Generalized Gaussian Scale-Space Axiomatics Comprising Linear Scale-Space, Affine Scale-Space and Spatio-Temporal Scale-Space

Tony Lindeberg

Abstract This paper describes a generalized axiomatic scale-space theory that makes it possible to derive the notions of linear scale-space, affine Gaussian scale-space and linear spatio-temporal scale-space using a similar set of assumptions (scale-space axioms).

The notion of non-enhancement of local extrema is generalized from previous application over discrete and rotationally symmetric kernels to continuous and more general non-isotropic kernels over both spatial and spatio-temporal image domains. It is shown how a complete classification can be given of the linear (Gaussian) scale-space concepts that satisfy these conditions on isotropic spatial, non-isotropic spatial and spatio-temporal domains, which results in a general taxonomy of Gaussian scale-spaces for continuous image data. The resulting theory allows filter shapes to be tuned from specific context information and provides a theoretical foundation for the recently exploited mechanisms of shape adaptation and velocity adaptation, with highly useful applications in computer vision.

It is also shown how time-causal spatio-temporal scale-spaces can be derived from similar assumptions. The mathematical structure of these scale-spaces is analyzed in detail concerning transformation properties over space and time, the temporal cascade structure they satisfy over time as well as properties of the resulting multi-scale spatio-temporal derivative operators. It is also shown how temporal deriv-
atives with respect to transformed time can be defined, leading to the formulation of a novel analogue of scale normalized derivatives for time-causal scale-spaces.

The kernels generated from these two types of theories have interesting relations to biological vision. We show how filter kernels generated from the Gaussian spatio-temporal scale-space as well as the time-causal spatio-temporal scale-space relate to spatio-temporal receptive field profiles registered from mammalian vision. Specifically, we show that there are close analogies to space-time separable cells in the LGN as well as to both space-time separable and non-separable cells in the striate cortex. We do also present a set of plausible models for complex cells using extended quasi-quadrature measures expressed in terms of scale normalized spatio-temporal derivatives.

The theories presented as well as their relations to biological vision show that it is possible to describe a general set of Gaussian and/or time-causal scale-spaces using a unified framework, which generalizes and complements previously presented scale-space formulations in this area.

Keywords Scale-space · Multi-scale representation · Scale-space axioms · Non-enhancement of local extrema · Causality · Scale invariance · Gaussian kernel · Gaussian derivative · Spatio-temporal · Affine · Spatial · Temporal · Time-recursive · Receptive field · Diffusion · Computer vision · Image processing

T. Lindeberg (✉)
School of Computer Science and Communication, KTH (Royal Institute of Technology), 100 44 Stockholm, Sweden
e-mail: tony@csc.kth.se

1 Introduction

When analyzing sensory data, such as images, a fundamental issue arises from the fact that real-world objects may appear in different ways depending upon the scale of observation. This insight is a major motivation for the development

of multi-scale representations such as *pyramids* [8, 10] and *scale-space representation* [20, 31, 32, 35, 39, 40, 65, 67, 73]. Indeed by studying the problem of how to construct a multi-scale representation, a general and multi-purpose theory for early visual operations can be stated, where the Gaussian kernel and its derivatives arise as a canonical family of image operators given natural requirements of a *visual front-end*. Complementary works have demonstrated that these filters can serve as a basis for expressing a large number of visual operations, including feature detection, stereo matching, computation of optic flow, tracking, estimation of shape cues and view-based object recognition [47].

Traditionally, however, most works on multi-scale representations have been concerned with image data defined on *isotropic spatial domains*, characterized by the fact that image data are accessible in all directions and moreover that all directions are equally treated. During recent years, these ideas have been extended to non-isotropic spatial domains in terms of *affine Gaussian scale-space*, where different amounts of smoothing may be performed in different directions, for example to account for the linear (affine) component of perspective deformations [39, 51], as well as to *spatio-temporal scale-space*, where space and time are intrinsically different dimensions and in addition Galilean motions may occur [43]. In particular, when dealing with temporal and/or spatio-temporal data in an on-line situation, we have to accept the fact that we cannot allow filters to extend into the future. For this reason, the smoothing filters have to be *time-causal* [33, 49]. The subject of this article is to show how a previously stated scale-space formulation in terms of *non-enhancement of local extrema* [38, 41] can be used for deriving Gaussian scale-spaces over non-isotropic as well as time-casual domains in an axiomatic way, by relaxing the requirement of rotationally symmetric filters that has been used in previous applications of non-enhancement of local extrema as a scale-space axiom.

