Information measures in scale-spaces

Sporring, Jon; Weickert, Joachim

Published in:
IEEE Transactions on Information Theory

Publication date:
1999

Document Version
Early version, also known as pre-print

Citation for published version (APA):
Sporring, J., & Weickert, J. (1999). Information measures in scale-spaces. IEEE Transactions on Information Theory, 45(3), 1051-1058.

Many nonlinear methods have been proposed for constrained locally adaptive smoothing, such as local monotonicity [21], regularization [9], and wavelet denoising. These filters may be followed by multiscale-gradient estimation for edge detection, and represent an interesting higher complexity alternative to the linear smoothing filters studied here.

ACKNOWLEDGMENT

The authors are grateful to R. Kozick for his assistance with the EM algorithm, and to F. Gini for material related to the CRB's.

REFERENCES

[1] I. E. Abdou and W. K. Pratt, "Quantitative design and evaluation of enhancement/thresholding edge detectors," *Proc. IEEE*, vol. 67, no. 5, pp. 753–763, 1979.
[2] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes, Theory and Application.* Englewood Cliffs; NJ: Prentice-Hall, 1993.
[3] R. S. Blum, R. J. Kozick, and B. M. Sadler, "An adaptive spatial diversity receiver for non-Gaussian interference and noise," *IEEE Trans. Signal Processing*, to be published..
[4] A. W. Brown and T. J. Tukey, "Some distributions of sample means," *Annals of Math. Stat.*, vol. 17, no. 1, pp. 1–12, 1946.
[5] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 679–698, Nov. 1986.
[6] Z. Cvetkovic and M. Vetterli, "Discrete-time wavelet extreme representation: Design and consistent reconstruction," *IEEE Trans. Signal Processing*, vol. 43, pp. 681–693, Mar. 1995.
[7] A. Denjean and F. Castanie, "Mean value jump detection: A survey of conventional and wavelet based methods," in *Wavelets: Theory, Algorithms, and Applications*, C. Chui, L. Montefusco, and L. Puccio, Eds. Newkoski Academic, 1994.
[8] M. Ghogho and A. Swami, "Fast computation of the exact FIM for deterministic signals in colored noise," *IEEE Trans. Signal Processing*, vol. 47, pp. 52–61, Jan. 1999.
[9] M. Gokmen and C. C. Li, "Edge detection and surface reconstruction using refined regularization," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 5, pp. 492–499, 1993.
[10] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[11] R. Kakarala and A. O. Hero, "On achievable accuracy in edge localization," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 777–781, July 1992.
[12] J. Koplowitz and V. Greco, "On the edge location error for local maximum and zero-crossing edge detectors," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 12, pp. 1207–1212, 1994.
[13] R. Kozick, R. Blum, and B. M. Sadler, "Signal processing in non-Gaussian noise using mixture distributions and the EM algorithm," *in 31st Asilomar Conf. on Sig., Syst., and Comp.*, Pacific Grove, CA, Nov. 1997, pp. 438–442.
[14] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 1, pp. 21–28, 1995.
[15] L. A. Liporace, "Maximum likelihood estimation for multivariate observations of Markov sources," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 5, pp. 729–734, 1982.
[16] L. Ljung, *System Identification: Theory for the User* Englewood Cliffs, NJ: Prentice-Hall, 1987.
[17] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, pp. 617–643, 1992.
[18] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 710–732, July 1992.
[19] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley, 1997.
[20] K. S. Miller, *Multidimensional Gaussian Distributions*. New York: Wiley, 1964.
[21] A. Restrepo and A. C. Bovik, "Locally monotonic regression," *IEEE Trans. Signal. Processing.*, vol. 41, pp. 2796–2810, 1993.
[22] A. M. Reza and M. Doroodchi, "Cramer-Rao lower bound on locations of sudden changes in a steplike signal," *IEEE Trans. Signal Processing*, vol. 44, no. 10, pp. 2551–2556, 1996.
[23] A. Rosenfeld, "A nonlinear edge detection technique," *Proc. IEEE* pp. 814–816, May 1970.
[24] A. Rosenfeld, Y. H. Lee, and R. B. Thomas, "Edge and curve detection for texture discrimination," in *Picture Processing and Psychopictorics*. New York: Academic, 1970, pp. 381–393.
[25] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. 20, pp. 562–569, May 1971.
[26] B. M. Sadler, T. Pham, and L. C. Sadler, "Optimal and wavelet-based shockwave detection and estimation," *J. Acoust. Soc. Amer.*, vol. 104, no. 2, pt. 1, pp. 955–963, Aug. 1998.
[27] B. M. Sadler and A. Swami, "On multiscale wavelet analysis for step estimation," *in Int. Conf. on Acoustics, Speech, and Signal Processing* (ICASSP'98), Seattle, WA, vol. 3, pp. 1517–1520, Apr. 1998.
[28] ——, "Analysis of wavelet transform multiscale products for step detection and estimation," Army Research Laboratory Tech. Rep., ARL-TR-1664, 1998.
[29] A. Swami, "Cramer-Rao bounds for deterministic signals in additive and multiplicative noise," *Signal Processing*, vol. 53, pp. 231–244, 1996.
[30] A. Swami and B. M. Sadler, "Step-change localization in additive and multiplicative noise via multiscale products," *in 32nd Asilomar Conf. on Sig., Syst., and Comp.*, Pacific Grove, CA, Nov. 1998.
[31] ——, "Cramer-Rao bounds for step-change localization in additive and multiplicative noise," *in 9th IEEE Statistical Signal and Array Processing Workshop*, Portland, OR, Sept. 1998, pp. 403–406.
[32] H. D. Tagare and R. J. P. deFigueiredo, "On the localization performance measure and optimal edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 12, pp. 1186–1190, 1990.
[33] Y. Xu, J. Weaver, D. Healy, and J. Lu, "Wavelet transform domain filters: A spatially selective noise filtration technique," *IEEE Trans. Image. Processing*, vol. 3, pp. 747–758, Nov. 1994.

