

# Scale-Space for N-Dimensional Discrete Signals\*

*Tony Lindeberg*

Computational Vision and Active Perception Laboratory (CVAP)  
Department of Numerical Analysis and Computing Science  
Royal Institute of Technology, S-100 44 Stockholm, Sweden  
Email: tony@bion.kth.se

*In: Y. O. Ying, A. Toet, D. Foster, H. Heijmanns and P. Meer (eds.) (1994) Shape in Picture: Mathematical Description of Shape in Grey-Level Images, (Proc. of workshop in Driebergen, Netherlands, Sep. 7-11, 1992). NATO ASI Series F, vol. 126, Springer-Verlag, pp. 571-590.*

**Abstract.** This article shows how a (linear) scale-space representation can be defined for discrete signals of arbitrary dimension. The treatment is based upon the assumptions that (i) the scale-space representation should be defined by convolving the original signal with a one-parameter family of symmetric smoothing kernels possessing a semi-group property, and (ii) local extrema must not be enhanced when the scale parameter is increased continuously.

It is shown that given these requirements the scale-space representation must satisfy the differential equation  $\partial_t L = \mathcal{A}_{ScSp} L$  for some linear and shift invariant operator  $\mathcal{A}_{ScSp}$  satisfying locality, positivity, zero sum, and symmetry conditions. Examples in one, two, and three dimensions illustrate that this corresponds to natural semi-discretizations of the continuous (second-order) diffusion equation using different discrete approximations of the Laplacean operator. In a special case the multi-dimensional representation is given by convolution with the one-dimensional discrete analogue of the Gaussian kernel along each dimension.

**Keywords:** scale, scale-space, diffusion, Gaussian smoothing, multi-scale representation, wavelets, image structure, causality

## 1 Introduction

Image structures are intrinsically of a multi-scale nature. Objects in the world and, hence, image features only exist as meaningful entities over certain ranges of scale. The idea behind a *scale-space* is to explicitly cope with this inherent property of measured data, by embedding a given signal into a family of gradually smoothed and simplified signals, in which the fine scale information is successively suppressed. Each member of the scale-space family should be associated with a specific value of a so-called *scale parameter*, somehow describing the current level of scale. A natural requirement of such an embedding is that features at coarser scales should correspond to (abstractions of) features at finer scales — they should not be just accidental phenomena created by the smoothing method.

---

\* The support from the Swedish National Board for Industrial and Technical Development, NUTEK, is gratefully acknowledged.

This property has been formalized in different ways by different authors. When Witkin [21] introduced the term “scale-space” he observed a decreasing number of zero-crossings when subjecting a signal to Gaussian smoothing. Koenderink [9] showed that natural constraints; causality, homogeneity, and isotropy, necessarily imply that the scale-space of a two-dimensional signal must satisfy the diffusion equation. Other formulations were given by Yuille and Poggio [22], regarding the zero-crossings of the Laplacean, Babaud *et al.* [2], and Lindeberg [12] who combined a decreasing number of local extrema with a semi-group structure on the smoothing transformation. Recently, Florack *et al.* [6] showed that the uniqueness of the Gaussian kernel for scale-space representation can be derived under weaker assumptions by imposing scale invariance on a semi-group of convolution kernels.

From the similarity of these results it can by now be regarded as well established that within the class of linear transformations the natural way to construct a scale-space  $L : \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  of a *continuous* signal  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  is by convolution with the Gaussian kernel

$$L(\cdot; t) = g(\cdot; t) * f(\cdot) \quad , \quad (1)$$

where

$$g(x; t) = \frac{1}{\sqrt{2\pi t}^N} e^{-(x_1^2 + \dots + x_N^2)/2t} \quad , \quad (2)$$

or equivalently, by solving the diffusion equation

$$\partial_t L = \frac{1}{2} \nabla^2 L \quad (3)$$

with initial condition  $L(\cdot; 0) = f$ . In contrast to many other multi-scale representations like pyramids (see, e.g., Burt [3]) or orthogonal wavelets (see, e.g., Mallat [17]), structures in the scale-space representation can be easily related across scales, since it is described by a differential equation (see, e.g., [14, 16]).

When applying scale-space theory in practice it should, however, be noted that real-life signals from standard detectors are *discrete*. The subject of this paper is to develop how scale-space theory can be discretized while still maintaining the scale-space properties exactly.

## 2 Scale-Space Theory for 1-D Discrete Signals

For one-dimensional signals it is possible to develop a complete discrete theory based on the assumption that the number of local extrema in a signal must not increase with scale. Below, are briefly summarized some of the main results from earlier work on this [12, 13]. The hasty reader may proceed directly to Sec. 3, where higher-dimensional signals are treated.

**Definition 1 Discrete scale-space kernel (1-D).** A kernel  $K : \mathbb{Z} \rightarrow \mathbb{R}$  is said to be a scale-space kernel if for any signal  $f_{in} : \mathbb{Z} \rightarrow \mathbb{R}$  the number of local extrema in  $f_{out} = K * f_{in}$  does not exceed the number of local extrema in  $f_{in}$ .

Using classical results (mainly by Edrei and Schoenberg; see Karlin [8]) it is possible to completely classify those kernels that satisfy this definition.

**Theorem 2 Classification of discrete scale-space kernels (1-D).** *A kernel  $K : \mathbb{Z} \rightarrow \mathbb{R}$  is a scale-space kernel if and only if its generating function  $\varphi_K(z) = \sum_{n=-\infty}^{\infty} K(n)z^n$  is of the form*

$$\varphi_K(z) = c z^k e^{(q_{-1}z^{-1} + q_1z)} \prod_{i=1}^{\infty} \frac{(1 + \alpha_i z)(1 + \delta_i z^{-1})}{(1 - \beta_i z)(1 - \gamma_i z^{-1})} ,$$

$$c > 0; \quad k \in \mathbb{Z}; \quad q_{-1}, q_1, \alpha_i, \beta_i, \gamma_i, \delta_i \geq 0 \quad \beta_i, \gamma_i < 1; \quad \sum_{i=1}^{\infty} (\alpha_i + \beta_i + \gamma_i + \delta_i) < \infty.$$

The interpretation of this result is that discrete scale-space kernels obey the following decomposition property:

**Corollary 3 Primitive discrete smoothing transformations (1-D).** *For discrete signals  $\mathbb{Z} \rightarrow \mathbb{R}$  there are five primitive types of linear and shift-invariant smoothing transformations, of which the last two are trivial;*

- two-point weighted average or generalized binomial smoothing

$$\begin{aligned} f_{out}(x) &= f_{in}(x) + \alpha_i f_{in}(x-1) & (\alpha \geq 0), \\ f_{out}(x) &= f_{in}(x) + \delta_i f_{in}(x+1) & (\delta_i \geq 0), \end{aligned}$$

- moving average or first order recursive filtering

$$\begin{aligned} f_{out}(x) &= f_{in}(x) + \beta_i f_{out}(x-1) & (0 \leq \beta_i < 1), \\ f_{out}(x) &= f_{in}(x) + \gamma_i f_{out}(x+1) & (0 \leq \gamma_i < 1), \end{aligned}$$

- infinitesimal smoothing or diffusion smoothing (see Theorem 4 for an example),
- rescaling, and
- translation.

It follows that a discrete kernel is a scale-space kernel if and only if it can be decomposed into the above primitive transformations. Moreover, the only non-trivial smoothing kernels of *finite support* arise from binomial smoothing.

If Definition 1 is combined with a requirement that the family of smoothing transformations must possess a semi-group property and have a *continuous* scale parameter, then the result is that there is in principle only one way to construct a scale-space for discrete signals.

