

# Multiscale Covariance Fields, Local Scales, and Shape Transforms

Diego H. Diaz Martinez<sup>1</sup>, Facundo Mémoli<sup>2</sup>, and Washington Mio<sup>1,\*</sup>

<sup>1</sup> Department of Mathematics, Florida State University  
Tallahassee, FL 32306-4510 USA, [mio@math.fsu.edu](mailto:mio@math.fsu.edu)

<sup>2</sup> School of Computer Science, The University of Adelaide, InnoVA21 Building  
Adelaide, SA 5005, Australia, [facundo.memoli@gmail.com](mailto:facundo.memoli@gmail.com)

**Abstract.** We introduce multiscale covariance fields associated with probability measures on Euclidean space and use them to define local scales at a point and to construct shape transforms. Local scales at  $x$  may be interpreted as scales at which key geometric features of data organization around  $x$  are revealed. Shape transforms are used to identify points that are most salient in terms of the local-global geometry of a probability distribution, yielding compact geometric summaries of the distribution.

**Keywords:** covariance fields, local scales, shape features, feature transforms

## 1 Introduction

The problem of extracting image features from a multiscale analysis has received a great deal of attention (cf. [3, 5, 6, 9]) as such features are useful in many image analysis and processing tasks. Techniques such as scale invariant feature transform (SIFT) employ a scale-space representation obtained by convolution with the Gaussian kernel to identify salient local features in images [6]. As important image features may be detected at different scales, a closely related problem is that of identifying *local scales* at a point  $x$ , that is, scales at which critical information about the local characteristics of the image near  $x$  can be obtained. The study of local scales of a shape, in contrast, still is incipient. In [8] there appears a study of local scales on parametric curves; in [7], a single local scale is defined for points in a shape, and [4] investigates local scales of submanifolds of Euclidean space.

Shape data often are acquired as noisy samplings of objects, so it is natural to treat these data as samples from probability distributions. In this paper, we introduce a notion of local scales of (Borel) probability measures  $\mu$  on Euclidean space  $\mathbb{R}^d$ , as well as shape transforms from which geometrically salient points of a distribution can be identified. Local scales and shape transforms will be derived from multiscale covariance fields associated with  $\mu$ . Scale dependence

---

\* Partially supported by NSF grant DBI-1052942

will be controlled by a kernel function  $K(x, y, \sigma) \geq 0$ , where  $x, y \in \mathbb{R}^d$  and  $\sigma > 0$  is the scale parameter. The idea is that, at scale  $\sigma$ , the kernel delimits the horizon of an observer positioned at  $x$  by attributing weight  $K(x, y, \sigma)$  to the point  $y$ . Covariation of the weighted data is measured relative to every point  $x \in \mathbb{R}^d$ , not just relative to the mean as is common practice, thus giving rise to a multiscale covariance field. Localized empirical covariance has been used in data analysis in *ad hoc* ways, but we develop a general framework for the formulation and treatment of such problems. Global covariance fields have been studied in [1]; however, for probability distributions on Euclidean space, they provide no information other than that already contained in the usual covariance tensor relative to the mean. Thus, a multiscale formulation is essential for gaining additional insight on local and regional data geometry.

Curvature carries rich information about the local geometry of embeddings of manifolds in Euclidean space. To offer evidence that multiscale covariance is effective in capturing the geometry of data sets, we show that the curvature of smooth curves in the plane and the principal curvatures of smooth surfaces in  $\mathbb{R}^3$  can be recovered from the small-scale covariance of singular measures induced by the volume form; that is, arc length of curves and area of surfaces. However, an advantage of the multiscale covariance approach is that these infinitesimal measures of geometric complexity easily extend to all scales and apply to more general probability distributions and measures, not just those supported on smooth objects.

The paper is organized as follows. In Section 2, we define multiscale covariance fields. Section 3 shows that the curvature of planar curves and the principal curvature of surfaces in  $\mathbb{R}^3$  can be recovered from small-scale covariance calculated with respect to a particular kernel function. In Section 4, we use covariance fields to define local scales and shape transforms that let us detect salient points in shapes. We also provide several illustrations using synthetic data and data obtained from the MPEG7 shape database.