Information Measures in Scale-Spaces

Jon Sporring and Joachim Weickert

***Abstract*— This correspondence investigates Rényi's generalized entropies under linear and nonlinear scale-space evolutions of images. Scale-spaces are useful computer vision concepts for both scale analysis and image restoration. We regard images as densities and prove monotony and smoothness properties for the generalized entropies. The scale-space extended generalized entropies are applied to global scale selection and size estimations. Finally, we introduce an entropy-based fingerprint description for textures.**

***Index Terms*—Scale-space, Shannon–Wiener entropy, Rényi's generalized entropies, scale selection, size estimation, texture.**

I. INTRODUCTION

In recent years multiscale techniques have gained a lot of attention in the image processing community. Typical examples are pyramid and wavelet decompositions. They represent images at a small

Manuscript received February 10, 1998; revised October 15, 1998. This work was funded in part by the Real World Computing Partnership, the Danish National Research Council, and the EU-TMR Project VIRGO.

J. Sporring was with the Image Group, Department of Computer Science, University of Copenhagen, DK-2100 Copenhagen, Denmark. He is now with the Foundation for Research and Technology–Hellas (FORTH), GR 711 10 Heraklion, Crete, Greece.

J. Weickert is with the Image Group, Department of Computer Science, University of Copenhagen, DK-2100 Copenhagen, Denmark.

Publisher Item Identifier S 0018-9448(99)02269-5.

number of scales and have proven their use for many image-processing tasks. Another important class of multiscale techniques consists of so-called *scale-space representations* [1]–[4]. They embed an original image into a continuous family of subsequently simpler versions. Many scale-spaces can be formulated as the evolution of the initial image under a suitable linear or nonlinear diffusion process. Such an image evolution is useful for tasks such as feature extraction, scale selection, and segmentation, see [5]–[7] and references therein.

Information-theoretical concepts such as the Shannon–Wiener entropy [8], [9], Rényi's generalized entropies [10]–[12], and the Kullback–Leibler distance [13] have made contributions to image analysis; for instance, Brink and Pendock [14], Brink [15], and Sahoo *et al.* [17] have used them for local image thresholding, and Vehel *et al.* [17] and Chaudhuri and Sarkar [18] study images in a multifractal setting. It is not difficult to see that the generalized entropies, the multifractal spectrum, the gray-value moments and the gray-value histogram itself are equivalent representations: they can be transformed into each other by one-to-one mappings. More details can be found in the Appendix.

Since scale-spaces simplify images, it is only natural to investigate their simplification properties in terms of information measures. Already in 1949, Shannon mentioned that the Shannon–Wiener entropy decreases under averaging transformations [8, p. 52]. In 1993, Illner and Neunzert [19] studied a biased diffusion process, where the original image evolves toward a background image b along a path where its Kullback–Leibler distance with respect to b increases monotonically. Jägersand [20] used the Kullback–Leibler distance in linear scale-space for focus of attention. Oomes and Snoeren [21] used the entropy relative to a background measure to estimate the size of objects in images. Sporring [22] applied the Shannon–Wiener entropy in linear scale-space to perform scale selection in textures and showed the monotone behavior using concepts from thermodynamics. Weickert [23] proved monotony of the Shannon–Wiener entropy in linear and nonlinear diffusion scale-spaces by regarding it as a Lyapunov functional. Lyaponov functionals have been used for scale-space synchronization [24] and for a uniform sampling of the scale axis with respect to its information content [25]. The fractal dimension in scale-spaces has been investigated by Peleg *et al.* [26], Barth *et al.* [27], and Pei *et al.* [28]. Relations between Shannon–Wiener entropy and multiscale concepts in terms of wavelets have been established by Krim and Brooks [29], where inequality theory was applied to propose optimal measures for feature-directed segmentation.

The present correspondence extends previous work in this field by studying both theoretical aspects and the practical potential of generalized entropies in a linear and nonlinear multiscale setting. Generalized entropies are complete in the sense that they allow for a reconstruction of the gray-value histogram (see the Appendix). A scale-space extension is used to complement the entropies with spatial information. We prove monotony and smoothness properties with respect to the information order and the scale parameter. We use the scale-space behavior of generalized entropies for scale selection and size estimation, and we introduce a fingerprint-like description for textures. The results indicate that our extensions broaden the potential use of entropy methods in image analysis. Some preliminary results in this correspondence have been presented at conferences [22], [30].

Throughout this correspondence we identify an image by its two-dimensional distribution of light on a rectangular image domain. It should be noted that this representation is invariant under multiplication with, but not under addition of, a constant. It is important to note that this two-dimensional distribution is *not* the gray-value histogram.

The outline of our correspondence is as follows. In Section II we will give a brief introduction to linear and nonlinear scale-spaces. Then in Section III we will investigate a scale-space extension of the generalized entropies. Finally, in Section IV we will describe some applications in image processing. A conclusion is given in Section V.

II. A Short Introduction to Scale-Spaces

The images considered in this work are all discrete, but for simplicity we will in this section introduce two scale-spaces in the continuous setting. Discrete scale-space aspects are discussed by Lindeberg [5] for the linear framework, and by Weickert [23] for the nonlinear setting. Scale-spaces can be considered as an alternative to traditional smoothing methods from statistics [31].

In scale-space theory one embeds an image $p(\boldsymbol{x})$: $\mathbb{R}^2 \rightarrow \mathbb{R}$ into a continuous family $\{p(\boldsymbol{x}, t) | t \geq 0\}$ of gradually smoother versions of it. The original image corresponds to the scale $t = 0$, and increasing the scale should simplify the image without creating spurious structures. Since a scale-space creates a hierarchy of the image features, it constitutes an important step from a pixel-related image description to a semantical image description.

It has been shown that partial differential equations are the suitable framework for scale-spaces [32]. The oldest and best studied scale-space obtains a simplified version $p(\boldsymbol{x}, t)$ of $p(\boldsymbol{x})$ as the solution of the linear diffusion process with $p(\boldsymbol{x})$ as initial value

$$\partial_t p = \partial_{x_1 x_1} p + \partial_{x_2 x_2} p \tag{1}$$

$$p(\boldsymbol{x}, 0) = p(\boldsymbol{x}) \tag{2}$$

where $\boldsymbol{x} = (x_1, x_2)^T$. It is well known from the mathematical literature that the solution $p(\boldsymbol{x}, t)$ can be calculated by convolving $p(\boldsymbol{x})$ with a Gaussian of standard deviation $\sigma = \sqrt{2t}$

$$p(\boldsymbol{x}, t) = (G_t * p)(\boldsymbol{x}) \tag{3}$$

$$G_t(\boldsymbol{x}) := \frac{1}{4\pi t}\, e^{-|\boldsymbol{x}|^2/(4t)}. \tag{4}$$

This process is called *Gaussian scale-space* or *linear scale-space*. It was first discovered by Iijima [1], [2] and became popular two decades later through the work of Witkin [3] and Koenderink [4]. A detailed treatment of the various aspects of Gaussian scale-space theory can be found in [5], [33], [7], and the references therein.