**Theorem 4 Scale-space for discrete signals; Necessity and sufficiency.**

*Given any signal  $f : \mathbb{Z} \rightarrow \mathbb{R}$ , let  $L : \mathbb{Z} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be a one-parameter family of functions defined by  $L(x; 0) = f(x)$  ( $x \in \mathbb{Z}$ ) and*

$$L(x; t) = \sum_{n=-\infty}^{\infty} T(n; t) f(x-n) , \quad (4)$$

( $x \in \mathbb{Z}$ ,  $t > 0$ ), where  $T : \mathbb{Z} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is a one-parameter family of symmetric functions satisfying the semi-group property  $T(\cdot; s) * T(\cdot; t) = T(\cdot; s+t)$  and the normalization criterion  $\sum_{n=-\infty}^{\infty} T(n; t) = 1$ . For any signal  $f$  and any  $t_2 > t_1$  it is required that the number of local extrema (zero-crossings) in  $L(x; t_2)$  must not exceed the number of local extrema (zero-crossings) in  $L(x; t_1)$ . Then, necessarily (and sufficiently),

$$T(n; t) = e^{-\alpha t} I_n(\alpha t) \quad (5)$$

for some non-negative real  $\alpha$ , where  $I_n$  are the modified Bessel functions of integer order. This kernel  $T$  is called the discrete analogue of the Gaussian kernel.

Similar arguments in the continuous case uniquely lead to the Gaussian kernel.

The term “diffusion smoothing” can be understood by noting that the scale-space family  $L$  satisfies a semi-discretized version of the diffusion equation:

**Theorem 5 Diffusion formulation of the discrete scale-space.** *The representation  $L : \mathbb{Z} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  given by (4) with  $T : \mathbb{Z} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  according to (5) and  $\alpha = 1$  satisfies the system of ordinary differential equations*

$$\partial_t L(x; t) = \frac{1}{2}(L(x+1; t) - 2L(x; t) + L(x-1; t)) = \frac{1}{2}(\nabla_2^2 L)(x; t) \quad (6)$$

with initial condition  $L(x; 0) = f(x)$  for any discrete signal  $f : \mathbb{Z} \rightarrow \mathbb{R}$  in  $l_1$ .

Despite the completeness of these results, they cannot be extended directly to higher dimensions, since in two (and higher) dimensions there are no non-trivial kernels guaranteed to never increase the number of local extrema in a signal. One example of this, originally due to Lifshitz and Pizer [11], can be found in [12] (see also Yuille [23]). Anyway, an important point about this study, is that it gives a deep understanding of what one-dimensional linear transformations can be regarded as smoothing transformations. It also shows that the only reasonable way to convert the one-dimensional scale-space theory from continuous signals to discrete signals is by discretizing the diffusion equation.

### 3 Selecting Scale-Space Axioms in Higher Dimensions

Koenderink [9] derives the scale-space for two-dimensional continuous images from three assumptions; causality, homogeneity, and isotropy. The main idea is that it should be possible to trace every greylevel at a coarse scale to a corresponding greylevel at a finer scale. In other words, no new level curves should be created when the scale parameter increases. Using differential geometry he shows that these requirements uniquely lead to the diffusion equation.

It is of course impossible to apply these ideas directly in the discrete case, since there are no direct correspondences to level curves or differential geometry for discrete signals. Neither can the scaling argument by Florack *et al.* [6] be carried out in a discrete situation. An alternative way of expressing the first property, however, is by requiring that *if for some scale level  $t_0$  a point  $x_0$  is a local maximum for the scale-space representation at that level (regarded as a*

function of the space coordinates only) then its value must not increase when the scale parameter increases. Analogously, if a point is a local minimum then its value must not decrease when the scale parameter increases.

It is clear that this formulation is equivalent to the formulation in terms of level curves for continuous images, since if the greylevel value at a local maximum (minimum) would increase (decrease) then a new level curve would be created. Conversely, if a new level curve is created then some local maximum (minimum) must have increased (decreased). An intuitive description of this requirement is that it prevents local extrema from being enhanced and from “popping up out of nowhere”. In fact, this is closely related to the maximum principle for parabolic differential equations (see, e.g., Widder [20]).

In next section it will be shown that this condition combined with a continuous scale parameter means a strong restriction on the smoothing method also in the discrete case, and again it will lead to a discretized version of the diffusion equation. In a special case, the scale-space representation will be reduced to the family of functions generated by separated convolution with the discrete analogue of the Gaussian kernel,  $T(n; t)$ .

### 3.1 Basic Definitions

Given a point  $x \in \mathbb{Z}^N$  denote its neighbourhood of connected points by  $N(x) = \{\xi \in \mathbb{Z}^N : (\|x - \xi\|_\infty \leq 1) \wedge (\xi \neq x)\}$  (corresponding to what is known as eight-connectivity in the two-dimensional case). The corresponding set including the central point  $x$  is written  $N_+(x)$ . Define (weak) extremum points as follows:

**Definition 6 Discrete local maximum.** A point  $x \in \mathbb{Z}^N$  is said to be a (weak) local maximum of a function  $g : \mathbb{Z}^N \rightarrow \mathbb{R}$  if  $g(x) \geq g(\xi)$  for all  $\xi \in N(x)$ .

**Definition 7 Discrete local minimum.** A point  $x \in \mathbb{Z}^N$  is said to be a (weak) local minimum of a function  $g : \mathbb{Z}^N \rightarrow \mathbb{R}$  if  $g(x) \leq g(\xi)$  for all  $\xi \in N(x)$ .

The following operators are natural discrete correspondences to the Laplacean operator  $\nabla^2$  in one ( $\nabla_3^2$ ), two ( $\nabla_5^2, \nabla_{\times 2}^2$ ) and three ( $\nabla_7^2, \nabla_{+3}^2, \nabla_{\times 3}^2$ ) dimensions respectively (below the notation  $f_{-1,0,1}$  stands for  $f(x-1, y, z+1)$  etc.):

$$\begin{aligned} (\nabla_3^2 f)_0 &= f_{-1} - 2f_0 + f_1, \\ (\nabla_5^2 f)_{0,0} &= f_{-1,0} + f_{+1,0} + f_{0,-1} + f_{0,+1} - 4f_{0,0}, \\ (\nabla_{\times 2}^2 f)_{0,0} &= 1/2(f_{-1,-1} + f_{-1,+1} + f_{+1,-1} + f_{+1,+1} - 4f_{0,0}), \\ (\nabla_7^2 f)_{0,0,0} &= f_{-1,0,0} + f_{+1,0,0} + f_{0,-1,0} + f_{0,+1,0} + f_{0,0,-1} + f_{0,0,+1} - 6f_{0,0,0}, \\ (\nabla_{+3}^2 f)_{0,0,0} &= 1/4 \quad (f_{-1,-1,0} + f_{-1,+1,0} + f_{+1,-1,0} + f_{+1,+1,0} + \\ &\quad f_{-1,0,-1} + f_{-1,0,+1} + f_{+1,0,-1} + f_{+1,0,+1} + \\ &\quad f_{0,-1,-1} + f_{0,-1,+1} + f_{0,+1,-1} + f_{0,+1,+1} - 12f_{0,0,0}), \\ (\nabla_{\times 3}^2 f)_{0,0,0} &= 1/4 \quad (f_{-1,-1,-1} + f_{-1,-1,+1} + f_{-1,+1,-1} + f_{-1,+1,+1} + \\ &\quad f_{+1,-1,-1} + f_{+1,-1,+1} + f_{+1,+1,-1} + f_{+1,+1,+1} - 8f_{0,0,0}). \end{aligned}$$

## 4 Axiomatic Discrete Scale-Space Formulation

Given that the task is to state an axiomatic formulation of the first stages of visual processing, *the visual front end*, a list of desired properties may be long; *linearity, translational invariance, rotational symmetry, mirror symmetry, semi-group, causality, positivity, unimodality, continuity, differentiability, normalization* to one, nice *scaling* behaviour, *locality*, rapidly *decreasing* for large  $x$  and  $t$ , existence of an *infinitesimal generator* (explained below), and *invariance* with respect to certain *greylevel transformations*, etc. Such a list will, however, contain redundancies, as does this one. Here, a (minimal) subset of these properties will be taken as axioms. In fact, it can be shown that all the other above-mentioned properties follow from the selected subset (see also [15, 16]).