## 2 Multiscale Covariance Fields

We begin with the definition of multiscale covariance fields associated with a (Borel) probability measure  $\mu$  on Euclidean space  $\mathbb{R}^d$ , with scale controlled by a kernel function  $K: \mathbb{R}^d \times \mathbb{R}^d \times (0, \infty) \rightarrow \mathbb{R}^+$ . In this paper, we focus on two types of kernels: the isotropic Gaussian kernel

$$G(x, y, \sigma) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right), \quad (1)$$

and the truncation kernel

$$T(x, y, \sigma) = \chi_{x, \sigma}(y), \quad (2)$$

where  $\chi_{x, \sigma}$  is the characteristic function of the ball of radius  $\sigma$  centered at  $x$ .

To define covariance fields, let us introduce some notation. Bilinear forms  $\mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  are identified with elements of the tensor product  $\mathbb{R}^d \otimes \mathbb{R}^d$ . In this

representation, a decomposable element  $v \otimes w$  corresponds to the bilinear form  $v \otimes w(x, y) = \langle v, x \rangle \cdot \langle w, y \rangle$ , where  $\langle \cdot, \cdot \rangle$  denotes Euclidean inner product. The multiscale covariance field of  $\mu$ , denoted  $\Sigma: \mathbb{R}^d \times (0, \infty) \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ , is defined as

$$\Sigma(x, \sigma) = \int_{\mathbb{R}^d} (y - x) \otimes (y - x) K(x, y, \sigma) d\mu(y), \quad (3)$$

provided that the integral is convergent.  $\Sigma(x, \sigma)$  is a non-negative bilinear form for any  $(x, \sigma)$ . If  $y_1, \dots, y_n$  are i.i.d.  $\mathbb{R}^d$ -valued random variables with distribution  $\mu$ , standard arguments show the consistency of the estimator

$$\Sigma_n(x, \sigma) = \frac{1}{n} \sum_{i=1}^n (y_i - x) \otimes (y_i - x) K(x, y_i, \sigma). \quad (4)$$

The normalized multiscale covariance field is defined as

$$\hat{\Sigma}(x, \sigma) = \frac{1}{Z(x, \sigma)} \int_{\mathbb{R}^d} (y - x) \otimes (y - x) K(x, y, \sigma) d\mu(y), \quad (5)$$

where  $Z(x, \sigma) = \int_{\mathbb{R}^d} K(x, y, \sigma) d\mu(y)$ , with the convention that  $\hat{\Sigma}(x, \sigma) = 0$  if  $Z(x, \sigma) = 0$ . Similarly, the empirical normalized covariance is given by

$$\hat{\Sigma}_n(x, \sigma) = \frac{1}{Z_n(x, \sigma)} \sum_{i=1}^n (y_i - x) \otimes (y_i - x) K(x, y_i, \sigma), \quad (6)$$

where  $Z_n(x, \sigma) = \frac{1}{n} \sum_{i=1}^n K(x, y_i, \sigma)$ .

*Remark.* Note that, for the truncation kernel  $T$ , the (normalized) covariance fields can be defined for any locally finite measure  $\mu$ ; that is, measures satisfying  $\mu(K) < \infty$  for any compact set  $K \subset \mathbb{R}^d$ .

### 3 Geometry of Curves and Surfaces

In this section, we show that much of the local geometry of planar curves and surfaces in  $\mathbb{R}^3$  can be recovered from covariance fields at small scales. The results support our more empirical use of these fields in the detection of local scales and in the construction of shape transforms.

#### 3.1 Planar Curves

Let  $C \subset \mathbb{R}^2$  be a smooth curve, that is, a 1-dimensional, smooth, properly embedded submanifold of  $\mathbb{R}^2$ . We consider the singular measure  $\mu$  on  $\mathbb{R}^2$  supported on  $C$  induced by arc length. This measure is locally finite because the embedding is proper. We calculate the small-scale normalized covariance at points on  $C$  for the truncation kernel and show that the curvature can be recovered from the eigenvalues of  $\hat{\Sigma}$ .