Unfortunately, Gaussian smoothing also blurs and dislocates semantically important features such as edges. This has triggered the study of nonlinear scale-spaces. Perona and Malik [34] proposed to replace the linear diffusion equation (1) by the nonlinear diffusion process

$$\partial_t p = \nabla \cdot (g(|\nabla p|) \nabla p) \tag{5}$$

where $\nabla = (\partial_x, \partial_y)^T$, and the diffusivity $g(|\nabla p|)$ is a decreasing function of $|\nabla p|$. The idea is to regard $|\nabla p|$ as an edge detector and to encourage smoothing within a region over smoothing across boundaries. Thus locations where the gradient is large have a large likelihood of being an edge, and the diffusivity is reduced.

In our experiments we consider a nonlinear diffusion process where the diffusivity is given by [35]

$$g(|\nabla p|) := \frac{1}{\sqrt{1 + |\nabla p|^2/\lambda^2}} \qquad (\lambda > 0). \tag{6}$$

Such a choice guarantees that the nonlinear diffusion filter is well-posed.

This is one of the simplest representative of nonlinear scale-spaces. Overviews of other nonlinear scale-spaces can be found in [6] and [23].

III. Generalized Entropies

Let us now consider a discrete image $\boldsymbol{p} = (p_1, \cdots, p_N)^T$, where $p_i > 0$ for all i. Note that a single index is used for the two-dimensional enumeration of pixels. Its family of generalized entropies is defined as

$$S_\alpha(\boldsymbol{p}) := \frac{1}{1-\alpha}\ \log \sum_{i=1}^{N} p_i^\alpha \tag{7}$$

for $\alpha \neq 1$. The limit from left and right at $\alpha = 1$ is the Shannon–Wiener entropy

$$S_1(\boldsymbol{p}) = -\sum_{i=1}^{N} p_i \log p_i \tag{8}$$

and we might as well consider it as part of the continuum. The parameter α is called information order.

Let the vector-valued function $\boldsymbol{p}(t) = (p_1(t), \cdots, p_N(t))^T$ be the linear or nonlinear scale-space extension, where the continuous parameter t denotes scale. These scale-spaces can be obtained by a spatial discretization of (1) or (5) with reflecting boundary conditions.

We will now discuss some details of the mathematical structure of generalized entropies.

Proposition III.1: The generalized entropies are decreasing in α.

Proof: Follows immediately from [10], [36]. □

Proposition III.2: The generalized entropies $S_\alpha(\boldsymbol{p}(t))$ are increasing in t for $\alpha > 0$, constant for $\alpha = 0$, and decreasing for $\alpha < 0$. For $t \to \infty$, they converge to S_0.

Proof: The proof is based on a result from [23, Theorem 5]. For a discrete image $\boldsymbol{p}(t)$, which is obtained from a spatially discrete diffusion scale-space, the following holds. The expression

$$\Phi(\boldsymbol{p}(t)) := \sum_{i=1}^{N} r(p_i(t)) \tag{9}$$

is decreasing in t for every smooth convex function r. Moreover, $\lim_{t\to\infty}\ p_i(t) = 1/N$ for all i.

Using this we first prove the monotony of S_α with respect to t. Let $\alpha > 1$ and $s > 0$. Since $r(s) = s^\alpha$ satisfies

$$r''(s) = \alpha(\alpha-1)s^{\alpha-2} > 0 \tag{10}$$

it follows that r is convex. Thus

$$\Phi(\boldsymbol{p}(t)) = \sum_{i=1}^{N} r(p_i(t)) = \sum_{i=1}^{N} p_i^\alpha(t) \tag{11}$$

is decreasing in t and

$$S_\alpha(\boldsymbol{p}(t)) = \frac{1}{1-\alpha}\ \log \Phi(\boldsymbol{p}(t)) \tag{12}$$

is increasing in t.

Similar reasoning can be applied to establish monotony for the cases $0 < \alpha < 1$ and $\alpha < 0$.

For $\alpha = 1$ we obtain the Shannon–Wiener entropy for which monotony has already been shown in [23].

Let $\alpha = 0$. Then

$$S_0(\boldsymbol{p}(t)) = \log \sum_{i=1}^{N} p_i^0(t) = \log N = \text{const.} \quad \forall t. \tag{13}$$

To verify the asymptotic behavior of the generalized entropies we utilize $\lim_{t\to\infty}\ p_i(t) = 1/N$. For $\alpha \neq 1$ this gives

$$\lim_{t\to\infty}\ S_\alpha(\boldsymbol{p}(t)) = \frac{1}{1-\alpha}\ \log \sum_{i=1}^{N} \frac{1}{N^\alpha} = \log N = S_0 \tag{14}$$

and $\alpha = 1$ yields

$$\lim_{t\to\infty}\ S_1(\boldsymbol{p}(t)) = -\sum_{i=1}^{N} \frac{1}{N}\ \log \frac{1}{N} = \log N = S_0. \tag{15}$$

This completes the proof. □

The following smoothness results constitute the basis for studying derivatives of generalized entropies as will be done in Section IV.

Proposition III.3: The generalized entropies are C^∞ for $\alpha \neq 1$ and at least C^1 in $\alpha = 1$. For linear scale-space they are C^∞ in t, and for the nonlinear scale-space they are C^1 in t.

Proof: In order to prove smoothness with respect to α, we first consider the case $\alpha \neq 1$. Then S_α is the product of the two C^∞ functions $1/(1-\alpha)$ and $\log \Sigma_{i=1}^{N}\ p_i^\alpha$, and thus also C^∞ in α. The smoothness in $\alpha = 1$ is verified by applying l'Hôpital's rule. Straightforward calculations show that

$$\lim_{\alpha\to 1}\ \frac{\partial S_\alpha}{\partial \alpha} = \frac{\sum\limits_{i=1}^{N} p_i(\log p_i)^2 - \left(\sum\limits_{i=1}^{N} p_i \log p_i\right)^2}{2}. \tag{16}$$

Thus $\partial S_\alpha/\partial\alpha$ exists and S_α is in C^1.