The scale-space representation for higher-dimensional signals is constructed analogously to the one-dimensional case. To start with, postulate that the scale-space should be generated by convolution with a one-parameter family of kernels, i.e.,  $L(x; 0) = f(x)$  and

$$L(x; t) = \sum_{\xi \in \mathbb{Z}^N} T(\xi; t) f(x - \xi) \quad (t > 0) . \quad (7)$$

This form of the smoothing formula corresponds to natural requirements about *linear shift-invariant smoothing* and the existence of a *continuous scale parameter*. It is natural to require that all coordinate directions should be handled identically. Therefore all kernels should be *symmetric*. Impose also a *semi-group* condition on the family  $T$ . This means that all scale levels will be treated similarly, that is, the smoothing operation does not depend on the scale value, and the transformation from a lower scale level to a higher scale level is always given by convolution with a kernel from the family:

$$\begin{aligned} L(\cdot; t_2) &= \{\text{definition}\} = T(\cdot; t_2) * f = \{\text{semi-group}\} = \\ &= (T(\cdot; t_2 - t_1) * T(\cdot; t_1)) * f = \{\text{associativity}\} = \\ &= T(\cdot; t_2 - t_1) * (T(\cdot; t_1) * f) = \{\text{definition}\} = T(\cdot; t_2 - t_1) * L(\cdot; t_1) . \end{aligned} \quad (8)$$

As smoothing criterion the *non-enhancement* requirement for local extrema is taken. It is convenient to express it as a condition of the derivative of the scale-space family with respect to the scale parameter. In order to ensure a proper statement of this condition, where differentiability is guaranteed, it is necessary to state a series of preliminary definitions leading to the desired formulation.

### 4.1 Definitions

Let us summarize this (minimal) set of basic properties, which a family should satisfy in order to be a candidate family for generating a (linear) scale-space.

**Definition 8 Pre-scale-space family of kernels.** A one-parameter family of kernels  $T : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is said to be a pre-scale-space family of kernels if it satisfies

- $T(\cdot; 0) = \delta(\cdot)$ ,
- the semi-group property  $T(\cdot; s) * T(\cdot; t) = T(\cdot; s + t)$ ,
- the symmetry properties  $T(-x_1, x_2, \dots, x_N; t) = T(x_1, x_2, \dots, x_N; t)$  and  $T(P_k^N(x_1, x_2, \dots, x_N); t) = T(x_1, x_2, \dots, x_N; t)$  for all  $x = (x_1, x_2, \dots, x_N) \in \mathbb{Z}^n$ , all  $t \in \mathbb{R}_+$ , and all possible permutations  $P_k^N$  of  $N$  elements, and
- the continuity requirement  $\|T(\cdot; t) - \delta(\cdot)\|_1 \rightarrow 0$  when  $t \downarrow 0$ .

**Definition 9 Pre-scale-space representation.** Let  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  be a discrete signal and let  $T : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be a pre-scale-space family of kernels. Then, the one-parameter family of signals  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  given by (7) is said to be the pre-scale-space representation of  $f$  generated by  $T$ .

Provided that the input signal  $f$  is sufficiently regular, these conditions on the family of kernels  $T$  guarantee that the representation  $L$  is differentiable and satisfies a system of linear differential equations.

**Lemma 10 A pre-scale-space representation is differentiable.** Let  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be the pre-scale-space representation of a signal  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  in  $l_1$ . Then  $L$  satisfies the differential equation

$$\partial_t L = \mathcal{A}L \quad (9)$$

for some linear and shift-invariant operator  $\mathcal{A}$ .

*Proof.* If  $f$  is sufficiently regular, e.g., if  $f \in l_1$ , define a family of operators  $\{\mathcal{T}_t, t > 0\}$ , here from  $l_1$  to  $l_1$ , by  $\mathcal{T}_t f = T(\cdot; t) * f$ . Due to the conditions imposed on the kernels it will satisfy the relation

$$\lim_{t \rightarrow t_0} \|(\mathcal{T}_t - \mathcal{T}_{t_0})f\|_1 = \lim_{t \rightarrow t_0} \|(\mathcal{T}_{t-t_0} - \mathcal{I})(\mathcal{T}_{t_0}f)\|_1 = 0, \quad (10)$$

where  $\mathcal{I}$  is the identity operator. Such a family is called a strongly-continuous semigroup of operators (see Hille and Phillips [7] p. 58-59). A semi-group is often characterized by its *infinitesimal generator*  $\mathcal{A}$  defined by

$$\mathcal{A}f = \lim_{h \downarrow 0} \frac{\mathcal{T}_h f - f}{h}. \quad (11)$$

The set of elements  $f$  for which  $\mathcal{A}$  exists is denoted  $\mathcal{D}(\mathcal{A})$ . This set is not empty and never reduces to the zero element. Actually, it is even dense in  $l_1$  (see Hille and Phillips [7] p. 307). If this operator exists then

$$\begin{aligned} \lim_{h \downarrow 0} \frac{L(\cdot, \cdot; t+h) - L(\cdot, \cdot; t)}{h} &= \lim_{h \downarrow 0} \frac{\mathcal{T}_{t+h} f - \mathcal{T}_t f}{h} = \\ \lim_{h \downarrow 0} \frac{\mathcal{T}_h(\mathcal{T}_t f) - (\mathcal{T}_t f)}{h} &= \mathcal{A}(\mathcal{T}_t f) = \mathcal{A}L(\cdot; t). \end{aligned} \quad (12)$$

According to a theorem by Hille and Phillips ([7] p. 308) strong continuity implies  $\partial_t(\mathcal{T}_t f) = \mathcal{A}\mathcal{T}_t f = \mathcal{T}_t \mathcal{A}f$  for all  $f \in \mathcal{D}(\mathcal{A})$ . Hence, the scale-space family  $L$  must obey the differential equation  $\partial_t L = \mathcal{A}L$  for some linear operator  $\mathcal{A}$ . Since  $L$  is generated from  $f$  by a convolution operation it follows that  $\mathcal{A}$  must be shift-invariant.  $\square$

This property makes it possible to formulate the previously indicated scale-space property in terms of derivatives of the scale-space representation with respect to the scale parameter. As in the maximum principle, the greylevel value in every local maximum point must not increase, while the greylevel value in every local minimum point must not decrease.

**Definition 11 Pre-scale-space property: Non-enhancement of extrema.**

A differentiable one-parameter family of signals  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is said to possess pre-scale-space properties, or equivalently not to enhance local extrema, if for every value of the scale parameter  $t_0 \in \mathbb{R}_+$  it holds that if  $x_0 \in \mathbb{Z}^N$  is a local extremum point for the mapping  $x \mapsto L(x; t_0)$  then the derivative of  $L$  with respect to  $t$  in this point satisfies

$$\partial_t L(x_0; t_0) \leq 0 \quad \text{if } x_0 \text{ is a local maximum point,} \quad (13)$$

$$\partial_t L(x_0; t_0) \geq 0 \quad \text{if } x_0 \text{ is a local minimum point.} \quad (14)$$

Now it can be stated that a pre-scale-space family of kernels is a scale-space family of kernels if it satisfies this property for *any* input signal.

