

Temporal Reflection Symmetry of Human Actions: A Riemannian Analysis

Qiao Wang¹²
qiao.wang@asu.edu
Rushil Anirudh¹²
ranirudh@asu.edu
Pavan Turaga²¹
pturaga@asu.edu

¹ School of Electrical, Computer and
Energy Engineering
Arizona State University
Tempe, Arizona, USA

² School of Arts, Media and Engineering
Arizona State University
Tempe, Arizona, USA

Abstract

While spatial symmetry of objects is studied frequently, the problem of measuring temporal symmetry has rarely been explored. In this paper we present a mathematical framework that utilizes a recently developed differential geometric approach to quantify temporal reflection symmetry in movements from video data, in which human movements are modeled as trajectories on Riemannian manifolds. Using the recently proposed transport square-root velocity functions (TSRVF) representation, the amounts of temporal asymmetry in movements are decomposed into a spatial component and a temporal one, each of which is quantified by an asymmetry score.

1 Introduction

Reflection symmetry exists ubiquitously in the world around us and attracts the attention of humans and animals [8]. It is often found to be correlated with quality in both practical and aesthetic senses [10], and is used by experts in various applications including scientific research, medical diagnosis, among many others (e.g. [12]). However, research so far has been mostly focused on spatial symmetry, both in computer vision and biomedical communities. The problem of quantifying temporal symmetry has rarely been explored. We argue that this is because of the lack of an advanced tool that can properly analyze temporal symmetry from high-dimensional signals, and a framework for such analytical tools is presented in this paper.

In current applications that do investigate temporal reflection symmetry, it is mostly associated with one-dimensional signals, two synthetic examples of which are given in Figure 1. The signals can be read as, for example, speed profiles of the hand in a reaching task, and temporally symmetric profiles represent the optimal movement strategy and are often observed in young and healthy subjects. The quantification of temporal asymmetry in this case can be used as an indicator of the motor performance decline due to aging [9]. Such quantification is often obtained by comparing the rising and falling time of the one-dimensional signal, where the rising and falling time is defined as the transition time between two threshold values (20% and 80% of the peak value, for example).

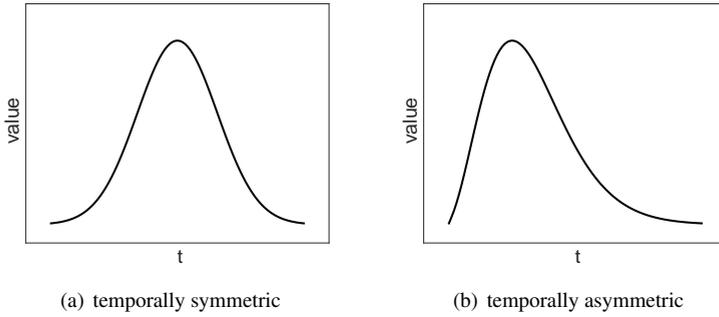


Figure 1: Temporal symmetry for one-dimensional signals

If a one-dimensional signal $f(t) \in \mathbb{R}, t \in [0, 1]$ meets the following criteria: f is continuous, $f(t), t \in [0, 1]$ consists of a monotonically increasing part and a monotonically decreasing part, and $f(0) = f(1)$, for any point $t \in [0, 1]$, $\exists t' \in [0, 1]$, where $f(t') = f(t)$ and $t' \neq t$ (except for the minimum or maximum point). In other words, exact spatial correspondences between points in f can always be obtained, thus the only asymmetry in f (if there is any) comes from the temporal component.

The same does not hold true for higher-dimensional signals, in which exact spatial correspondences cannot always be obtained. Figure 2 gives several synthetic examples of two-dimensional signals representing a reaching movement. The first part of the movement is