For linear scale-space, C^∞ in t follows directly from the fact that $G_t(x)$ is in C^∞ with respect to t. In the nonlinear case, C^1 in t is a consequence of the fact that the solution $\boldsymbol{p}(t)$ is in C^1 with respect to t. This is proven in [23, Theorem 4]. □

Fig. 1 illustrates the monotony of the generalized entropies both in scale and order for both scale-spaces. The figures have been created by finite difference algorithms which preserve the monotony properties established in this section [37].

IV. Experiments

We will in this section demonstrate some applications for the generalized entropies in image processing. We will consider the change of entropies by logarithmic scale

$$c_\alpha(\boldsymbol{p}(t)) := \frac{\partial S_\alpha(\boldsymbol{p}(t))}{\partial(\log\ t)} \tag{17}$$

since this appears to be the natural parameter (at least for linear scale-space) [4], [38], [5, Sec. 8.7.1], [30]. We emphasize that the generalized entropies are global measures and are thus best suited for images with homogeneous textures.

A. Shannon–Wiener Entropy and Zooming

This section analyzes the zooming behavior of the Shannon–Wiener entropy in linear scale-space.

Fig. 2 (top left and right) shows images from a laboratory experiment. The camera is placed fronto-parallel to a plane with a simple texture: pieces of paper with discs arranged in a regular manner. A sequence is produced as a series of increasing zoom values. In Fig. 2 (bottom) we plot the scale $\sigma = \sqrt{2t}$ of the point of maximum entropy change against the mean size of the discs. As can be seen, the relation is close to linear. It appears that in linear scale-space the point of maximal entropy change by logarithmic scale corresponds to the size of the dominating image structures.

B. Spatial Extent of Structures

In this section we show that the scaling behavior in linear scale-space carries over to the generalized entropies, and that they can be used to simultaneously measure the size of light and dark structures. We shall also see that the latter cannot be done with the Shannon–Wiener entropy.

The idea is as follows: The definition of the generalized entropies implies that entropies for large positive α focus on high gray-values (white areas), while for large negative value they analyze low gray-values (dark areas).

We expect that $c_\alpha(\boldsymbol{p}(t))$ is especially high for structures of diameter d, when the variance $\sigma^2 = t/2$ of the Gaussian is close to the variance

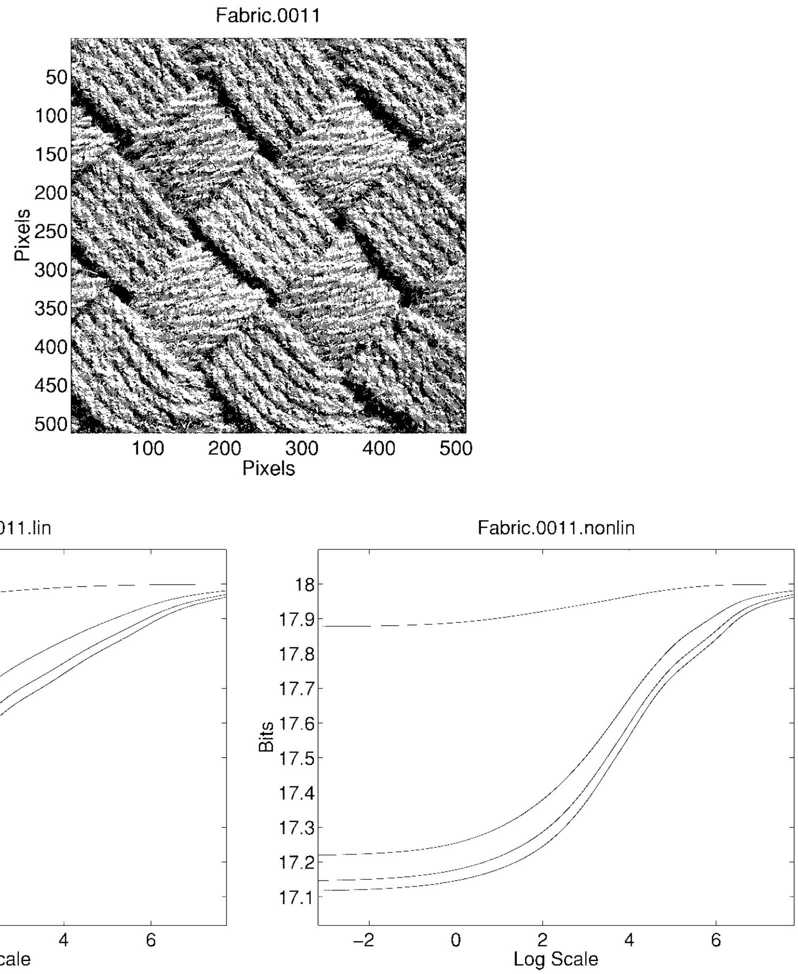

Fig. 1. Examples of some generalized entropies. Top: A 512×512 gray-valued image. Bottom Left: Generalized entropies in linear scale-space. From top to bottom $\alpha = 1, 34, 67, 100$. Bottom Right: The same for nonlinear scale-space.

of the structures. Let us, for simplicity, consider a random variable with uniform probability density function whose support is a disc of diameter d. Its variance is

$$\sigma^2 = \int_0^{2\pi} \int_0^{d/2} \frac{r^2}{\pi (d/2)^2}\, r\, dr\, d\phi = \frac{d^2}{8}. \tag{18}$$

Hence we expect a light (or dark) structure of diameter d to have a significant entropy change by logarithmic scale at time $\sigma^2/2 = d^2/16$. This size estimate remains qualitatively correct for nondisc structures. In this case, it gives the size of the largest minimal diameter.

Fig. 3 shows the result of a performance analysis. The size estimate (18) has been applied to a number of simple sinusoidal images with structures (half-wavelengths) between 1 and 257 pixels. It can be seen in the bottom graph, for sufficiently large structures, that the estimated sizes are close to the true size. Although by definition, the generalized entropies are not symmetric in order, both have a similar scaling behavior which is close to linear.

In Fig. 4 we show an experiment on a texture with a more complicated periodicity. This real image has been created by the Belousov–Zhabotinsky reaction [39]. From orders ± 20 we find dominating low-intensity values corresponding to a diameter 7.2, while the dominating high-intensity values suggest structures of diameter 3.5. From this we conclude that the distance between the light spiral arms in the mean is approximately 7.2 pixels, and the width of the spiral arms is approximately 3.5 pixels. In spite of the fact that the disc model (18) is not very appropriate for the line-like structure, the size estimates are in the correct order of magnitude.