**Definition 12 Scale-space family of kernels.** A one-parameter family of pre-scale-space kernels  $T : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is said to be a scale-space family of kernels if for any signal  $f : \mathbb{Z}^N \rightarrow \mathbb{R} \in l_1$  the pre-scale-space representation of  $f$  generated by  $T$  possesses pre-scale-space properties, i.e., if *for any* signal local extrema are never enhanced.

**Definition 13 Scale-space representation.** A pre-scale-space representation  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  of a signal  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  generated by a family of kernels  $T : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , which are scale-space kernels, is said to be a scale-space representation of  $f$ .

In the next section it will be shown how these requirements strongly restrict the possible class of kernels and scale-space representations. For example, they will lead to a number of restrictions on the operator  $\mathcal{A}$  in Lemma 10:

**Definition 14 Infinitesimal scale-space generator.** A shift-invariant linear operator  $\mathcal{A}$  from  $l_1$  to  $l_1$

$$(\mathcal{A}L)(x; t) = \sum_{\xi \in \mathbb{Z}^N} a_\xi L(x - \xi; t) , \quad (15)$$

is said to be an infinitesimal scale-space generator, denoted  $\mathcal{A}_{ScSp}$ , if the coefficients  $a_\xi \in R$  satisfy

- the locality condition  $a_\xi = 0$  if  $\xi \notin N_+(0)$ ,
- the positivity constraint  $a_\xi \geq 0$  if  $\xi \neq 0$ ,
- the zero sum condition  $\sum_{\xi \in \mathbb{Z}^N} a_\xi = 0$ , as well as
- the symmetry requirements  $a_{(-\xi_1, \xi_2, \dots, \xi_N)} = a_{(\xi_1, \xi_2, \dots, \xi_N)}$  and  $a_{P_k^N(\xi_1, \xi_2, \dots, \xi_N)} = a_{(\xi_1, \xi_2, \dots, \xi_N)}$  for all  $\xi = (\xi_1, \xi_2, \dots, \xi_N) \in \mathbb{Z}^N$  and all possible permutations  $P_k^N$  of  $N$  elements.



## 4.2 Necessity

It will first be shown that these conditions necessarily imply that the family  $L$  satisfies a semi-discretized version of the diffusion equation.

**Theorem 15 Scale-space for discrete signals: Necessity.** *A scale-space representation  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  of a signal  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  satisfies the differential equation*

$$\partial_t L = \mathcal{A}_{ScSp} L \quad (16)$$

with initial condition  $L(\cdot; 0) = f(\cdot)$  for some infinitesimal scale-space generator  $\mathcal{A}_{ScSp}$ . In one, two and three dimensions respectively (16) reduces to

$$\partial_t L = \alpha_1 \nabla_3^2 L \quad , \quad (17)$$

$$\partial_t L = \alpha_1 \nabla_5^2 L + \alpha_2 \nabla_{\times 2}^2 L \quad , \quad (18)$$

$$\partial_t L = \alpha_1 \nabla_7^2 L + \alpha_2 \nabla_{+3}^2 L + \alpha_3 \nabla_{\times 3}^2 L \quad , \quad (19)$$

for some constants  $\alpha_1 \geq 0$ ,  $\alpha_2 \geq 0$  and  $\alpha_3 \geq 0$ .

*Proof.* The proof consists of two parts. The first part has already been presented in Lemma 10, where it was shown that the requirements on the kernels imply that the family  $L$  obeys a linear differential equation. Because of the shift invariance  $\mathcal{A}L$  can be written in the form (15). In the second part counterexamples will be constructed from various simple test functions in order to delimit the class of possible operators.

The extremum point conditions (13), (14) combined with Definitions 12-13 mean that  $\mathcal{A}$  must be *local*, i.e., that  $a_\xi = 0$  if  $\xi \notin N_+(0)$ . This is easily understood by studying the following counterexample: First, assume that  $a_{\xi_0} > 0$  for some  $\xi_0 \notin N_+(0)$  and define a function  $f_1 : \mathbb{Z}^N \rightarrow \mathbb{R}$  by

$$f_1(x) = \begin{cases} \varepsilon > 0 & \text{if } x = 0, \\ 0 & \text{if } x \in N(0), \\ 1 & \text{if } x = \xi_0, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \quad (20)$$

Obviously, 0 is a local maximum point for  $f_1$ . From (9) and (15) one obtains  $\partial_t L(0; 0) = \varepsilon a_0 + a_{\xi_0}$ . It is clear that this value can be positive provided that  $\varepsilon$  is chosen small enough. Hence,  $L$  cannot satisfy (13). Similarly, it can also be shown that  $a_{\xi_0} < 0$  leads to a violation of the non-enhancement property (14) (let  $\varepsilon < 0$ ). Consequently,  $a_\xi$  must be zero if  $\xi \notin N_+(0)$ .

Moreover, the *symmetry* conditions imply that permuted and reflected coefficients must be equal, i.e.,  $a_{(-\xi_1, \xi_2, \dots, \xi_N)} = a_{(\xi_1, \xi_2, \dots, \xi_N)}$  and  $a_{P_k^N(\xi_1, \xi_2, \dots, \xi_N)} = a_{(\xi_1, \xi_2, \dots, \xi_N)}$  for all  $\xi = (\xi_1, \xi_2, \dots, \xi_N) \in \mathbb{Z}^N$  and all possible permutations  $P_k^N$  of  $N$  elements. For example, the two-dimensional version of (15) reads

$$\partial_t L = \begin{pmatrix} a & b & a \\ b & c & b \\ a & b & a \end{pmatrix} L \quad (21)$$

for some  $a$ ,  $b$  and  $c$ . Then, consider the function given by

$$f_2(x, y) = \begin{cases} 1 & \text{if } x \in N_+(0), \text{ and} \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

With the given (weak) definitions of local extremum points it is clear that 0 is both a local maximum point and a local minimum point. Hence  $\partial_t L(0; 0)$  must be zero, and the *coefficients sum to zero*  $\sum_{\xi \in \mathbb{Z}^N} a_\xi = 0$ , which in two dimensions reduces to  $4a + 4b + c = 0$  in (21). Obviously, (15) can be written

$$\partial_t L = (\mathcal{A}L)(x; t) = \sum_{\xi \in N(0)} a_\xi (L(x - \xi; t) - L(x; t)) \ , \quad (23)$$

and the two-dimensional special case (21) reduces to

$$\partial_t L = \alpha_1 \begin{pmatrix} 1 & & \\ 1 & -4 & 1 \\ & & 1 \end{pmatrix} L + \alpha_2 \begin{pmatrix} 1/2 & 1/2 \\ & -2 \\ 1/2 & 1/2 \end{pmatrix} L = \alpha_1 \nabla_5^2 L + \alpha_2 \nabla_{\times 2}^2 L \ . \quad (24)$$

Finally, by considering the test function

$$f_3(x, y) = \begin{cases} \epsilon > 0 & \text{if } x = 0, \\ -1 & \text{if } x = \tilde{\xi}, \text{ and} \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

for some  $\tilde{\xi}$  in  $N(0)$  one easily realizes that  $a_\xi$  must be *non-negative* if  $\xi \in N(0)$ . It follows that  $\alpha_1 \geq 0$  and  $\alpha_2 \geq 0$  in (24), which proves (18). (17) and (19) follow from similar straightforward considerations. The initial condition  $L(\cdot; 0) = f$  is a direct consequence of the definition of pre-scale-space kernel.  $\square$

### 4.3 Sufficiency

The reverse statement of Theorem 15 is also true.