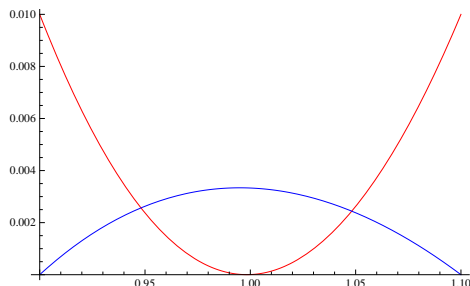
Let  $x_0 \in C$ . The arc-length parametrization of  $C$  near  $x_0$  may be written as  $X(s) = s - \frac{\kappa^2 s^3}{6} + O(s^4)$  and  $Y(s) = \frac{\kappa s^2}{2} + \frac{\kappa_s s^3}{6} + O(s^4)$ , where  $X(s)$  and  $Y(s)$

are coordinates along the tangent and normal to  $C$  at  $x_0$ , respectively. Here, the curvature  $\kappa$  and its derivatives  $\kappa_s$ , and  $\kappa_{ss}$  are evaluated at  $x_0$ . A calculation shows that in these coordinates

$$\hat{\Sigma}(x_0, \sigma) = \begin{pmatrix} \frac{\sigma^2}{3} - \frac{7\kappa^2\sigma^4}{180} + O(\sigma^6) & \frac{\kappa_s\sigma^4}{30} + O(\sigma^6) \\ \frac{\kappa_s\sigma^4}{30} + O(\sigma^6) & \frac{\kappa^2\sigma^4}{20} + O(\sigma^6) \end{pmatrix}. \quad (7)$$

The eigenvalues are  $\lambda_1 = \frac{\sigma^2}{3} - \frac{7\kappa^2\sigma^4}{180} + O(\sigma^6)$  and  $\lambda_2 = \frac{\kappa^2\sigma^4}{20} + O(\sigma^6)$ . Thus, one is able to recover curvature at  $x_0$  from small-scale covariance.

**The case of a circle** Let  $C_R$  be a circle of radius  $R$  centered at the origin in  $\mathbb{R}^2$ . If  $x_0$  is such that  $|\|x_0\| - R| > \sigma$  then  $\hat{\Sigma}(x_0, \sigma) = 0$ . Assume that  $x_0 \neq (0, 0)$  and  $\sigma > 0$  are such that  $\|x_0\| \in [R - \sigma, R + \sigma]$ . In this case, in the coordinate system given by the directions  $n := x_0/\|x_0\|$  and  $t := n^\perp$ , a calculation shows that  $\tilde{\Sigma}(x_0, \sigma)$  is diagonal with entries  $f_n(x_0, \sigma) = \frac{\phi(R^2 + 2\|x_0\|^2) + R(R \cos \phi - 4\|x_0\|) \sin \phi}{2\phi}$  and  $f_t(x_0, \sigma) = \frac{R^2(\phi - \sin \phi \cos \phi)}{2\phi}$ , where  $\phi = \arccos\left(\frac{R^2 + \|x_0\|^2 - \sigma^2}{2R\|x_0\|}\right)$ . In particular, this means that  $f_n$  is the eigenvalue corresponding to the eigenvector  $n$  along the normal direction, and  $f_t$  is the eigenvalue corresponding to the eigenvector  $t$  along the tangent direction – see Fig. 1.



**Fig. 1.** Plot of the eigenvalues  $f_n$  (in red) and  $f_t$  (in blue) for  $R = 1$  and  $\sigma = 0.1$  as a function of  $\|x_0\| \in [0.9, 1.1]$ . For  $\|x_0\| = 1$ ,  $f_n \simeq 5 \times 10^{-6}$  and  $f_t \simeq 0.033$ . Notice that starting from  $\|x_0\| = 1$ , as  $\|x_0\|$  grows, the two eigenvalues first become equal and then, as  $\|x_0\|$  approaches 1.1,  $f_t$  becomes dominant.

### 3.2 Surfaces in $\mathbb{R}^3$

Let  $S \subset \mathbb{R}^3$  be a smooth compact surface. Consider the singular measure  $\mu$  on  $\mathbb{R}^3$  supported on  $S$  induced by the area measure on  $S$ . We calculate the small-scale normalized covariance at points on  $S$  for the truncation kernel and show how its principal curvatures can indeed be recovered from the spectrum of  $\hat{\Sigma}$ .

Given a non-umbilic point  $x_0 \in S$ , one can choose a Cartesian coordinate system centered at  $x_0$  so that the  $x$ -axis is along the direction of maximal curvature at  $x_0$ , the  $y$ -axis is along the direction of minimal curvature at  $x_0$ , and the  $z$ -axis is along the normal at  $x_0$ .