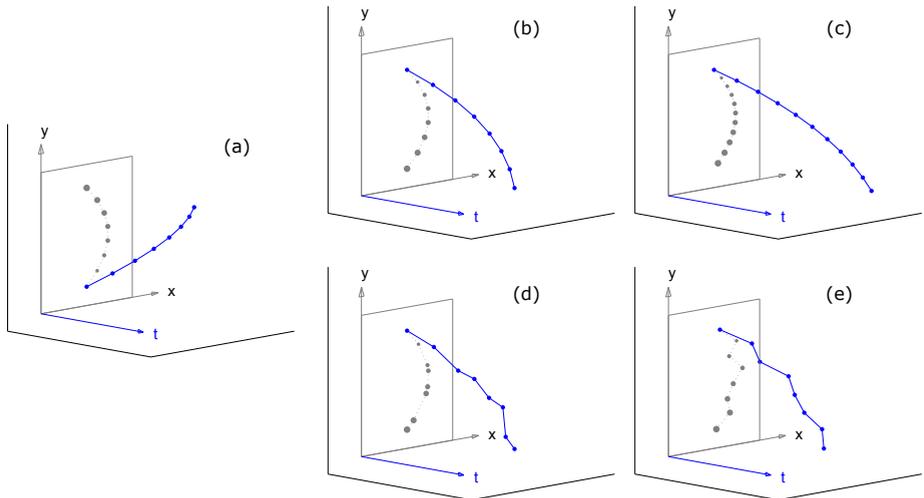


Figure 2: Synthetic examples of 2D signals showing that temporal asymmetry consists of both spatial and temporal discrepancies (In each subfigure, the blue solid line shows the sequence of samples $P(t) = (x, y)$ in $x - y - t$ space, whereas the gray dotted line shows the actual spatial trajectory on the $x - y$ plane. Dots are data points sampled at uniform intervals and larger gray dots represents points that are more recent).

given in (a), and the second part is in (b), (c), (d), or (e). It can be seen that exact spatial correspondences can only be obtained in the composite signal of (a) + (b), (a) + (c), or (a) + (d), but not in (a) + (e). The difference between two parts in the former cases are solely caused by different moving rates, while the difference in (a) + (e) is caused by spatial discrepancies. By using the one-dimensional examples in Figure 1 as analogies, we can say that signal (a) + (b) is similar to the ideally symmetric signal in Figure 1(a), while signals (a) + (c) and (a) + (d) are similar to the asymmetric signal in Figure 1(b), and (a) + (e) presents a new type of asymmetric signals that do not exist in one-dimensional space. Apparently it is also possible to have both temporal and spatial discrepancies between two parts in one signal. If we simply compare the sequences of samples between two parts, we will not be able to differentiate the two different types of asymmetry. However, since the two types are caused by different mechanisms, there are benefits in separating them. For example, in stroke rehabilitation, spatial alignment and speed variation are often computed separately to give patients more customized guidance [9].

As far as we know, there were few prior works that investigated the problem of quantifying temporal reflection symmetry, except for one-dimensional signal applications discussed above. For spatial symmetry, [18] proposed an early approach for automatically quantifying the degree of spatial symmetry of 2D shapes, by comparing a shape with its nearest symmetric counterpart. Some more recent papers presented more generalized frameworks for 2D shapes and 3D surfaces using various shape representations and optimization methods [6, 8, 9, 10]. Moreover, [7] utilized 2D spatial asymmetry quantifications to construct a shape descriptor that can be used for retrieval applications. Our method is similar to them in the sense that we also use the distance between a trajectory and its reflected version as a measure of asymmetry, and we use a trajectory aligning method that bears some similarity with the elastic deformation method used in [10]. However, our work differs in that we decompose temporal asymmetry into two components, which do not exist for spatial symmetries. Also our method analyzes movements as trajectories on Riemannian manifolds, while prior works investigated shapes/surfaces in 2D/3D Euclidean spaces.

2 Mathematical framework

There have been significant advances in applying differential geometry methods in computer vision in recent years. Non-Euclidean features have been developed and it has been demonstrated that they can often achieve better performances in various applications. A human action can generally be considered as a trajectory on a Riemannian manifold M with a Riemannian metric $\langle \cdot, \cdot \rangle$. For example, at any time instance, the human body can be represented as a shape silhouette, thus a point on the Grassmann manifold [14], or the combination of skeletal joints as a point in a Lie group $SE(3) \times \dots \times SE(3)$ [10]. The whole action is therefore a trajectory of such points on M . Using similar notations as in [13], let $\alpha : [0, 1] \rightarrow M$ denote a trajectory on M and we assume that α is smooth. Let \mathcal{M} denote the set of all such trajectories, and $A : \mathcal{M} \rightarrow \mathbb{R}_+$ be a function that maps a trajectory to a non-negative real number which represents an asymmetry score for the input trajectory. Also, let R denote the temporal reflection mapping, i.e. $R(t) = 1 - t, t \in [0, 1]$. The asymmetry score function A should have the following properties:

$$A(\alpha) = 0, \quad \text{if } \alpha = \alpha \circ R \quad (1)$$

$$A(\alpha) = A(\alpha \circ R) \quad (2)$$

For the spatial component of asymmetry (denoted as A_s) discussed in Section 1, it should have the additional property

$$A_s(\alpha) = A_s(\alpha \circ \gamma) \quad (3)$$

where $\gamma \in \tilde{\Gamma} = \{\gamma: [0, 1] \rightarrow [0, 1] \mid \gamma(0) = 0, \gamma(1) = 1, \gamma \text{ is nondecreasing, absolutely continuous}\}$.

Here γ is a warping function and $\alpha \circ \gamma$ is a trajectory that follows the same path of α on the manifold but has a different evolution rate. Thus Eq. (3) implies that the difference in evolution rates in different parts of α is not reflected in the quantification result of A_s . Therefore A_s only quantifies the spatial component of the discrepancy between α and $\alpha \circ R$.

Using the transport square-root velocity functions (TSRVF) representation recently proposed in [13], the TSRVF representation h of a trajectory α is a parallel transport of a vector field to a reference point $c \in M$:

$$h_\alpha(t) = \frac{\dot{\alpha}(t)_{\alpha(t) \rightarrow c}}{\sqrt{|\dot{\alpha}(t)|}} \quad (4)$$

and the distance between two TSRVFs is

$$d_h(h_{\alpha_1}, h_{\alpha_2}) = \left(\int_0^1 |h_{\alpha_1}(t) - h_{\alpha_2}(t)|^2 dt \right)^{1/2} \quad (5)$$

The equivalent class of h_α is defined as:

$$[h_\alpha] = \{h_{\alpha \circ \gamma} \mid \gamma \in \tilde{\Gamma}\} \quad (6)$$

where $\tilde{\Gamma}$ is the same as defined above.

We define the spatial component of our asymmetry scores (A_s) as the shortest distance between the equivalent classes of h_α and $h_{\alpha \circ R}$:

$$A_s(\alpha) = d_s([h_\alpha], [h_{\alpha \circ R}]) = \inf_{\gamma_1, \gamma_2 \in \tilde{\Gamma}} \left(\int_0^1 \left| h_\alpha(\gamma_1(t)) \sqrt{\dot{\gamma}_1(t)} - h_{\alpha \circ R}(\gamma_2(t)) \sqrt{\dot{\gamma}_2(t)} \right|^2 dt \right)^{1/2} \quad (7)$$

Here we prove that the function A_s defined in Eq. (7) has the desired properties defined in Eqs. (1-3):

First, if $\alpha = \alpha \circ R$, $A_s(\alpha) = d_s([h_\alpha], [h_{\alpha \circ R}]) = d_s([h_\alpha], [h_\alpha]) = 0$.

Second, $A_s(\alpha \circ R) = d_s([h_{\alpha \circ R}], [h_{\alpha \circ R \circ R}]) = d_s([h_{\alpha \circ R}], [h_\alpha])$. Because d_s is a proper distance [13], $d_s([h_{\alpha \circ R}], [h_\alpha]) = d_s([h_\alpha], [h_{\alpha \circ R}])$, thus $A_s(\alpha \circ R) = A_s(\alpha)$.