**Theorem 16 Scale-space for discrete signals: Sufficiency.** *Let  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  be a discrete signal in  $l_1$ , let  $\mathcal{A}_{ScSp}$  be an infinitesimal scale-space generator, and let  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be the representation generated by the solution to the differential equation*

$$\partial_t L = \mathcal{A}_{ScSp} L$$

*with initial condition  $L(\cdot; 0) = f(\cdot)$ . Then,  $L$  is a scale-space representation of  $f$ .*

*Proof.* It follows almost trivially that  $L$  possesses pre-scale-space properties, i.e., that  $L$  does not enhance local extrema, if the differential equation is rewritten in the form

$$\partial_t L = (\mathcal{A}L)(x; t) = \sum_{\xi \in N(0)} a_\xi (L(x - \xi; t) - L(x; t)) \ . \quad (26)$$

If at some scale level  $t$  a point  $x$  is a local maximum point then all differences  $L(x - \xi; t) - L(x; t)$  are non-positive, which means that  $\partial_t L(x; t) \leq 0$  provided that  $a_\xi \geq 0$ . Similarly, if a point is a local minimum point then the differences are all non-negative and  $\partial_t L(x; t) \geq 0$ .

What remains to be verified is that  $L$  actually satisfies the requirements for being a pre-scale-space representation. Since  $L$  is generated by a linear differential equation, it follows that  $L$  can be written as the convolution of  $f$  with some kernel  $T$ , i.e.,  $L(\cdot; t) = T(\cdot; t) * f$ . The requirements of pre-scale-space kernels can be shown to hold by letting the input signal  $f$  be the discrete delta function. The semi-group property of the kernels follows from the fact that the coefficients  $\xi$  are constant, and the solution at a time  $s + t$  hence can be computed from the solution at an earlier time  $s$  by letting the time increase by  $t$ . The symmetry properties of the kernel are obvious from the symmetry of the differential equation. The continuity at the origin follows directly from the differentiability.  $\square$

These results show that *a one-parameter family of discrete signals is a scale-space representation if and only if it satisfies the differential equation (16) for some infinitesimal scale-space generator.*

## 5 Parameter Determination

For simplicity, from now on mainly two-dimensional signals will be considered. If (18) is rewritten in the form

$$\partial_t L = C \left( (1 - \gamma) \nabla_5^2 L + \gamma \nabla_{\times 2}^2 L \right) = C \nabla_\gamma^2 L \quad , \quad (27)$$

the interpretation of the parameter  $C$  is just a trivial rescaling of the scale parameter. Thus, without loss of generality  $C$  may be set to  $\frac{1}{2}$  in order to get the same scaling constant as in the one-dimensional case. What is left to investigate is how the remaining degree of freedom in the parameter  $\gamma \in [0, 1]$  affects the scale-space representation.

If  $\gamma = 1$  then a undesirable situation appears. Since the cross-operator only links diagonal points, the system of ordinary differential equations given by (27) can then be split into two *uncoupled* systems, one operating on the points with even coordinate sum  $x + y$  and the other operating on the points with odd coordinate sum. It is clear that this is really an unwanted behaviour, since then even after a substantial amount of “blurring”, for certain types of input signals the “smoothed” greylevel landscape may still have a rather saw-toothed shape.

### 5.1 Derivation of the Fourier Transform

Further arguments showing that  $\gamma$  must not be too large can be obtained by studying the Fourier transform of the corresponding scale-space family of kernels.

**Proposition 17 Fourier transform of the 2D discrete scale-space.** *Let  $L : \mathbb{Z}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be the scale-space representation of a discrete signal  $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$*

generated by (27) with initial condition  $L(\cdot; 0) = f(\cdot)$ . Assume that  $f \in l_1$ . Then the generating function of the kernel describing the transformation from the original signal to the representation at a certain scale  $t$  is given by

$$\varphi_T(z, w) = \sum_{(m,n) \in \mathbb{Z}^2} T(m, n; t) z^m w^n = e^{-(2-\gamma)t + \frac{(1-\gamma)}{2}(z^{-1} + z + w^{-1} + w) + \frac{\gamma}{4}(z^{-1}w^{-1} + z^{-1}w + zw^{-1} + zw)} . \quad (28)$$

Its Fourier transform is

$$\varphi_T(z, w) = \psi_T(e^{-iu}, e^{-iv}) = e^{-(2-\gamma)t + (1-\gamma)(\cos u + \cos v)t + (\gamma \cos u \cos v)t} . \quad (29)$$

*Proof.* Discretizing (27) further in scale using Euler's explicit method with scale step  $\Delta t$ , gives an iteration formula of the form

$$\begin{aligned} L_{i,j}^{k+1} &= (1 - (2 - \gamma)\Delta t) L_{i,j}^k + \\ &\quad \frac{(1 - \gamma)\Delta t}{2} (L_{i-1,j}^k + L_{i+1,j}^k + L_{i,j-1}^k + L_{i,j+1}^k) + \\ &\quad \frac{\gamma\Delta t}{4} (L_{i-1,j-1}^k + L_{i-1,j+1}^k + L_{i+1,j-1}^k + L_{i+1,j+1}^k), \end{aligned} \quad (30)$$

where the subscripts  $i$  and  $j$  denote the spatial coordinates  $x$  and  $y$  respectively, and the superscript  $k$  denotes the iteration index. The generating function describing one iteration with this transformation is

$$\begin{aligned} \varphi_{step}(z, w) &= (1 - (2 - \gamma)\Delta t) + \frac{(1 - \gamma)\Delta t}{2}(z^{-1} + z + w^{-1} + w) + \\ &\quad \frac{\gamma\Delta t}{4}(z^{-1}w^{-1} + z^{-1}w + zw^{-1} + zw) . \end{aligned} \quad (31)$$

Assume that the scale-space representation at a scale level  $t$  is computed using  $n$  iterations with a scale step  $\Delta t = \frac{t}{n}$ . Then, the generating function describing the composed transformation can be written  $\varphi_{composed,n}(z, w) = (\varphi_{step}(z, w))^n$ . After substitution of  $\Delta t$  for  $\frac{t}{n}$  and using  $\lim_{n \rightarrow \infty} (1 + \frac{\alpha_n}{n})^n = e^\alpha$  if  $\lim_{n \rightarrow \infty} \alpha_n = \alpha$ , it follows that  $\varphi_{composed,n}(z)$  tends to  $\varphi_T(z, w)$  according to (28) when  $n \rightarrow \infty$ , provided that the discretization (30) converges to the actual solution of (27).  $\square$

## 5.2 Unimodality in the Fourier Domain

It is easy to verify that the Fourier transform is unimodal if and only if  $\gamma \leq \frac{1}{2}$ .

### Proposition 18 Unimodality of the Fourier transform (2D).

The Fourier transform (29) of the kernel describing the transformation from the original signal to the smoothed representation at a coarser level of scale is unimodal if and only if  $\gamma \leq \frac{1}{2}$ .

*Proof.* Differentiation of (29) gives  $\partial_u \psi = -\psi(u, v) \sin u (1 - \gamma(1 + \cos v)) t$  and  $\partial_v \psi = -\psi(u, v) \sin v (1 - \gamma(1 + \cos u)) t$ . The Fourier transform decreases with  $|u|$  and  $|v|$  for all  $u$  and  $v$  in  $[-\pi, \pi]$  if and only if the factors  $(1 - \gamma(1 + \cos v))$  and  $(1 - \gamma(1 + \cos u))$  are non-negative for all  $u$  and  $v$ , i.e., if and only if  $\gamma \leq \frac{1}{2}$ . Then, any directional derivative away from the origin is negative.  $\square$

### 5.3 Separability

The transformation kernel is separable if and only if its Fourier transform is separable, that is, if and only if  $\psi_T(u, v)$  can be written on the form  $U_T(u)V_T(v)$  for some functions  $U_T$  and  $V_T$ . From (29) it is realized that this separation is possible if and only if  $\gamma = 0$ . Hence,

**Proposition 19 Separability of the 2D discrete scale-space.** *The convolution kernel associated with the scale-space representation defined by  $L(x, y; 0) = f(x, y)$  and*

$$\partial_t L = \frac{1}{2} ((1 - \gamma)\nabla_x^2 L + \gamma\nabla_{x^2}^2) \quad (32)$$

is separable if and only if  $\gamma = 0$ . Then  $L$  is given by

$$L(x, y; t) = \sum_{m=-\infty}^{\infty} T(m; t) \sum_{n=-\infty}^{\infty} T(n; t) f(x - m, y - n) \quad (t > 0), \quad (33)$$

where  $T(n; t) = e^{-t} I_n(t)$  and  $I_n$  are the modified Bessel functions of integer order.