**Proposition 1.** *Let  $\sigma > 0$  be small and  $x_0 \in S$  be non-umbilic. Then, in the coordinate system described above,  $\hat{\Sigma}(x_0, \sigma)$  is given by*

$$\begin{pmatrix} \frac{\sigma^2}{4} - \frac{(3\kappa_1 + \kappa_2)^2}{384}\sigma^4 + O(\sigma^5) & O(\sigma^5) & O(\sigma^4) \\ O(\sigma^5) & \frac{\sigma^2}{4} - \frac{(\kappa_1 + 3\kappa_2)^2}{384}\sigma^4 + O(\sigma^5) & O(\sigma^4) \\ O(\sigma^4) & O(\sigma^4) & \frac{3\kappa_1^2 + 2\kappa_1\kappa_2 + 3\kappa_2^2}{96}\sigma^4 + O(\sigma^5) \end{pmatrix},$$

where  $\kappa_1$  and  $\kappa_2$  are the principal curvatures of  $S$  at  $x_0$ .

It follows from this result that, for  $\sigma > 0$  small, the determinant and trace of  $\hat{\Sigma}(x_0, \sigma)$  equal

$$\frac{3\kappa_1^2 + 2\kappa_1\kappa_2 + 3\kappa_2^2}{1536}\sigma^8 + O(\sigma^9) \quad \text{and} \quad \frac{\sigma^2}{2} + \frac{(\kappa_1 - \kappa_2)^2}{192}\sigma^4 + O(\sigma^5),$$

respectively. As a consequence,  $\kappa_1$  and  $\kappa_2$  can be recovered from the spectrum of  $\hat{\Sigma}(x_0, \sigma)$  as a function of  $\sigma$ . Indeed, from the Taylor expansions of these two functions one can extract the values of  $(\kappa_1 - \kappa_2)^2$  and  $3\kappa_1^2 + 2\kappa_1\kappa_2 + 3\kappa_2^2$  from which one can determine the values of  $\kappa_1$  and  $\kappa_2$ .

*Proof (Proof of Proposition 1).* By considering cylindrical coordinates in the reference system that we have chosen, one can parametrize the patch  $S \cap B(0, \sigma)$  as  $(\rho \cos \phi, \rho \sin \phi, z(\rho, \phi))$  for  $\rho \in [0, \rho_\sigma(\phi)]$ ,  $\phi \in [0, 2\pi]$ , where  $\rho_\sigma(\phi) = \sigma - \frac{1}{8}(\kappa_1(\cos \phi)^2 + \kappa_2(\sin \phi)^2)^2\sigma^3 + O(\sigma^4)$  and  $z(\rho, \phi) = \frac{\rho^2}{2}(\kappa_1(\cos \phi)^2 + \kappa_2(\sin \phi)^2) + O(\sigma^3)$ . The area element on that surface patch is given by

$$dA = \left( \rho + \frac{\rho^3}{2}(\kappa_1^2(\cos \phi)^2 + \kappa_2^2(\sin \phi)^2) + O(\rho^5) \right) d\rho d\phi$$

so that the area of  $S \cap B(0, \sigma)$  equals  $Z(0, \sigma) = \pi\sigma^2 + \frac{\pi}{32}(\kappa_1 - \kappa_2)^2\sigma^4 + O(\sigma^5)$ . Now we have all the ingredients we need in order to compute  $\Sigma(0, \sigma)$ . For example, to calculate the  $(1, 1)$ -entry one computes  $\iint_{S \cap B(0, \sigma)} x^2 dA$  as

$$\int_0^{2\pi} \int_0^{\rho_\sigma(\phi)} \left( \rho^3 \cos^2(\phi) + \frac{1}{2}\rho^5 \cos^2(\phi) (\kappa_2^2 \sin^2(\phi) + \kappa_1^2 \cos^2(\phi)) + O(\rho^6) \right) d\rho d\phi,$$

which after a simple but tedious calculation yields the desired result. The computation of other entries of the matrix follows the same steps.

## 4 Local Scales and Shape Transforms

Let  $\mu$  be a (Borel) probability measure on  $\mathbb{R}^d$ . In a logarithmic representation of the scale parameter, the multiscale covariance field will be denoted  $\Lambda: \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d \otimes \mathbb{R}^d$ , where  $\Lambda(x, r) = \Sigma(x, e^r)$ .