It will be shown that a parameterized family of linear scale-spaces is obtained by this construction, including (i) the traditional linear scale-space representation for rotationally symmetric spatial domains, (ii) the affine Gaussian scale-space for non-isotropic spatial image domains that may be subject to local or global affine image deformations and (iii) spatio-temporal linear scale-space for spatio-temporal image domains subject to local Galilean motions. Spatio-temporal scale-spaces will be developed for both non-causal and time-causal spatio-temporal domains. Compared to the isotropic linear scale-space, these affine and spatio-temporal scale-spaces give rise to non-separable and elongated filter kernels over space and time, which allow for *shape adaptation* in space [5, 6, 51, 55, 61, 69] and *velocity adaptation* along the direction of motion [36, 37, 43, 52, 56]. Notably the receptive field profiles generated by this scale-space concept have high qualitative similarity to receptive

fields profiles recorded from biological vision [11, 70] in analogy with previously established relations between spatial receptive fields and Gaussian derivative operators [74, 75] with extensions to spatio-temporal data in [76, 77].

1.1 Outline of the Presentation

This paper is organized as follows: Section 2 gives a review of related work with emphasis on axiomatic derivations of linear scale-spaces. Section 3 defines the notion of non-enhancement of local extrema and gives formal proofs showing, by both necessity and sufficiency, how this requirement in combination with a semi-group structure implies both existence and a restriction on a corresponding infinitesimal generator. Specifically, it is shown how this algebraic structure implies that the scale-space has to satisfy a family of parabolic differential equations, where the second-order term is determined by a positive semi-definite covariance matrix. The consequences of this result with regard to different types of image domains are developed in Sect. 4 with emphasis on either rotationally symmetric, affine spatially anisotropic or non-causal spatio-temporal domains.

Up to this point, we have throughout assumed non-causal temporal image data, which is relevant for processing static spatial domains or spatio-temporal data in off-line situations. Then, to be able to handle time-causal data in on-line scenarios, where information about the future is not available, Sect. 5 shows how the previously stated scale-space formulation can be reformulated to handle time-causal data, where we do only have access to the past.

Interestingly, there are close relations between these scale-space theories and biological vision. Section 6 shows how the multi-scale spatio-temporal derivative operators from the two types of spatio-temporal scale-spaces relate to receptive fields registered in mammalian vision. Finally, Sect. 7 concludes with a summary and discussion about some of the main results.

Since the mathematical structure of the resulting time-causal scale-space turns out to be more complicated than for the Gaussian scale-space, we find it important to analyze and describe its properties in more detail. Therefore, Appendix D includes a brief review of closely related theory for heat conduction in solids, which is then transferred to spatio-temporal scale-space.

To simplify the flow through the presentation, we have also put some other more technical material in appendix sections. Appendix A develops the specific regularity properties of the scale-space kernels over scales that are used when deriving formal necessity results. Appendix B describes relations between the proposed notion of non-enhancement of local extrema and the maximum principle. Appendix C gives a theory for how Galilean invariant fixed points can be constructed in a spatio-temporal scale-space representation as a way of interpreting the output from an ensemble

of velocity-adapted scale-space filters. Finally, Appendix E gives the formal definitions and proofs regarding the axiomatic derivation of the time-causal spatio-temporal scale-space.

2 Related Work on Axiomatic Scale-Space Formulations

The Gaussian scale-space concept satisfies a number of useful properties[1] that make it particularly attractive for generating a scale-space representation; (i) linearity, (ii) shift invariance, (iii) semi-group property, (iv) existence of an infinitesimal generator, (v) non-creation of local extrema or zero-crossings in the one-dimensional case, (vi) non-enhancement of local extrema in any number of dimensions, (vii) rotational symmetry, (viii) positivity, (ix) unit normalization, and (x) scale invariance. In fact, it can be shown that the Gaussian kernel *by necessity* is a *unique* choice for a number of different combinations of subsets of these scale-space axioms [4, 20, 32, 38, 39, 41, 47, 57, 67, 72, 78]

$$g(x; s) = \frac{1}{(2\pi s)^{N/2}} e^{-(x_1^2 + \cdots + x_N^2)/2s}. \tag{1}$$

The Gaussian function is also special in the respect that it (xi) minimizes the uncertainty relation and (xii) is the probability density function with maximum entropy. The maximum entropy result can be interpreted as the Gaussian kernel making minimal use of information. These properties are also desirable when constructing a scale-space representation, since the uncertainty relation makes the smoothing operation well localized over space and scales, while the maximum entropy result means that the Gaussian kernel is maximally *uncommitted*.