*Proof.* The Fourier transform  $\psi_T(u, v)$  can be written in the form  $U_T(u)V_T(v)$  for some functions  $U_T$  and  $V_T$  if and only if the term with  $\cos u \cos v$  can be eliminated from the argument of the exponential function, i.e., if and only if  $\gamma$  is zero. In that case the Fourier transform reduces to

$$\psi_T(u, v) = e^{(-2 + \cos u + \cos v)t} = e^{(-1 + \cos u)t} e^{(-1 + \cos v)t} \quad (34)$$

which corresponds to separated smoothing with the one-dimensional discrete analogue of the Gaussian kernel along each coordinate direction.

It can also be verified directly that (33) satisfies (32). Consider the possible scale-space representation of an  $N$ -dimensional signal generated by separable convolution with the one-dimensional discrete analogue of the Gaussian kernel; i.e., given  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  define  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  by

$$L(x; t) = \sum_{x \in \mathbb{Z}^N} T_N(\xi; t) f(x - \xi) \quad (t > 0), \quad (35)$$

where  $T_N : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is given by

$$T_N(\xi; t) = \prod_{i=1}^N T_1(\xi_i; t) , \quad (36)$$

$\xi = (\xi_1, \dots, \xi_N)$ , and  $T_1 : \mathbb{Z} \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is the discrete analogue of the Gaussian kernel,  $T_1(n; t) = e^{-t} I_n(t)$ . It will be shown that this representation satisfies a semi-discretized version of the two-dimensional diffusion equation

$$\partial_t L = \frac{1}{2} \nabla_{2N+1}^2 L , \quad (37)$$

where

$$(\nabla_{2N+1}^2 L)(x; t) = \sum_{i=1}^N L(x + e_i; t) - 2L(x; t) + L(x - e_i; t) , \quad (38)$$

and  $e_i$  denotes the unit vector in the  $i$ th coordinate direction. Consider

$$(\partial_i T_N)(\xi; t) = \sum_{i=1}^N (\partial_i T_1)(\xi_i; t) \prod_{j \neq i} T_1(\xi_j; t) . \quad (39)$$

Since  $T_1$  satisfies (6), this expression can be written

$$(\partial_i T_N)(\xi; t) = \sum_{i=1}^N \frac{1}{2} (T_1(\xi_i - 1; t) - 2T_1(\xi_i; t) + T_1(\xi_i + 1; t)) \prod_{j \neq i} T_1(\xi_j; t),$$

which is obviously equivalent to

$$\partial_i T_N = \frac{1}{2} \nabla_{2N+1}^2 T_N . \quad (40)$$

The same relation holds for  $L$  provided that the differentiation and infinite summation operators commute.  $\square$

In other words, in the separable case the resulting higher-dimensional discrete scale-space corresponds to repeated application of the one-dimensional scale-space concept along each coordinate direction.

#### 5.4 Discrete Iterations

The discretization of (27) in (30) using Euler's explicit method with scale step  $\Delta t$  corresponds to iterating with a kernel with the computational molecule

$$\begin{pmatrix} \frac{\gamma \Delta t}{4} & \frac{(1-\gamma) \Delta t}{2} & \frac{\gamma \Delta t}{4} \\ \frac{(1-\gamma) \Delta t}{2} & 1 - (2-\gamma) \Delta t & \frac{(1-\gamma) \Delta t}{2} \\ \frac{\gamma \Delta t}{4} & \frac{(1-\gamma) \Delta t}{2} & \frac{\gamma \Delta t}{4} \end{pmatrix} . \quad (41)$$

Clearly, this kernel is unimodal if and only if  $\gamma \leq \frac{2}{3}$ . It is separable if and only if  $\gamma = \Delta t$  (see below). In that case, the corresponding one-dimensional kernel is a discrete scale-space kernel in the sense of Definition 1 if and only if  $\Delta t \leq \frac{1}{2}$  (see Proposition 10 in [12]). This gives a further indication that  $\gamma$  should not exceed the value  $\frac{1}{2}$ .

**Proposition 20 Separability of the iteration kernel.** *The iteration kernel (41), corresponding to discrete forward iteration with Euler's explicit method, is separable if and only if  $\gamma = \Delta t$ . In that case, the corresponding one-dimensional kernel is a discrete scale-space kernel if and only if  $0 \leq \gamma \leq 1/2$ .*

*Proof.* Since the kernel is symmetric and the coefficients sum to one, the kernel is separable if and only if it can be written as a kernel  $(a, 1 - 2a, a)$  convolved with itself, i.e., if and only if there exists an  $a \geq 0$  such that  $a^2 = \gamma\Delta t/4$ ,  $a(1 - a) = (1 - \gamma)\Delta t/2$ , and  $(1 - a)^2 = 1 - (2 - \gamma)\Delta t$ . The first equation has one non-negative root  $a = \sqrt{\gamma\Delta t}/2$ . Insertion into the second equation gives two conditions for  $\Delta t$ ; either  $\Delta t = 0$  or  $\Delta t = \gamma$ . One verifies that these roots satisfy the third equation. The kernel  $(a, 1 - 2a, a)$  is a discrete scale-space kernel if and only if  $a \leq \frac{1}{2}$  (see Equations (30) and (31) in [12]; compare also with Theorem 2).  $\square$

The boundary case  $\gamma = \Delta t = \frac{1}{2}$  gives the iteration kernel in Fig. 1(a) corresponding to separated convolution with the one-dimensional binomial kernel  $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$  frequently used in pyramid generation (see, e.g., Crowley [4]).

$$\begin{array}{cccc}
 \begin{pmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/4 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{pmatrix} & 
 \begin{pmatrix} 1/8 & 2/8 & 1/8 \\ 2/8 & -12/8 & 2/8 \\ 1/8 & 2/8 & 1/8 \end{pmatrix} & 
 \begin{pmatrix} 1/36 & 1/9 & 1/36 \\ 1/9 & 4/9 & 1/9 \\ 1/36 & 1/9 & 1/36 \end{pmatrix} & 
 \begin{pmatrix} 1/6 & 4/6 & 1/6 \\ 4/6 & -20/3 & 4/6 \\ 1/6 & 4/6 & 1/6 \end{pmatrix} \\
 \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)}
 \end{array}$$

**Fig. 1.** Computational molecules corresponding to (from left to right); (a) discrete iteration with  $\Delta t = \gamma = \frac{1}{2}$ , (b) the Laplacean operator when  $\gamma = \frac{1}{2}$ , (c) discrete iteration with  $\Delta t = \gamma = \frac{1}{3}$ , and (d) the Laplacean operator when  $\gamma = \frac{1}{3}$ .