Third, for any $\gamma \in \tilde{\Gamma}$, $A_s(\alpha \circ \gamma) = d_s([h_{\alpha \circ \gamma}], [h_{\alpha \circ \gamma \circ R}])$. Since $R \circ \gamma \circ R(t) = 1 - \gamma \circ R(t) \in \tilde{\Gamma}$, $\exists \gamma' \in \tilde{\Gamma}, s.t. R \circ \gamma' = \gamma \circ R$. Thus $A_s(\alpha \circ \gamma) = d_s([h_{\alpha \circ \gamma}], [h_{\alpha \circ R \circ \gamma'}]) = d_s([h_\alpha], [h_{\alpha \circ R}]) = A_s(\alpha)$. \square

It is also noteworthy that since $\dot{R}(t) = -1$,

$$h_{\alpha \circ R} = -h_\alpha \circ R \quad (8)$$

Thus we can equivalently define A_s as:

$$A_s(\alpha) = d_s([h_\alpha], [-h_\alpha \circ R]) = \inf_{\gamma_1, \gamma_2 \in \tilde{\Gamma}} \left(\int_0^1 \left| h_\alpha(\gamma_1(t)) \sqrt{\dot{\gamma}_1(t)} + h_\alpha(1 - \gamma_2(t)) \sqrt{\dot{\gamma}_2(t)} \right|^2 dt \right)^{1/2} \quad (9)$$

In practice, instead of finding two warping functions γ_1 and γ_2 , we can find a $\gamma^* \in \Gamma = \{\gamma: [0, 1] \rightarrow [0, 1] | \gamma(0) = 0, \gamma(1) = 1, \gamma \text{ is a diffeomorphism}\}$ that aligns $h_{\alpha \circ R}$ with h_α , *i.e.* a warping function that minimizes $d_h(h_\alpha, h_{\alpha \circ R \circ \gamma^*})$ over Γ . The optimization is performed using the dynamic programming (DP) algorithm [10] and the time complexity is $O(N^2)$, N being the number of discrete samples in the interval $[0, 1]$.

We denote $\gamma_{(h_{\alpha_1}, h_{\alpha_2})}^*$ as the optimal warping function described above that aligns h_{α_2} with h_{α_1} , and the temporal component of our asymmetry scores is defined as:

$$A_t(\alpha) = \frac{1}{2} \int_0^1 \left(\left| \dot{\gamma}_{(h_\alpha, h_{\alpha \circ R})}^*(t) - 1 \right| + \left| \dot{\gamma}_{(h_{\alpha \circ R}, h_\alpha)}^*(t) - 1 \right| \right) dt \quad (10)$$

Intuitively, A_t measures the amounts of deviation of γ^* s from the ideal warping function $\gamma(t) = t$ (*i.e.* in cases where no warping is needed for optimal matching). It is apparent that A_t also meets the two criteria for asymmetry measures in Eqs. (1) and (2).

3 Experimental results

We test our proposed method on the UMD common activities dataset [14]. The main reason we choose this dataset for our preliminary experiment is because there are a decent number of action repetitions (10) for each unique combination of actor and action. The dataset is also suitable for visualization of actions and illustration of results.

Each action sample of the dataset is represented as a sequence of shapes, extracted from the actor’s silhouettes. Each shape is represented by an $m \times 2$ matrix L that contains the positions of m 2D points, uniformly sampled on the contour of the shape. The positions of the points are normalized so that the centroid of the m points is $(0, 0)$, and the norms of column vectors of L are 1. The subspace spanned by the column vectors of L is invariant to affine transforms, and can be considered as a point on a Grassmann manifold [15]. Therefore, an action sample is a trajectory on a Grassmann manifold. We refer interested readers to [15] and [10] for more details on the properties of the manifold.

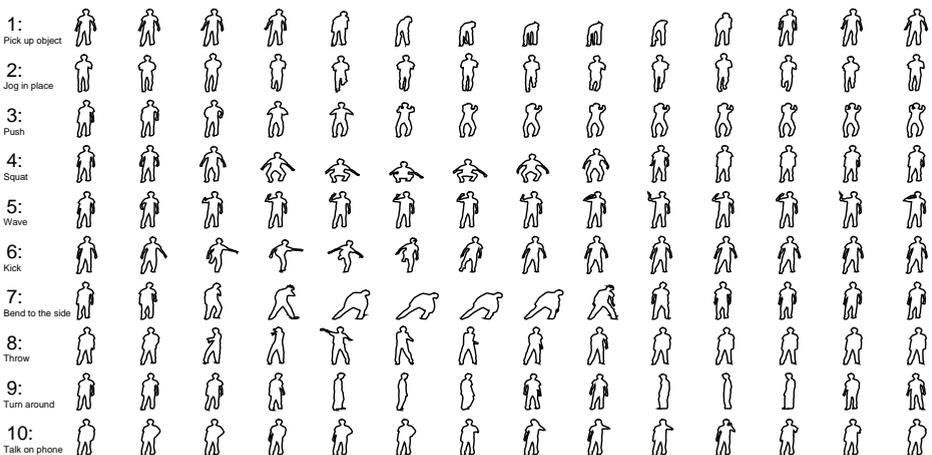


Figure 3: Actions whose temporal reflection symmetries are quantified (from the UMD common activities dataset)