When Witkin [73] coined the term "scale-space", he was concerned with one-dimensional signals and observed that new local extrema cannot be created under Gaussian convolutions. Specifically, he applied this property to zero-crossings of the second-order derivative to construct so-called "fingerprints". This observation shows that Gaussian convolution satisfies certain sufficiency results for being a smoothing operation. The first proof in the Western literature of the necessity of Gaussian smoothing for generating a scale-space was given by Koenderink [32], who also gave a formal extension of the scale-space theory to higher dimensions. He introduced the concept of *causality*, which means that new level surfaces must not be created in the scale-space representation when the scale parameter is increased. By combining causality with the notions of *isotropy* and *homogeneity*, which essentially mean that all spatial positions

and all scale levels must be treated in a similar manner, he showed that the scale-space representation must satisfy the diffusion equation

$$\partial_t L = \frac{1}{2} \nabla^2 L. \tag{2}$$

Related necessity results were given by Babaud et al. [4] and by Yuille and Poggio [78].

Lindeberg [38] considered the problem of characterizing those kernels in one dimension that share the property of not introducing new local extrema or new zero-crossings in a signal under convolution. Such scale-space kernels can be completely classified using classical results by Schoenberg [62, 63]. For continuous signals, it can be shown that all such non-trivial scale-space kernels can be decomposed into Gaussian kernels and truncated exponential functions. By imposing a *semi-group* structure on scale-space kernels, the Gaussian kernels will then be singled out as a unique choice. For discrete signals, the corresponding result is that all discrete scale-space kernels can be decomposed into generalized binomial smoothing, moving average or first-order recursive filtering and infinitesimal smoothing with the discrete analogue of the Gaussian kernel. To express a corresponding theory for higher-dimensional signals, Lindeberg [38] reformulated Koenderink's causality requirement into *non-enhancement of local extrema* and combined this requirement with a semi-group structure as well as an *infinitesimal generator* and showed that all such discrete scale-spaces must satisfy semi-discrete diffusion equations. A corresponding scale-space formulation for continuous signals based on non-enhancement of local extrema for rotationally symmetric smoothing kernels was presented in [41].

A formulation by Florack et al. [21] with continued work by Pauwels et al. [57] shows that the class of allowable scale-space kernels can also be restricted by combining a semi-group structure of convolution operations with *scale invariance* and rotational symmetry. When Florack et al. [21] studied this approach, they used separability in Cartesian coordinates as an additional constraint and showed that this lead to the Gaussian kernel. If the requirement about separability on the other hand is relaxed, Pauwels et al. [57] showed that this leads to a one-parameter family of scale-spaces, with Fourier transforms of the form

$$\hat{h}(\omega; s) = e^{-\alpha|\sigma\omega|^p}, \tag{3}$$

where $\sigma = \sqrt{s}$. Within this class, it can furthermore be shown that only the exponents p that are even integers lead to differential equations that have local infinitesimal generators of a classical form.[2] Specifically, out of this countable

---

[1]Please, refer to Lindeberg [48] for detailed mathematical definitions of these scale-space properties.

[2]Of the form stated in (11) and made more precise in Sect. 3.

set in turn, only the choice $p = 2$ gives rise to a non-negative convolution kernel, which leads to the Gaussian kernel.

There are, however, also possibilities of defining scale-space representations for other values of p. The specific case with $p = 1$ has been studied by Felsberg and Sommer [18], who show that the corresponding scale-space representation is in the two-dimensional case given by convolution with Poisson kernels of the form

$$P(x; s) = \frac{s}{2\pi((\frac{s}{2})^2 + |x|^2)^{3/2}}. \tag{4}$$

Duits et al. [12, 13] have investigated the cases with other non-integer values of p in the range $]0, 2[$ and showed that such families of self-similar α-scale-spaces (with $\alpha = p/2$) can be modelled so-called pseudo-partial differential equations of the form

$$\partial_s L = -\frac{1}{2}(-\Delta)^{p/2}L. \tag{5}$$

These scale-spaces can be related to the theory of Lévy processes and infinitely divisible distributions. For example, according to this theory a non-trivial probability measure on $\mathbb{R}^N$ is α-stable with $0 < \alpha \leq 2$ if and only if its Fourier transform is of the form (3) with $p = \alpha$ [60, page 86]. These scale-space do, however, not obey non-enhancement of local extrema, and we will not consider such self-similar scale-spaces with non-classical infinitesimal generators further, since the main subject of this article is to develop a more general theory for Gaussian scale-spaces corresponding to $p = 2$.