## 5.5 Spatial Isotropy

Another aspect that might affect the selection of  $\gamma$  is spatial isotropy. It is not clear that rotational invariance is a primary quality to be aimed at in the discrete case, since then one is locked to a fixed square grid. It is also far from obvious what should be meant by spatial isotropy in a discrete situation. Possibly, it is better to talk about the lack of spatial isotropy, spatial anisotropy, or rotational asymmetry. However, since the Fourier transform is a continuous function of  $u$  and  $v$ , one can regard its variation as a function of the polar angle, given a fixed value of the radius, as one measure of this property. By expressing  $\psi_T(u, v)$  in polar coordinates  $u = \omega \cos \phi$ ,  $v = \omega \sin \phi$  and examining the resulting expression,

$$\psi_T(\omega \cos \phi, \omega \sin \phi) = e^{h(\omega \cos \phi, \omega \sin \phi)t} , \quad (42)$$

where

$$\begin{aligned}
 h(\omega \cos \phi, \omega \sin \phi) = & -(2 - \gamma) + (1 - \gamma)(\cos(\omega \cos \phi) + \cos(\omega \sin \phi)) + \\
 & \gamma \cos(\omega \cos \phi) \cos(\omega \sin \phi) , \quad (43)
 \end{aligned}$$

one realizes that the value of  $\gamma$  that gives the smallest angular variation for a fixed value of  $\omega$ , depends on  $\omega$ . Hence, with this formulation, the ‘‘rotational invariance’’ is scale dependent. At coarse scales one obtains:

**Proposition 21 Rotational invariance in the Fourier domain (2D).**

The value of  $\gamma$  that gives the least rotational asymmetry for large scale phenomena in the solution to the differential equation (27) is  $\gamma = \frac{1}{3}$ .

*Proof.* The Taylor expansion of  $h$  for small values of  $\omega$  is (see [13] Appendix A.2.3)

$$h(\omega \cos \phi, \omega \sin \phi) = -\frac{1}{2}\omega^2 + \frac{1}{24}(1 + (6\gamma - 2) \cos^2 \phi \sin^2 \phi)\omega^4 + O(\omega^6) , \quad (44)$$

where the  $O(\omega^6)$  term depends on both  $\phi$  and  $\gamma$ . Observe that if  $\gamma = \frac{1}{3}$  then the  $\phi$ -dependence decreases with  $\omega$  as  $\omega^6$  instead of as  $\omega^4$ .  $\square$

This means that  $\gamma = \frac{1}{3}$  *asymptotically*, i.e., with increasing spatial scale, gives the most isotropic smoothing effect on coarse-scale events. The reason why spatial isotropy is desired at coarse scales rather than at fine scales is because the grid effects become smaller for coarse-scale phenomena, which in turn makes it more meaningful to talk about rotational invariance. This selection of  $\gamma$  corresponds to approximating the Laplacean operator with the “the nine-point operator” (see Fig. 1(d) and Dahlquist [5]). Note that when the separability is violated by using a non-zero value of  $\gamma$ , the discrete scale-space representation can anyway be computed efficiently in the Fourier domain using (29).

## 6 Summary and Discussion

The proper way to apply the scale-space theory to discrete signals is apparently by discretizing the diffusion equation. Starting from a requirement that local extrema must not be enhanced when the scale parameter is increased continuously, it has been shown that within the class of linear transformations a necessary and sufficient condition for a one-parameter family of representations  $L : \mathbb{Z}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  to be a scale-space family of a discrete signal  $f : \mathbb{Z}^N \rightarrow \mathbb{R}$  is that it satisfies the differential equation

$$\partial_t L = \mathcal{A}_{ScSp} L , \quad (45)$$

with initial condition  $L(\cdot; 0) = f(\cdot)$  for some infinitesimal scale-space generator  $\mathcal{A}_{ScSp}$ . In one, two and three dimensions respectively it can equivalently be stated that a family is a scale-space family if and only if for some linear reparametrization of the scale parameter  $t$  and for some  $\gamma_i \in [0, 1]$  it satisfies

$$\partial_t L = \frac{1}{2} \nabla_3^2 L , \quad (46)$$

$$\partial_t L = \frac{1}{2} ((1 - \gamma_1) \nabla_5^2 L + \gamma_1 \nabla_{\times 2}^2 L) , \quad (47)$$

$$\partial_t L = \frac{1}{2} ((1 - \gamma_1 - \gamma_2) \nabla_7^2 L + \gamma_1 \nabla_{+3}^2 L + \gamma_2 \nabla_{\times 3}^2 L) . \quad (48)$$

The essence of (45)–(48) is that these equations correspond to discretizations of *first-order* differential operators in *scale*, and *second-order* differential operators in *space*.



The effect of using different values of  $\gamma_1$  in the two-dimensional case has been analyzed in detail. Nevertheless, the question about definite selection is left open. Unimodality considerations indicate that  $\gamma$  must not exceed  $\frac{1}{2}$ , while  $\gamma = \frac{1}{3}$  gives the least degree of rotational asymmetry in the Fourier domain.

The family of scale-space kernels is separable if and only if  $\gamma = 0$ . In this case the scale-space family is given by convolution with the one-dimensional discrete analogue of the Gaussian kernel along each dimension. For this parameter setting the closed-form expressions for several derived entities simplify (see, e.g., [12, 15]). Observe also that  $\gamma = 0$  arises a necessary consequence if the neighbourhood concept (defined in Sec. 3.1) is redefined as  $N(x) = \{\xi \in Z^N : (\|x_i - \xi\|_1 \leq 1) \wedge (\xi \neq x)\}$  (corresponding to what is known as four-connectivity in the two-dimensional case), since then necessarily  $\alpha_i = 0$  ( $i > 1$ ) in (18) and (19). Similar results hold in higher dimensions. A possible disadvantage with choosing  $\gamma = 0$  is that it emphasizes the role of the coordinate axes as being special directions.

Finally, it should be remarked that if a linear and shift-invariant operator  $\mathcal{L}$ , commuting with the smoothing operator  $T*$ , is applied to the scale-space representation  $L$  of a signal  $f$ , then  $\mathcal{L}L$  will be a scale-space representation of  $\mathcal{L}f$ . One consequence of this is that multi-scale discrete derivative approximations defined by linear filtering of the smoothed signal preserve the scale-space properties. This property, which provides a natural way to discretize the multi-scale N-jet representation proposed by Koenderink and van Doorn [10], is developed in [15].

## 7 Possible Extensions

The treatment so far has been restricted to signals defined on infinite and uniformly sampled square grids using uniform smoothing of all grid points. Below the ways in which these notions can be generalized are outlined.

### 7.1 Anisotropic Smoothing

Perona and Malik [19] propose anisotropic smoothing as a way to reduce the shape distortions arising in edge detection by smoothing across object boundaries (see also Nordström [18]). The suggested methodology is to modify the diffusion coefficients in order to favour intraregion smoothing over interregion smoothing.

Using the maximum principle they show that the resulting anisotropic scale-space representation possesses a suppression property for local extrema similar to that used in Koenderink's [9] continuous scale-space formulation and this discrete treatment. From the proofs of Theorems 15-16 it is obvious that the discrete scale-space concept can easily be extended to such anisotropic diffusion by letting the coefficients in the operator  $\mathcal{A}_{ScSp}$  depend upon the input signal. By this, the locality, positivity, and zero sum conditions will be preserved, while the symmetry requirements must be relaxed. Introducing such an anisotropic diffusion equation, however, violates the convolution form of smoothing as well as the semi-group property. Therefore, when proving the necessity of the representation a certain form of the smoothing formula may have to be assumed, for

example, of the form (9) with the filter coefficients depending upon the input signal. Note that, if the translational invariance and the symmetry with respect to coordinate interchanges are relaxed in (45), then this equation corresponds the (spatial) discretization of the (second-order) diffusion equation with *variable conductance*,  $c(\mathbf{x}; t)$ ,

$$(\partial_t L)(\mathbf{x}; t) = \nabla(c(\mathbf{x}; t) \nabla L(\mathbf{x}; t)) . \quad (49)$$

Throughout this work uniform smoothing has been used at the cost of possible smoothing across object boundaries. The motivation behind this choice has been the main interest in using scale-space for *detecting* image structures. Therefore, in the absence of any prior information, it is natural that the first processing steps should be as *uncommitted* as possible. The approach taken has been to first detect candidate regions of interest, and then, once candidates have been detected as regions, improve their localization. Possibly, variable conductance could be useful in the second step of this process. Another natural application is to avoid the negative effects of smoothing thin or elongated structures.