The Shannon–Wiener entropy cannot be used for size estimation since it is a mixture of information from both light and dark areas. Thus is does not allow for a distinction between fore- and background.

C. Fingerprints for Entropies in Scale-Space

Sections IV-A and IV-B have shown that the scales of extremal entropy change carry significant information for selected information orders. Thus it would be interesting to introduce a compact description of the extremal changes for the continuum of information orders. In analogy with edge analysis in linear scale-space [40] we call such a description a fingerprint image. In Fig. 5 fingerprint images for two textures are given, both in the linear and nonlinear scale-space. The fingerprint lines are the extrema of $c_\alpha(\boldsymbol{p}(t))$ in t. Our monotony results immediately imply the following consequences: If there is only one fingerprint line for a given positive order, then it corresponds to a maximum (likewise, to a minimum for negative orders); see also Fig. 3. For almost all orders there will be an odd number of fingerprint lines, that correspond to alternating maxima and minima. This can be seen, for instance, in the middle right graph in Fig. 5. For information order 60, the leftmost line is a maximum followed by alternating minima and maxima.

It appears that the location of the fingerprint lines is more stable over information orders for the nonlinear scale-space than for the

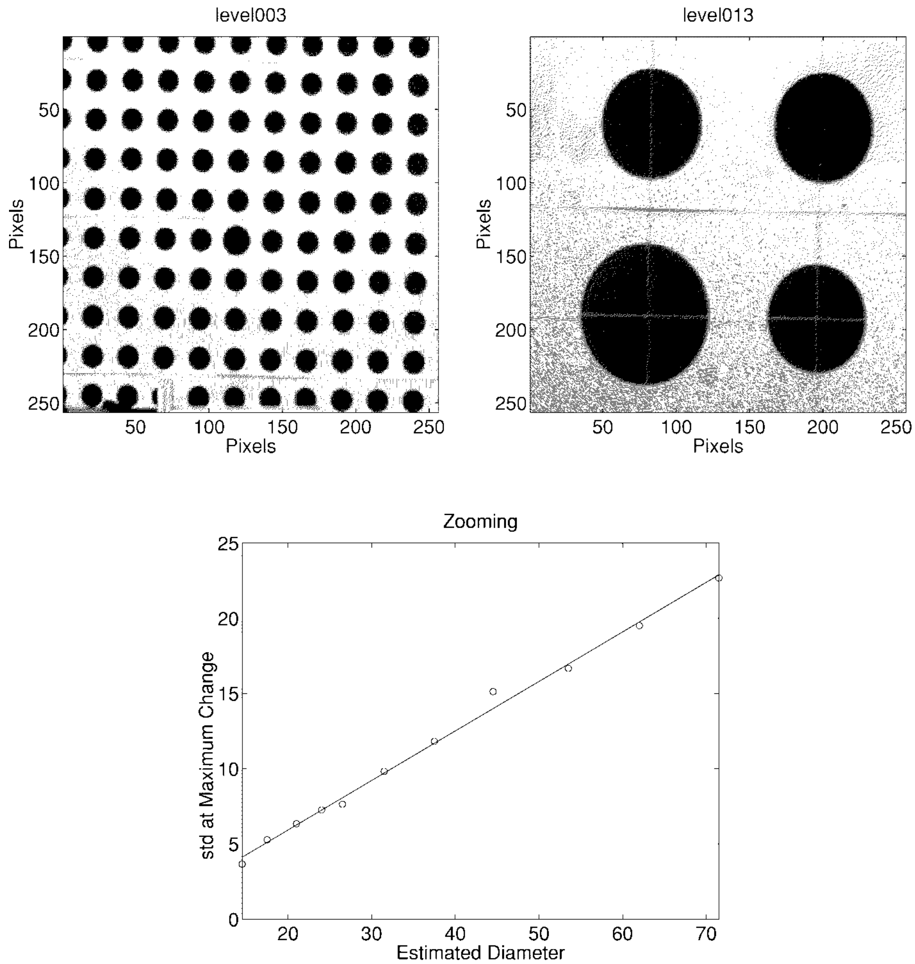

Fig. 2. A zooming sequence. Top: First and last image. Bottom: The $\sigma = \sqrt{t/2}$ values maximizing $c_1(\boldsymbol{p}(t))$ versus the estimated disc sizes.

linear one. Due to the reduced diffusivity of the nonlinear scale-space, the fingerprint lines are shifted toward higher scales.

V. Conclusions

In this correspondence we have investigated entropies as a means for extracting information from scale-spaces. This has lead to the following contributions.

- Monotony and smoothness properties for the Shannon–Wiener entropy and Rényi's generalized entropies have been proven for the linear and nonlinear diffusion scale-spaces. The proofs hold also for all other nonlinear diffusion scale-spaces treated in [23].
- We have illustrated that the generalized entropies can be used to perform size measurements for periodic textures. This is not possible with the Shannon–Wiener entropy. We have proceeded to define a fingerprint image for entropies in scale-space and analyzed some of its basic properties. The localization of the fingerprint lines can be improved using nonlinear instead of linear scale-space.

The following topics appear promising for future work.

- In the context of texture analysis, it would be interesting to perform an in-depth study on the relation between the fingerprint topology and the structure of the texture.
- This correspondence has focused on the maximal entropy change by scale to estimate the size of image structures. The minimal change by scale, however, indicates especially stable scales with respect to evolution time. We expect these scales to be good candidates for stopping times in nonlinear diffusion scale-spaces.
- The entropies in this correspondence are global measures. For topics such as focus-of-attention it would be interesting to study local variants of them.

It should be emphasized that the analysis carried out in this correspondence is directly transferable to the analysis of multifractals, gray-value moments, and gray-value histograms.

Appendix
Relations to Gray-Value Moments, Histograms, and Multifractal Spectra

The gray-value moments of an image are defined as

$$m_\alpha(\boldsymbol{p}) = \sum_{i=1}^{N} p_i^\alpha. \tag{19}$$

From the definition of S_α in (7) it is clear that there is a one-to-one relation to m_α.