For the specific family of Gaussian scale-space representations Koenderink and van Doorn [35] carried out a closely related study, where they showed that Gaussian derivative operators are natural operators to derive from a scale-space representation, given the assumption of scale invariance. Axiomatic derivations of image processing operators based on scale invariance have also been given in earlier Japanese literature [72].

With regard to temporal data, the first proposal about a scale-space for temporal data was given by Koenderink [33] by applying Gaussian smoothing to a logarithmically transformed time axes. Such temporal smoothing filters have been considered in follow-up works by Florack [20] and ter Haar Romeny et al. [68]. These approaches, however, require infinite memory of the past and have so far not been developed for computational applications. To handle time-causality in a manner more suitable for real-time implementation, Lindeberg and Fagerström [49] expressed a strictly time-recursive space-time separable spatio-temporal scale-space model based on cascades of temporal scale-space kernels in terms of either truncated exponential functions or first-order recursive filters. These temporal scale-space models also had the attractive and memory saving property that

temporal derivatives could be computed from differences between temporal channels at different scales, thus eliminating the need for complimentary time buffering. A similar computation of temporal derivatives has been used by Fleet and Langley [19]. Early work on non-separable spatio-temporal scale-spaces with velocity adaptation was presented in Lindeberg [43, 46] with applications to Galilean invariant image descriptors and recognition of activities in [36, 37, 52]. More recently, Fagerström [16, 17] has studied scale-invariant continuous scale-space models that allows for the construction of continuous semi-groups over the internal memory representation and in a special case lead to a diffusion formulation.

Outside the class of linear operations, there is also a large literature on non-linear scale-spaces [66]. In particular, the works by Alvarez et al. [3] and Guichard [28] have many structural similarities to the linear/affine/spatio-temporal scale-space formulations in terms of semi-group structure, infinitesimal generator and invariance to rescalings and affine or Galilean transformations. Non-linear scale-space that obey similar properties as non-enhancement of local extrema have been studied in particular by Weickert [71]. With close relationship to non-enhancement of local extrema, the maximum principle has been used as a *sufficient* condition for defining linear or non-linear scale-space representations Alvarez et al. [3], Hummel and Moniot [30].

3 Generalizing Non-enhancement to Non-isotropic Domains

Out of the above mentioned large family of possibilities, we shall here start by exploring the richer structure that can be obtained from a scale-space family if the requirements about rotational symmetry are relaxed. To begin the treatment, let us start by restating the set of scale-space axioms that the analysis will be based on:

3.1 Structural Scale-Space Axioms

We would like to model an uncommitted visual front-end, that performs linear and shift-invariant operations. Hence, we assume that each scale level $L(\cdot; s)$ is generated by convolving the original signal f with a convolution kernel $T(\cdot; s)$,

$$L(\cdot; s) = T(\cdot; s) * f(\cdot), \tag{6}$$

in operator form written

$$L(\cdot; s) = \mathcal{T}_s f(\cdot). \tag{7}$$

Then, to ensure regularity with respect to scale, we assume that the family of scale-space kernels $T(\cdot; s)$ forms a *semi-group*

$$T(\cdot; s_1) * T(\cdot; s_2) = T(\cdot; s_1 + s_2). \tag{8}$$

This condition means that all scale levels are computed from conceptually similar operations and that, in addition, the transformation from any fine scale level to any coarser scale level is of the same form as the transformation from the original signal

$$L(\cdot;\ s_2) = T(\cdot;\ s_2 - s_1) * L(\cdot;\ s_1). \qquad (9)$$

Another important consequence of imposing the semi-group requirement on the family of convolution kernels is that if we assume reasonable continuity requirements of T with respect to variations of the scale parameter s (see below), then it follows from a general result in functional analysis [29] that there exists a limit case operator, the *infinitesimal generator*

$$\mathcal{A}f = \lim_{h\downarrow 0} \frac{T(\cdot;\ h) * f - f}{h}, \qquad (10)$$

such that the scale-space family satisfies a differential equation of the form

$$\partial_s L(\cdot;\ s) = \lim_{h\downarrow 0} \frac{L(\cdot;\ t + h) - L(\cdot;\ s)}{h} = \mathcal{A}L(\cdot;\ s). \qquad (11)$$

