC*. BMVA Press, 2014. 1