Let the image $(p_1, \cdots, p_N)^T$ consist of M distinct gray-values $v_1, \cdots, v_M$ occurring $f_1, \cdots, f_M$ times. We may use this gray-value histogram $\boldsymbol{f}$ to rewrite the moments as

$$m_\alpha(\boldsymbol{p}) := \sum_{j=1}^{M} f_j v_j^\alpha. \tag{20}$$

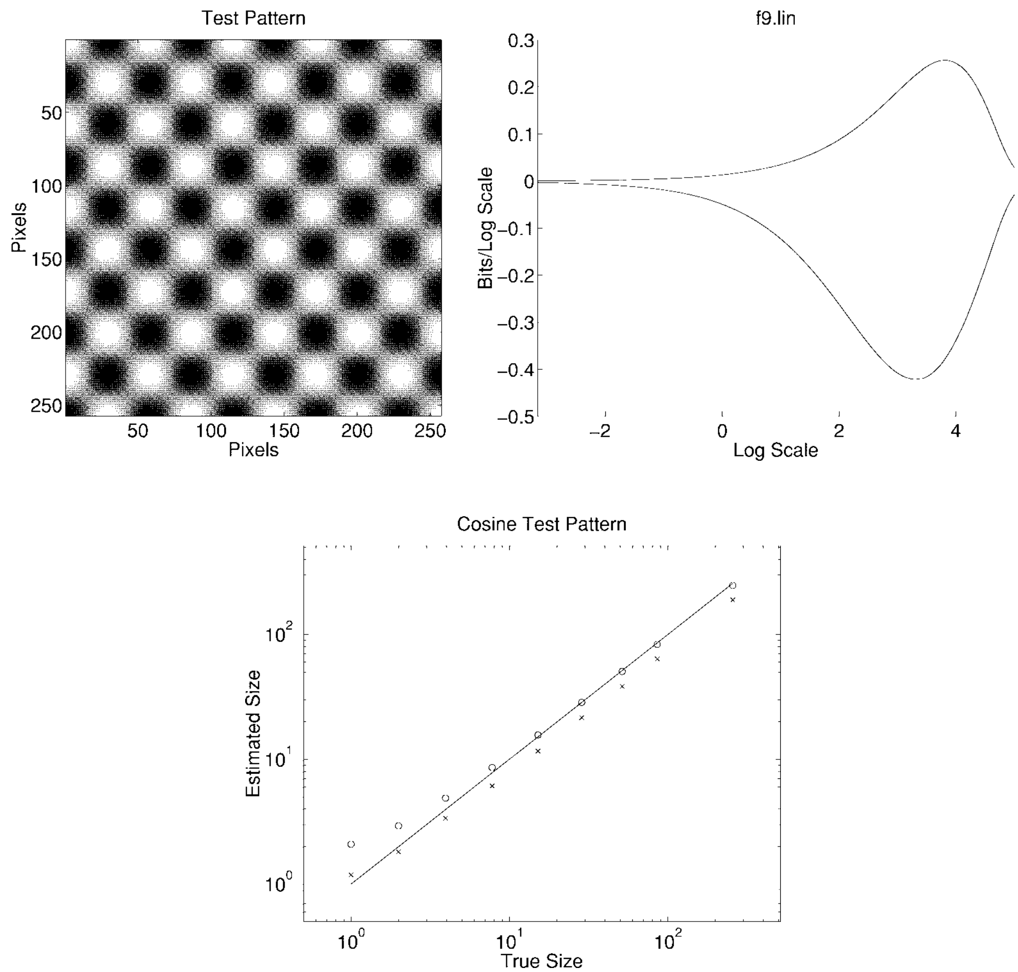

Fig. 3. Scaling behavior and size estimation with generalized entropies. Top Left: Test image generated by $257^{-2}(1 + 0.6 \cos(\omega x_1) \cos(\omega x_2))$ with $\omega = 9\pi/257$. Top Right: The corresponding $c_\alpha(\boldsymbol{p}(t))$ curves for $\alpha = \pm 100$. Top curve is for positive order and bottom curve for negative order. Bottom: A double logarithmic plot of the true size versus the estimated size for various ω. The straight line depicts the truth, the circles the estimation from order 100, and the crosses for order -100.

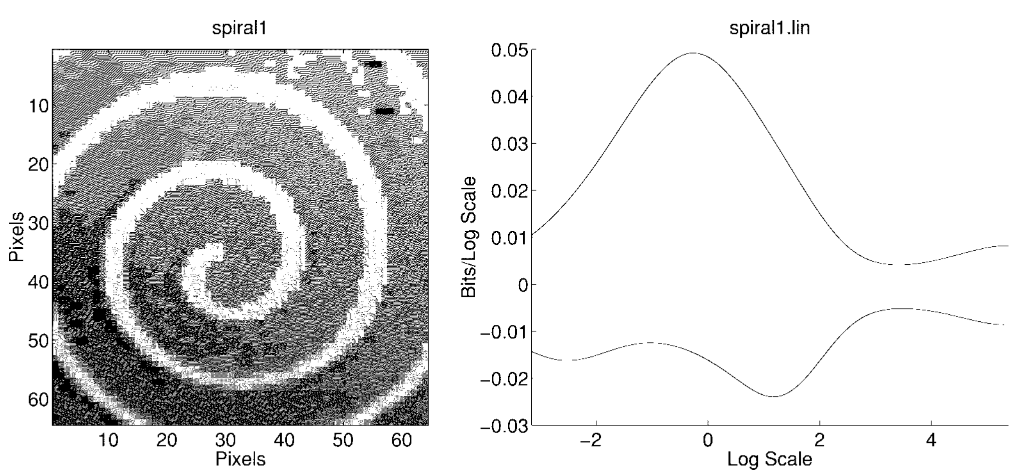

Fig. 4. Left: Spiral generated by a chemical reaction. RIght: Entropy changes for orders 20 (top curve) and -20 (bottom curve).