With regard to differential equations, it is natural to formalize the non-enhancement requirement of local extrema in terms of a sign condition on the derivative of the scale-space family with respect to the scale parameter. Hence, at any non-degenerate extremum point over the image domain, in other words for each local extremum at which the determinant of the Hessian matrix is non-zero, we require the following conditions[3] to hold (see Fig. 1 for an illustration):

$$\partial_s L \le 0 \quad \text{at any non-degenerate local maximum}, \qquad (12)$$

$$\partial_s L \ge 0 \quad \text{at any non-degenerate local minimum}. \qquad (13)$$

3.2 Formal Definitions

To be able to use tools from functional analysis, we will initially assume that both the original signal f and the family of convolution kernels $T(\cdot;\ s)$ are in the Banach space $L^2(\mathbb{R}^N)$, i.e., $f, T(\cdot;\ t) \in L^2(\mathbb{R}^N)$, with the norm

$$\|f\|_2^2 = \int_{x\in\mathbb{R}^N} |f|^2\, dx. \qquad (14)$$

---

[3]The careful reader may note that this formulation of non-enhancement of local extrema is expressed in terms of "vertical derivatives" ∂_s in scale-space. When the signal is subject to scale-space smoothing, however, the positions of the local extrema may change with scale by a drift velocity v. Hence, one could find it more natural to consider derivatives along such *extremum paths* $\partial_{\bar{s}} = \partial_s + v\,\partial_x$ instead of "vertical derivatives" ∂_s with respect to scale. At the position of a local extremum, however, the "horizontal component" $v\,\partial_x$ will be zero, and the two types of definitions are therefore equivalent.
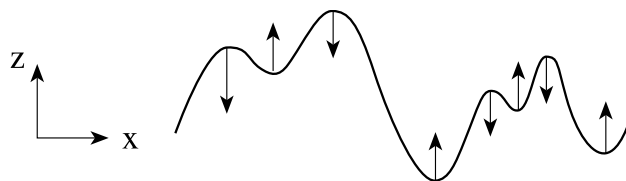
Fig. 1 The non-enhancement condition of local extrema means that the grey-level value of a local maximum must not increase with scale and that the grey-level of a local minimum must not decrease

Then also the scale-space representations $L(\cdot;\ s)$ will be in the same space, and we can define a family of scale-space smoothing operators $\mathcal{T}(s)$ from $L^2(\mathbb{R}^N)$ to itself by

$$L(\cdot;\ s) = \mathcal{T}(s)\, f. \qquad (15)$$

The semi-group structure under convolution transformations in (8) in combination with the initial condition $L(\cdot;\ 0) = f$ imply that we require a semi-group structure on the smoothing operators

$$\begin{cases} \mathcal{T}(s_1)\,\mathcal{T}(s_2) = \mathcal{T}(s_1 + s_2), \\ \mathcal{T}(0) = I, \end{cases} \qquad (16)$$

and in order to ensure sufficient regularity with respect to the scale parameter, we assume the semi-group to be *strongly continuous* (C_0) in the sense that

$$\lim_{s\downarrow s_0} \|(\mathcal{T}(s) - \mathcal{T}(s_0))(f)\|_2 = 0 \qquad (17)$$

should hold for each $f \in L^2(\mathbb{R}^N)$ and for any $s_0 \ge 0$ [29, page 59], [26, page 14], [58, page 4]. In terms of explicit convolution kernels, this requirement corresponds to continuity at the origin in the sense that

$$\lim_{s\downarrow 0} \|T(\cdot;\ s) * f - f\|_2 = 0 \qquad (18)$$

should hold for every $f \in L^2(\mathbb{R}^N)$. Given these assumptions, it follows from a general result in functional analysis [29, page 308], [26, page 14], [58, page 5] that there exists a limit case operator, the *infinitesimal generator*

$$\mathcal{A}f = \lim_{s\downarrow 0} \frac{\mathcal{T}(s)\, f - f}{s} = \partial_s \mathcal{T}(s)\, f|_{s=0}, \qquad (19)$$

such that the scale-space family satisfies a differential equation of the form

$$\partial_s L(\cdot;\ s) = \lim_{h\downarrow 0} \frac{L(\cdot;\ s + h) - L(\cdot;\ s)}{h}$$
$$= \mathcal{A}(\mathcal{T}(s)\, f(\cdot)) = \mathcal{A}L(\cdot;\ s). \qquad (20)$$