We apply Eqs. (7) and (10) to all the action samples in the dataset. The mean and standard deviation values of asymmetry scores are summarized in Figures 4 and 5. For comparison, the asymmetry scores before warping are defined as $d_h(h_\alpha, h_{\alpha \circ R})$, where d_h is defined in Eq. (5). For illustration purposes, only the first 79 frames (out of 80) of each action sample are used for calculating asymmetry scores and illustration. Figure 3 illustrates the actions whose asymmetries are quantified. For each action class, the one sample that has the closest $A_s(\alpha)$ to the class mean is illustrated in Figure 3.

From the results shown in Figures 4, we can see that after applying our method, the actions can approximately be divided into two categories: the more temporally symmetric ones (“Pick up object”, “Jog in place”, “Squat”, “Wave”) and the less temporally symmetric ones (“Push”, “Kick”, “Throw”, “Turn around”, “Talk on phone”), while action “Bend to the side” lies in the middle. Had we simply calculated the asymmetry score as $d_h(h_\alpha, h_{\alpha \circ R})$ (as shown in Figure 4(a)), they would not be distinguishable. Figure 5 shows that the actions

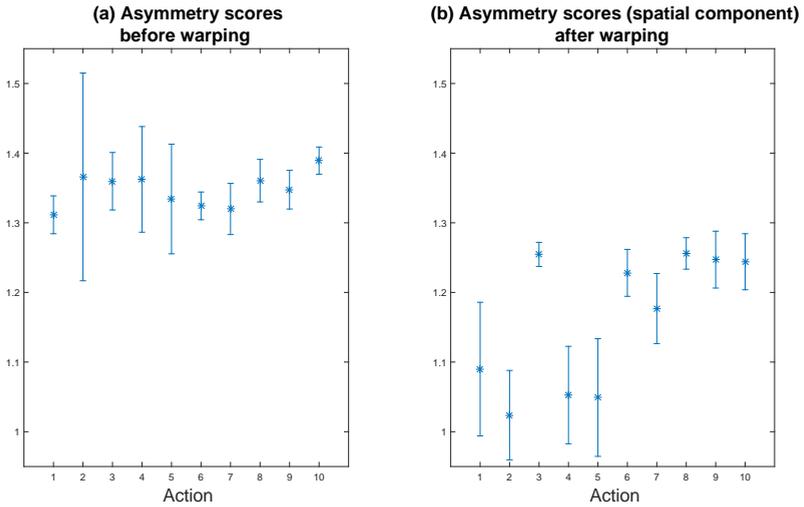


Figure 4: Means and standard deviations of asymmetry scores before and after warping (spatial components)

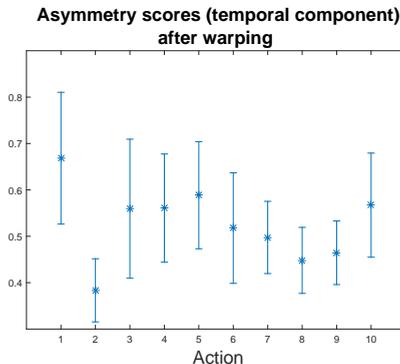


Figure 5: Means and standard deviations of asymmetry scores (temporal components)

have similar degrees of rate variation (reflected by the temporal component of asymmetry scores), except for the action of “Jog in place”, which has a lower degree of rate variation, and the difference is significant compared with a few other actions such as “Pick up object” or “Squat”.

To test the sensitivity to segmentation of these actions, we rerun the same algorithms using only the first 73 frames of each action sample instead of 79 (can be illustrated by removing the last column of shapes in Figure 3). The results remain almost the same, showing that our method is robust to small variation in action segmentations.