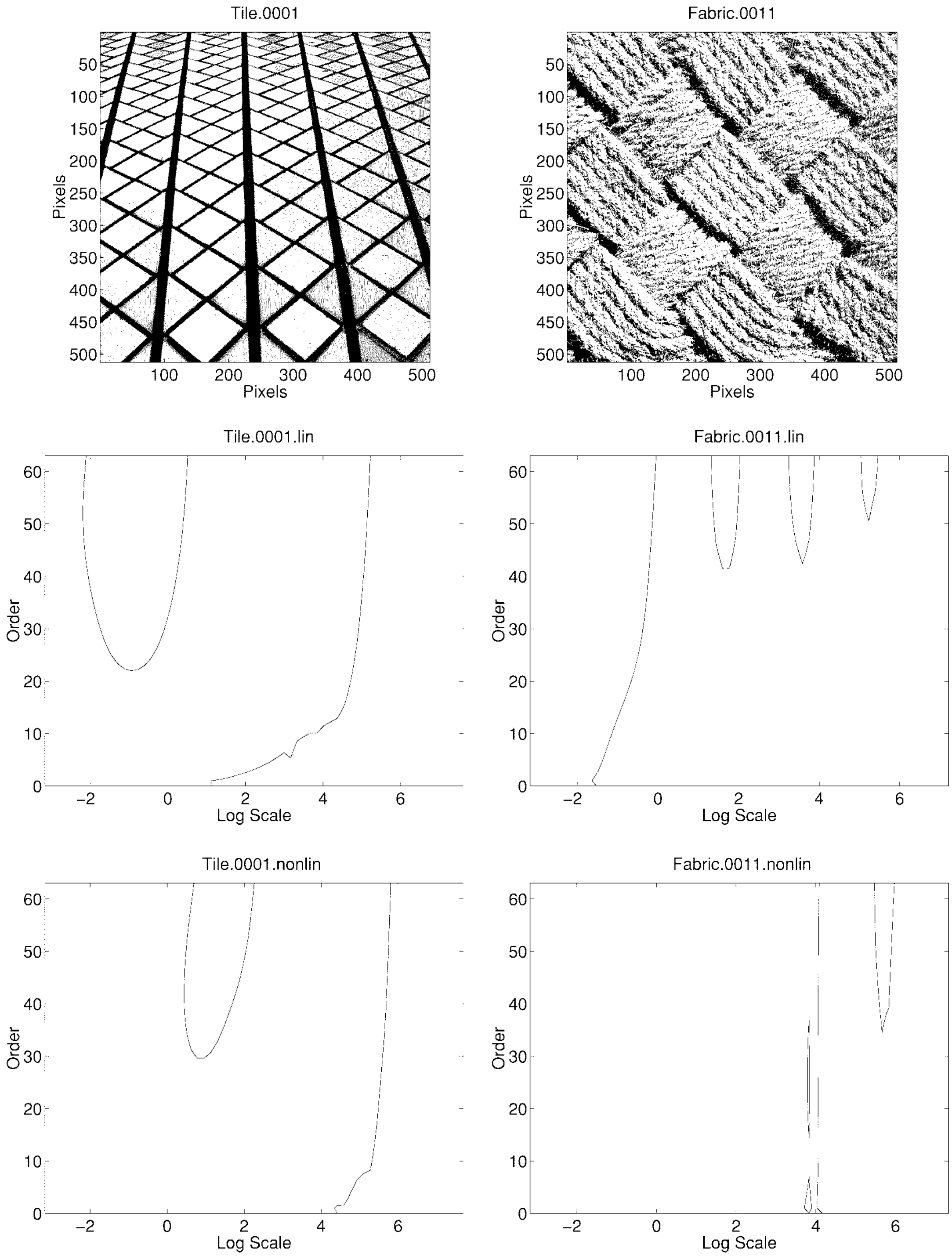

Fig. 5. Fingerprints of generalized entropies. Top Row: Two textures. Middle Row: Fingerprints for linear scale-space. Bottom Row: The same for nonlinear scale-space.

Considering the moments $m_0, \cdots, m_{M-1}$ gives the relation

$$\begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ \vdots \\ m_{M-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ v_1 & v_2 & \cdots & v_M \\ v_1^2 & v_2^2 & \cdots & v_M^2 \\ & & \ddots & \\ v_1^{M-1} & v_2^{M-1} & \cdots & v_M^{M-1} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_M \end{bmatrix}.$$

The system matrix is the so-called Vandermonde matrix. By induction over M the determinant can be shown to be $\prod_{1 \leq n < m \leq M} (v_m - v_n)$. Since $v_j, j = 1, \cdots, M$ are distinct, the matrix is invertible (but ill-conditioned). Thus there is a one-to-one relation between the moments $m_0, \cdots, m_{M-1}$ and the histogram $f_1, \cdots, f_M$.

The equivalence of the multifractal spectrum and the generalized entropies is discussed in [42].

Acknowledgment

We thank P. Johansen, M. Nielsen, L. Florack, O. F. Olsen, and R. Maas for many discussions on this topic. P. G. Sørensen from the Department of Chemistry at the University of Copenhagen [39] has supplied the spiral image in Fig. 4. The images in Fig. 5 are from the 'VisTex' collection [41].

REFERENCES

[1] T. Iijima, "Basic theory on normalization of a pattern (in case of typical one-dimensional pattern)," *Bull. Electrotech. Lab.*, vol. 26, pp. 368–388, 1962, in Japanese.

[2] J. Weickert, S. Ishikawa, and A. Imiya, "On the history of Gaussian scale-space axiomatics," in J. Sporring, M. Nielsen, L. Florack, and P. Johansen, Eds., *Gaussian Scale-Space Theory.* Dordrecht, The Netherlands: Kluwer, 1997, ch. 4, pp. 45–59.

[3] A. P. Witkin, "Scale–space filtering," in *Proc. 8th Int. Joint Conf. Artificial Intelligence (IJCAI '83)* (Karlsruhe, Germany), vol. 2, pp. 1019–1022, Aug. 1983.

[4] J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.

[5] T. Lindeberg, *Scale-Space Theory in Computer Vision* (Kluwer Int. Ser. in Engineering and Computer Science). Boston, MA: Kluwer, 1994.

[6] B. M. ter Haar Romeny, Ed., *Geometry-Driven Diffusion in Computer Vision* (Number 1 in the series *Computational Imaging and Vision*). Dordrecht, The Netherlands: Kluwer, 1994.

[7] J. Sporring, M. Nielsen, L. Florack, and P. Johansen, Eds., *Gaussian Scale-Space Theory.* Dordrecht, The Netherlands: Kluwer, 1997.

[8] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* Urbana, IL: Univ. Illinois Press, 1949.

[9] N. Wiener, *Cybernetics.* New York: Wiley, 1948.

[10] A. Rényi, "Some fundamental questions of information theory," in P. Turán, Ed., *Selected Papers of Alfréd Rényi.* Budapest, Hungary: Akadémiai Kiadó, 1976, pp. 526–552 (originally: *MTA III. Oszt. Közl.*, vol. 10, pp. 251–282, 1960).

[11] ——, "On measures of entropy and information," in P. Turán, Ed., *Selected Papers of Alfréd Rényi.* Budapest, Hungary: Akadémiai Kiadó, 1976, pp. 565–579 (originally in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, vol. 1. Berkeley, CA: Univ. California Press, 1961, pp. 547–561).