The set of elements $f \in L^2(\mathbb{R}^N)$ for which $\mathcal{A}$ exists is denoted $D(\mathcal{A})$. This set is not empty and never reduces to the zero element. Actually, $D(\mathcal{A})$ is even dense in $L^2(\mathbb{R}^N)$

[29, page 308], [58, page 5]. Hence, the scale-space family will satisfy a first-order differential equation with respect to scale.

To ensure sufficient regularity, we will also assume that the convolution kernels $T(\cdot;\, s) \in L^1(\mathbb{R}^N)$. This assumption implies that for all smooth functions $f \in C^\infty(\mathbb{R}^N)$ with compact support, spatial derivatives of the scale-space representation L do always exist by

$$\partial_{x_j} L(\cdot;\, s) = \partial_{x_j}(T(\cdot;\, s) * f(\cdot)) = T(\cdot;\, s) * (\partial_{x_j} f)(\cdot) \quad (21)$$

even at $s = 0$. By applying this property recursively, the scale-space representation of a smooth signal with compact support will also be guaranteed to be smooth, i.e., $L(\cdot;\, s) \in C^\infty(\mathbb{R}^N)$, also at $s = 0$.[4]

In the following, we will show that the requirement of non-enhancement of local extrema combined with semi-group structure and regularity properties with respect to the scale parameter imply that the scale-space family must satisfy a family of parabolic differential equations, in which the second-order terms are determined by a positive semi-definite covariance matrix, while the first-order terms may be arbitrary. Specifically, the result that we shall derive implies that (i) the infinitesimal operator $\mathcal{A}$ must be a local operator that depends on local derivatives only, (ii) we cannot allow for derivatives of higher order than two, and (iii) we cannot allow for any zero-order term either.

To be able to express the non-enhancement condition in a context where derivatives of the scale-space representation with respect to the scale parameter are well-defined pointwise and not only almost everywhere, we will make a further requirement explained in more detail in Appendix A that the semi-group $\mathcal{T}$ should obey the following regularity requirements with respect to variations of the scale parameter

$$\lim_{h \downarrow 0} \left\| \frac{1}{h} \int_{s=0}^{h} \mathcal{T}(s) f\, ds - f \right\|_{H^k(\mathbb{R}^N)}$$

$$= \lim_{h \downarrow 0} \left\| \frac{1}{h} \int_{s=0}^{h} (\mathcal{T}(s) - I) f\, ds \right\|_{H^k(\mathbb{R}^N)} = 0 \quad (22)$$

for some $k > N/2$ and for all smooth functions $f \in L^1(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^N)$, where

$$\|u\|_{H^k(\mathbb{R}^N)} = \left(\int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2 \right)^k |\hat{u}(\omega)|^2 d\omega \right)^{1/2}. \quad (23)$$

---

[4]Usually, one otherwise uses a family of smooth and rapidly decreasing smooth convolution kernels for defining derivatives of non-differentiable functions f, by the theory of Schwartz distributions $\partial_{x_j} L(\cdot;\, s) = \partial_{x_j}(T(\cdot;\, s) * f(\cdot)) = (\partial_{x_j} T(\cdot;\, s)) * f(\cdot)$. Here, however, we start by initially weaker assumptions on the family of convolution kernels, with the aim that the requirement of non-enhancement of local extrema will then imply stronger restrictions on the scale-space representation when combined with a semi-group structure and rather modest regularity requirements on the semi-group with respect to variations of the scale parameter s.

This continuity condition is referred to as C_1 continuity and is a stronger condition than the more commonly used C_0 continuity of semi-groups [29, page 322].

As pre-requisites to the treatment that shall be performed, let us first summarize the basic algebraic structure in terms of a definition and a lemma.