There are, however, some problems that need to be further analyzed. Modifying the diffusion coefficients requires some kind of a priori information concerning which structures in the image are to be smoothed and which are not. In the method by Perona and Malik there is a tuning function to be determined, giving the diffusion coefficient as function of the gradient magnitude. When the scale parameter  $t$  tends to infinity, the solution to the anisotropic diffusion equation tends to a function with various sharp edges. Hence, choosing a tuning function somehow implies an implicit assumption about a “final segmentation” of the image. It is not clear that such a concept exists or can be rigorously defined.

## 7.2 Finite Data

A practical problem always arising in linear filtering is what to do with pixels near the image boundary for which a part of the filter mask stretches outside the available image.

The most conservative outlook is, of course, to regard the output as undefined as soon as a computation requires image data outside the available domain. This is, however, hardly desirable for scale-space smoothing, since the (untruncated) convolution masks have infinite support, while the peripheral coefficients decrease towards zero very rapidly. A variety of ad hoc methods have been proposed to deal with this; extension methods, subtraction of steady-state components, solving the diffusion equation on a limited domain with (say, adiabatic) boundary conditions, etc. However, no such technique can overcome the problem with missing data. In some simple situations ad hoc extensions may do. But this requires some kind of a priori information about the contents of the image.

Inevitably, the peripheral image values of a smoothed finite image will be less reliable than the central ones. Instead, if accurate values really are required near the image boundary, then the vision system should try to acquire additional data such that the convolution operation becomes well-defined up to the

prescribed accuracy. This is easily achieved within the active vision paradigm simply by moving the camera so that more values become available in a sufficiently large neighbourhood of the object of interest. The task of analyzing an object manifesting itself at a certain scale requires input data in a region around the object. The width of this frame depends both on the current level of scale and the prescribed accuracy of the analysis.

Of course, a genuinely finite approach is also possible. In this presentation this subject has not been developed, since the associated problems are somehow artificial and difficult to handle in a consistent manner, although the non-enhancement property can be easily formulated for finite data and although in the one-dimensional case the concepts of sign-regularity and semi-groups of totally positive matrices [8] in principle provide possible tools for dealing with this issue. One way to avoid both the infiniteness and the boundary problems is by using a spherical camera. Then, the ordinary planar camera geometry appears as an approximate description for foveal vision, that is, small solid angles in the central field of vision.

### 7.3 Other Types of Grids

The assumption of a square grid is not a necessary restriction. The same type of treatment can be carried out on, for example, a hexagonal grid with the semi-group property preserved, and also on a grid corresponding to non-uniform spatial sampling provided that the diffusion coefficients are modified accordingly. In the latter case some a priori form of the smoothing formula may have to be adopted when proving the necessity of the representation. An interesting case to consider might actually be the non-uniformly sampled spherical camera.

### 7.4 Further Work

Finally, it should be pointed out that there is one main issue that has not been considered here, namely *scale-dependent spatial sampling*. This issue is certainly of importance in order to improve the computational efficiency both when computing the representation and for algorithms working on the data. The scale-space concept outlined here uses the same spatial resolution at all levels of scale. The pyramid representations (see, e.g., Burt [3]) on the other hand imply a fixed relation between scale and resolution beyond which refinements are not possible.

Since the smoothed images at coarser scales become progressively more redundant, it seems plausible that some kind of subsampling can be done at the coarser scales without too much loss of information. It would be interesting to carry out an analysis about how much information is lost by such an operation, and to which extent a subsampling operator can be introduced in this representation, while still maintaining the theoretical properties associated with having a continuous scale parameter, and without introducing any severe discontinuities along the scale direction that would be a potential source to numerical difficulties for algorithms working on the output from the representation.

## References

1. Abramowitz, M., Stegun, I.A. (1964). Handbook of Mathematical Functions, Applied Mathematics Series, **55**, National Bureau of Standards.
2. Babaud, J., Witkin, A.P., Baudin, M., Duda, R.O. (1986). Uniqueness of the Gaussian kernel for scale-space filtering, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 8:1, pp. 26–33.
3. Burt, P.J., Adelson, E.H. (1983). The Laplacian pyramid as a compact image code, *IEEE Trans. Comm.*, Vol. 9:4, pp. 532–540.
4. Crowley, J.L., Stern, R.M. (1984). Fast computation of the difference of low pass transform, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 6, pp. 212–222.
5. Dahlquist, G., Björk, Å., Anderson, N. (1974). Numerical Methods, Prentice-Hall.
6. Florack, L.M.J., ter Haar Romeny, B.M., Koenderink, J.J., Viergever, M.A. (1992). Scale-space: its natural operators and differential invariants, *Image and Vision Computing*, Vol. 10:6, pp. 376–388.
7. Hille, E., Phillips, R.S. (1957). Functional Analysis and Semi-Groups, *Am. Math. Soc. Coll. Publ.*, Vol. XXXI.
8. Karlin, S. (1968). Total Positivity, Vol.I, Stanford Univ. Press.
9. Koenderink, J.J. (1984). The structure of images, *Biol. Cyb.*, Vol. 50, pp. 363–370.
10. Koenderink, J.J., van Doorn, A.J. (1990). Receptive field families, *Biol. Cyb.*, Vol. 63, pp. 291–297.
11. Lifshitz, L.M., Pizer, S.M. (1987). A multiresolution hierarchical approach to image segmentation based on intensity extrema, Technical report, Depts. Comp. Sci. and Radiology, Univ. North Carolina, Chapel Hill, N.C., U.S.A.
12. Lindeberg, T.P. (1990). Scale-space for discrete signals, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 12:3, pp. 234–254.
13. Lindeberg, T.P. (1991). Discrete Scale-Space Theory and the Scale-Space Primal Sketch, Ph.D. Thesis, ISRN KTH/NA/P-91/8-SE, Dept. Num. Anal. Comp. Sci., Royal Inst. Tech., S-100 44 Stockholm, Sweden. A revised and extended version to appear in *The Kluwer International Series in Engineering and Computer Science*.
14. Lindeberg, T.P. (1992). Scale-space behaviour of local extrema and blobs, *J. Math. Imaging Vision*, Vol. 1, pp. 65–99.
15. Lindeberg, T.P. (1993). Discrete derivative approximations with scale-space properties, *J. Math. Imaging and Vision*, to appear.
16. Lindeberg T.P. (1993) Scale-space behaviour and invariance properties of differential singularities. In: This volume.
17. Mallat, S.G. (1989). A theory of multiresolution signal processing: The wavelet representation, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 11:6, pp. 674–693.
18. Nordström, N. (1990). Biased anisotropic diffusion — A unified regularization and diffusion approach to edge detection. In: Faugeras, O. (ed.), *Proc 1st Eur. Conf. Comp. Vision*, Antibes, France, Apr. 23-37, pp. 18–27, Springer-Verlag.
19. Perona, P., Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 12:7, pp. 629–639.
20. Widder, D.V. (1975). The Heat Equation, Academic Press, New York.
21. Witkin, A.P. (1983). Scale-space filtering, In: *Proc. 8th Int. Joint Conf. Art. Intell.*, Karlsruhe, Germany, Aug. 8-12, pp. 1019–1022.
22. Yuille, A., Poggio, T. (1986). Scaling theorems for zero-crossings, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol. 9:1, pp. 15–25.
23. Yuille, A.L. (1988). The creation of structure in dynamic shape, In: *Proc 2nd Int. Conf. Comp. Vision*, Tampa, Florida, Dec. 5-8, pp. 685–689.