[12] ——, "On the foundations of information theory," in P. Turán, Ed., *Selected Papers of Alfréd Rényi.* Budapest, Hungary: Akadémiai Kiadó, 1976, pp. 304–318 (originally: *Rev. Inst. Int. Statist.*, vol. 33, pp. 1–14, 1965).

[13] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, pp. 79–86, 1951.

[14] A. D. Brink and N. E. Pendock, "Minimum cross-entropy threshold selection," *Pattern Recogn.*, vol. 29, no. 1, pp. 179–188, Jan. 1996.

[15] A. D. Brink, "Using spatial information as an aid to maximum entropy image threshold selection," *Pattern Recogn. Lett.*, vol. 17, no. 1, pp. 29–36, Jan. 1996.

[16] P. Sahoo, C. Wilkins, and J. Yeager, "Threshold selection using Renyi's entropy," *Pattern Recogn.*, vol. 30, pp. 71–84, 1997.

[17] J. Lévy Véhel and P. Mignot, "Multifractal segmentation of images," *Fractals*, vol. 2, no. 3, pp. 379–382, June 1994.

[18] B. B. Chaudhuri and N. Sarkar, "Texture segmentation using fractal dimension," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 1, pp. 72–77, Jan. 1995.

[19] R. Illner and H. Neunzert, "Relative entropy maximization and directed diffusion equations," *Math. Meth. Appl. Sci.*, vol. 16, pp. 545–554, 1993.

[20] M. Jägersand, "Saliency maps and attention selection in scale and spatial coordinates: An information theoretic approach," in *5th Int. Conf. Computer Vision*. IEEE Computer Soc. Press, June 1995, pp. 195–202.

[21] A. Oomes and P. R. Snoeren, "Structural information in scale-space," in *Proc. Copenhagen Worksh. Gaussian Scale-Space Theory*, P. Johansen, Ed. (Universitetsparken 1, DK-2100 Copenhagen, Denmark, May 10–13 1996), pp. 48–57, DIKU-96/19.

[22] J. Sporring, "The entropy of scale-space," in *Proc. 13th Int. Conf. Pattern Recognition (ICPR '96)* (Vienna, Austria), vol. I, pp. 900–904, Aug. 1996.

[23] J. Weickert, *Anisotropic Diffusion in Image Processing*. Stuttgart, Germany: Teubner-Verlag, 1998.

[24] W. J. Niessen, K. L. Vincken, J. Weickert, and M. A. Viergever, "Nonlinear multiscale representations for image segmentation," *Comput. Vision Image Understanding*, vol. 66, pp. 233–245, Apr. 1997.

[25] J. Weickert, K. J. Zuiderveld, B. M. ter Haar Romeny, and W. J. Niessen, "Parallel implementations of AOS schemes: A fast way of nonlinear diffusion filtering," in *Proc. 1997 IEEE Int. Conf. Image Processing (ICIP-97,)* (Santa Barbara, CA, Oct. 26–29, 1997), vol. 3, pp. 396–399, 1997.

[26] S. Peleg, J. Naor, R. Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 518–523, 1984.

[27] E. Barth, C. Zetzsche, M. Ferraro, and I. Rentschler, "Fractal properties from 2D-curvature on multiple scales," in *Geometric Methods in Computer Vision*, B. Vemuri, Ed., vol. 2031 of *SPIE*, pp. 87–99, 1993.

[28] S.-C. Pei, C.-C. Tseng, and C.-Y. Lin, "Wavelet transform and scale space filtering of fractal images," *IEEE Trans. Image Processing*, vol. 4, pp. 682–687, May 1995.

[29] H. Krim and D. H. Brooks, "Feature-based segmentation of ECG signals," in *IEEE Int. Symp. Time Frequency/Scale Analysis* (Paris, France, 1996).

[30] J. Sporring and J. Weickert, "On generalized entropies and scale-space," in *Scale-Space Theory in Computer Vision, Proc. 1st Int. Conf.* (Utrecht, The Netherlands), *Lecture Notes in Computer Science*, vol. 1252. Berlin, Germany: Springer-Verlag, July 1997, pp. 53–64.

[31] J. S. Simonoff, *Smoothing Methods in Statistics* (Springer Series in Statistics). Berlin, Germany: Springer-Verlag, 1996.

[32] L. Alvarez, F. Guichard, P.-L. Lions, and J.-M. Morel, "Axioms and fundamental equations of image processing," *Arch. Ration. Mech. Anal.*, vol. 123, no. 3, pp. 199–257, Sept. 1993.

[33] L. Florack, *Image Structure, Computational Imaging and Vision.* Dordrecht, The Netherlands: Kluwer, 1997.

[34] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 629–639, 1990.

[35] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Two deterministic half-quadratic regularization algorithms for computed imaging," in *Proc. 1st Int. Conf. Image Processing*, vol. 2 (Austin, TX, Nov. 13–16, 1994), pp. 168–172.

[36] H. G. E. Hentschel and I. Procaccia, "The infinite number of generated dimensions of fractals and strange attractors," *Physical 8D*, vol. 8, pp. 435–444, 1983.

[37] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. Image Processing*, vol. 7, pp. 398–410, Mar. 1998.

[38] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, "Scale and the differential structure of images," *Image Vision Comput.*, vol. 10, no. 6, pp. 376–388, July/Aug. 1992.

[39] F. G. Jensen, J. Sporring, M. Nielsen, and P. G. Sørensen, "Analyzing the dynamics of target and spiral waves by image processing techniques," Tech. Rep. DIKU-98/16, Dep. Comput. Sci., Univ. Copenhagen, Denmark, 1998.

[40] A. L. Yuille and T. A. Poggio, "Scaling theorems for zero crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 15–25, 1986.

[41] R. Picard, C. Graczyk, S. Mann, J. Wachman, L. Picard, and L. Campbell, "Vistex," 1995. [Online]. Available FTP: ftp:whitechapel.media.mit.edu. Copyright 1995, Mass. Inst. Technol.

[42] T. C. Hasley, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, "Fractal measures and their singularities: The characterization of strange sets," *Phys. Rev. A*, vol. 33, pp. 1141–1151, 1986.

[43] P. Turán, Ed., *Selected Papers of Alfréd Rényi.* Budapest, Hungary: Akadémiai Kiadó, 1976.