We will use a simple measure of anisotropy to reduce the tensor field  $\Lambda$  to a scalar field  $h$ . Let  $A_{(x,r)}: \mathbb{R}^d \rightarrow \mathbb{R}^d$  be the non-negative self-adjoint operator associated with the 2-tensor  $\Lambda(x, r)$ . The operator  $A_{(x,r)}$  is characterized by

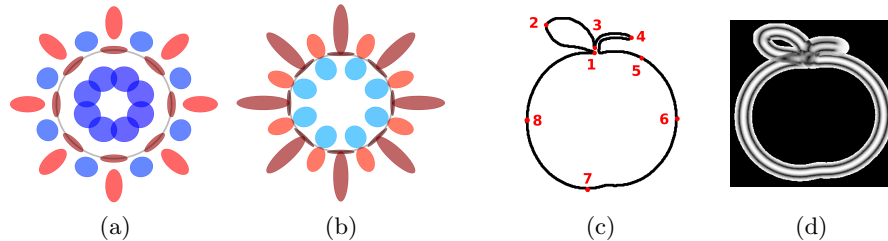
$A(x, r)(v, w) = \langle v, A_{(x,r)}(w) \rangle$ , for any  $v, w \in \mathbb{R}^d$ . If  $\lambda_1(x, r) \geq \dots \geq \lambda_d(x, r) \geq 0$  are the eigenvalues of  $A_{(x,r)}$ , define the anisotropy function  $h: \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  by

$$h(x, r) = 1 - \frac{\lambda_d(x, r)}{\lambda_1(x, r)}. \quad (8)$$

If  $\lambda_1(x, r) = 0$ , we set  $h(x, r) = 0$ . The function  $h$  takes values in  $[0, 1]$ , with 0 corresponding to fully isotropic covariance. Finer measures of anisotropy may be used, for example, to account for the decay of the eigenvalues of  $A_{(x,r)}$ . However, in this paper, we only consider  $h$  as defined in (8).

To describe some useful properties of anisotropy functions, we introduce the following notation. For any function  $f$  defined on  $\mathbb{R}^d \times \mathbb{R}$ , we let  $f_x$  be the function obtained by holding  $x$  fixed, so that  $f_x(r) = f(x, r)$ . Similarly, we define  $f_r(x) = f(x, r)$ .

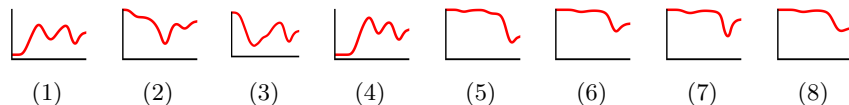
As a preliminary illustration, Figs. 2(a) and 2(b) show the normalized covariance fields (as fields of ellipses) – relative to the truncation and Gaussian kernels, respectively – at a fixed scale for a data set comprising 1,000 points on a circle sampled uniformly with respect to arc length. As expected, the field for the truncation kernel is highly anisotropic near the circle with variation essentially in directions tangential to the circle. Consistent with the calculations of Section 3, as  $x$  moves away from the circle,  $\Sigma_r(x)$  first becomes more isotropic, then becomes anisotropic again with high variation in the radial direction, finally vanishing at points that are at a distance from the circle larger than the value of the scale parameter. Fig. 2(c) shows the contour of an apple sampled at 872



**Fig. 2.** Normalized covariance field at a fixed scale for a uniform sampling of the circle calculated relative to the (a) truncation kernel and the (b) Gaussian kernel; (c) a sampling of the contour of an apple with eight points highlighted; (d) map of the anisotropy function  $h_r$  of the apple with respect to the truncation kernel.

points and Fig. 2(d) displays the anisotropy function  $h_r$  as an image. Fig. 3 shows the function  $h_x$ , calculated with respect to the Gaussian kernel, at the points on the apple highlighted in Fig. 2(c). The geometry of the curve near points 5, 6, 7, and 8 is simple and this is reflected in the fact that the small-scale covariance at those points is highly anisotropic with variation taking place primarily along the direction tangential to the curve. Only at larger scales those points “see” more

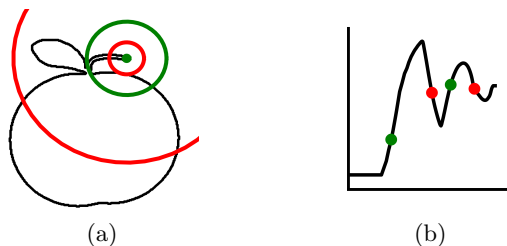
interesting shape. In contrast, there are richer geometric features near points 1, 2, 3, and 4, leading to a more complex oscillatory behavior of  $h_x$ .