4 Summary

In this paper we have presented a framework for quantifying temporal reflection symmetry of movements by decomposing the amount of asymmetry into a spatial component and a temporal one. We consider movements as trajectories on Riemannian manifolds, and use TSRVF to properly align a movement with its temporal reflection. The use of differential geometric methods has helped improve the results of quantification.

Acknowledgments

This work was supported in part by NSF grant 1320267 and NSF CAREER grant 1452163.

References

- [1] Rushil Anirudh, Pavan Turaga, Jingyong Su, and Anuj Srivastava. Elastic functional coding of human actions: From vector-fields to latent variables. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3147–3155, 2015.
- [2] Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena Scientific Belmont, MA, 1995.
- [3] Yinpeng Chen, He Huang, Weiwei Xu, Richard Isaac Wallis, Hari Sundaram, Thanassis Rikakis, Todd Ingalls, Loren Olson, and Jiping He. The design of a real-time, multi-modal biofeedback system for stroke patient rehabilitation. In *Proceedings of the 14th annual ACM international conference on Multimedia*, pages 763–772. ACM, 2006.
- [4] A-M Ferrandez and Normand Teasdale. *Changes in sensory motor behavior in aging*, volume 114. Elsevier, 1995.
- [5] Michael Kazhdan, Thomas Funkhouser, and Szymon Rusinkiewicz. Symmetry descriptors and 3D shape matching. In *Proceedings of the 2004 Eurographics/ACM SIG-GRAPH symposium on Geometry processing*, pages 115–123. ACM, 2004.
- [6] Gert Kootstra, Arco Nederveen, and Bart De Boer. Paying attention to symmetry. In *Proceedings of the British Machine Vision Conference (BMVC2008)*, pages 1115–1125. The British Machine Vision Association and Society for Pattern Recognition, 2008.

- [7] Sebastian Kurtek, Mo Shen, and Hamid Laga. Elastic reflection symmetry based shape descriptors. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, pages 293–300. IEEE, 2014.
- [8] Aurélien Martinet, Cyril Soler, Nicolas Holzschuch, and François X Sillion. Accurate detection of symmetries in 3D shapes. *ACM Transactions on Graphics (TOG)*, 25(2): 439–464, 2006.
- [9] Niloy J Mitra, Leonidas J Guibas, and Mark Pauly. Symmetrization. In *ACM Transactions on Graphics (TOG)*, volume 26, page 63. ACM, 2007.
- [10] Anders Pape Møller and John P Swaddle. *Asymmetry, developmental stability and evolution*. Oxford University Press, 1997.
- [11] Chafik Samir, Anuj Srivastava, Mohamed Daoudi, and Sebastian Kurtek. On analyzing symmetry of objects using elastic deformations. In *VISAPP (1)*, pages 194–200, 2009.
- [12] Mikkel B Stegmann, Karl Skoglund, and Charlotte Ryberg. Mid-sagittal plane and mid-sagittal surface optimization in brain MRI using a local symmetry measure. In *Medical Imaging*, pages 568–579. International Society for Optics and Photonics, 2005.
- [13] Jingyong Su, Sebastian Kurtek, Eric Klassen, Anuj Srivastava, et al. Statistical analysis of trajectories on Riemannian manifolds: bird migration, hurricane tracking and video surveillance. *The Annals of Applied Statistics*, 8(1):530–552, 2014.
- [14] Pavan Turaga and Rama Chellappa. Locally time-invariant models of human activities using trajectories on the Grassmannian. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2435–2441. IEEE, 2009.
- [15] Pavan Turaga, Ashok Veeraraghavan, and Rama Chellappa. Statistical analysis on Stiefel and Grassmann manifolds with applications in computer vision. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.
- [16] Ashok Veeraraghavan, Rama Chellappa, and Amit K Roy-Chowdhury. The function space of an activity. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 959–968. IEEE, 2006.
- [17] Raviteja Vemulapalli, Felipe Arrate, and Rama Chellappa. Human action recognition by representing 3D skeletons as points in a Lie group. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 588–595. IEEE, 2014.
- [18] Hagit Zabrodsky, Shmuel Peleg, and David Avnir. Symmetry as a continuous feature. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(12):1154–1166, 1995.