**Fig. 3.** Anisotropy  $h_x$  at the points on the contour of an apple highlighted in Fig. 2(c).

#### 4.1 Local Scales

For a fixed  $x \in \mathbb{R}^d$ , rapid growth or decay of the anisotropy function  $h_x$  reveals scales at which pronounced changes occur in the behavior of covariance tensors centered at  $x$ . This suggests that values of  $r$  that correspond to local maxima and minima of the derivative of  $h_x$  represent geometrically important scales from the viewpoint of  $x$ . These will be referred to as *local scales* at  $x$ . We order local scales at  $x \in \mathbb{R}^d$  in an increasing sequence that identify scales at which the complexity of the geometry of a data set changes most significantly from the viewpoint of an observer at  $x$  as the observational horizon expands. Fig. 4 shows an example based on the truncation kernel  $T$ . Panel (a) highlights the balls of radius  $\sigma$  that correspond to local scales detected at the tip of the stem of an apple and panel (b) shows the full anisotropy function  $h_x$  at that point highlighting local scales.



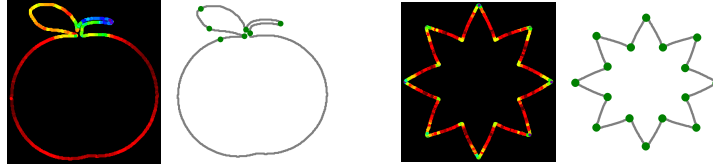
**Fig. 4.** (a) Local scales at the tip of the stem of an apple calculated with respect to the truncation kernel; (b) anisotropy function  $h_x$  at that point with local scales highlighted.

#### 4.2 Total Variation Transform

We use a reduction of the anisotropy function  $h$  to construct a *shape transform*  $TV_A: \mathbb{R}^d \rightarrow \mathbb{R}$ , where  $TV_A(x)$  is the total variation of  $h_x$  over a fixed finite interval.  $TV_A(x)$  provides a summary of the complexity of the geometry of the data distribution from the perspective of an observer positioned at  $x$ .

We calculated the total variation transform for a cyclically ordered sampling of the contour curve of the objects in Fig. 5. Panel (a) shows a smoothed map of the transform  $TV_A$  restricted to the data points. The local maxima of  $TV_A$  are ordered according to decreasing values of  $TV_A$  and panel (b) highlights the top

ranked points. The other figures show a similar example. As expected, the  $TV_A$  transform is able to detect points with most salient local-global geometry.



**Fig. 5.** Maps of the  $TV_A$  transform restricted to the contour of two objects and the most salient local maxima.

## 5 Summary and Discussion

We introduced the notion of multiscale covariance field modulated by a kernel function and employed it to identify local scales of data sets and probability distributions on Euclidean space. For probability measures supported on smooth data, we showed how curvature information can be recovered from small scale covariance. We also constructed shape transforms from covariance tensor fields that let us locate geometrically salient points in a probability measure. The present work only deals with probability measures on Euclidean space, but we will investigate extensions to a multiscale analysis of probability distributions and data on more general spaces such as Riemannian manifolds. Future plans also include study of other shape transforms derived from multiscale covariance fields.

## References

1. Balov, N.H.: Covariance fields. arXiv:0807.4690 (2008).
2. Feiszli, M., Jones, P.: Curve denoising by multiscale singularity detection and geometric shrinkage. *Applied and Computational Harmonic Analysis* 31(3), 392-409 (2011).
3. Jones, P., Le, T.: Local scales and multiscale image decompositions. *Applied and Computational Harmonic Analysis*, 26(3), 371–394 (2009)
4. Le, T., Mémoli, F.: Local scales on curves and surfaces. *Appl. Comput. Harmon. Anal.* 33, 401–437 (2012).
5. Lindeberg, T: Feature detection with automatic scale selection. *Int. J. Comput. Vis.* 30(2), 79–116 (1998).
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60(2), 91–110 (2004).
7. Luo, B., Aujol, J.F., Gousseau, Y., Ladjal, S., Matre, H.: Characteristic scale in satellite images. In: *ICASSP 2006*, vol. 2, 809–812 (2006).
8. Rosin, P.L.: Determining local natural scales of curves. *Pattern Recognition Letters*, 19(1):63–75 (1998).
9. Strong, D., Chan, T.: Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems* 19, S165–S187 (2003).