Definition 1 (Continuous pre-scale-space representation) Let $f \in L^2(\mathbb{R}^N)$ be a continuous signal and let $\mathcal{T}(s)$ with $s \in \mathbb{R}_+$ be a strongly continuous semi-group of linear and shift-invariant operators from $L^2(\mathbb{R}^N)$ to $L^2(\mathbb{R}^N)$ according to (16) and (18), where the convolution kernels $T(\cdot;\, s)$ are also required to be in $L^1(\mathbb{R}^N)$ and the semi-group is also for some value $k > N/2$ required to be C_1 continuous with respect to the L^2-based Sobolev norm $\| \cdot \|_{H^k(\mathbb{R}^N)}$ for all smooth functions $f \in L^1(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$. Then, the one-parameter family of signals $L: \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ given by

$$L(\cdot;\, s) = \mathcal{T}(s) f \quad (24)$$

is said to be the continuous pre-scale-space representation of f generated by $\mathcal{T}(s)$.

As we have described in connection with (19) and (11), this algebraic structure implies that the pre-scale-space representation will be differentiable with respect to the scale parameter and will possess an infinitesimal generator.

Lemma 2 (A continuous pre-scale-space representation is differentiable) *Let $L: \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ be the continuous pre-scale-space representation of a signal $f \in L^2(\mathbb{R}^N)$. Then, L satisfies the differential equation*

$$\partial_s L = \mathcal{A} L \quad (25)$$

for some linear and shift-invariant operator $\mathcal{A}$ from $L^2(\mathbb{R}^N)$ to $L^2(\mathbb{R}^N)$. The scale-space representation $L(\cdot;\, s)$ of a smooth function $f \in C^\infty(\mathbb{R}^N)$ of compact support is smooth as function over the spatial domain, and smooth functions f of compact support are in the support $D(\mathcal{A})$ implying that for smooth functions of compact support the partial derivatives $\partial_s L(x;\, s)$ are well-defined for every $(x;\, s) \in \mathbb{R}^N \times \mathbb{R}$.

Proof The basic structure follows from [29, page 308] and our previous treatment. Since L is generated from f by convolutions, it follows that $\mathcal{A}$ must also be shift-invariant and commute with the shift operator $(\mathcal{S}_{\Delta x} f)(x) = f(x - \Delta x)$. The operator $\mathcal{A}$ does, however, not need to be bounded.

The smoothness of L with respect to space follows from the assumption of the convolution kernels $T(\cdot;\, s) \in L^1(\mathbb{R}^N)$ and the discussion in connection with (21). As shown in Appendix A, the regularity requirements on the semi-group imply that smooth functions of compact support are in the domain $D(\mathcal{A})$ of $\mathcal{A}$ implying that the partial derivatives $\partial_s L(x;\, s)$ are well-defined for every $(x;\, s) \in \mathbb{R}^N \times \mathbb{R}_+$. $\square$

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, such that the grey-level value in every local maximum point must not increase, whereas the grey-level value in every local minimum point must not decrease.

Definition 3 (Pre-scale-space property: Non-enhancement of local extrema) A continuous pre-scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ of a smooth signal $f \in L^2(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$ is said to possess continuous non-enhancement pre-scale-space properties, or equivalently not to enhance local extrema, if for every value of the scale parameter $s_0 \in \mathbb{R}_+$ it holds that if $x_0 \in \mathbb{R}^N$ is a critical point for the mapping $x \mapsto L(x; s_0)$ and if the Hessian matrix at this point is non-degenerate, then the derivative of L with respect to s at this point has the same sign as the Hessian matrix, *i.e.,*

$$\text{sign}\,\partial_s L = \text{sign}\,\text{trace}\,\mathcal{H}L. \tag{26}$$

This condition is closely related to the *maximum principle* for elliptic and parabolic equations, however, it is not identical. The maximum principle refers to *global property* of a function concerning *the global maximum* (or minimum), while non-enhancement of local extrema refers to a *local property* concerning *every local extremum* (see Appendix B).

Now we can state that a semi-group of operators generates is a scale-space family representation if it leads to non-enhancement of local extrema for *any* input signal.

Definition 4 (Continuous non-enhancement scale-space representation) Let $\mathcal{T}$ be a strongly continuous semi-group of linear and shift-invariant operators from $L^2(\mathbb{R}^N)$ to $L^2(\mathbb{R}^N)$. Given a signal $f \in L^2(\mathbb{R}^N)$, the pre-scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ of f is said to be a continuous scale-space representation of f if and only if it for every smooth function $f' \in L^2(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$ of compact support it holds that the pre-scale-space representation $L': \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ of f' generated by the semi-group $\mathcal{T}(s)$ obeys non-enhancement of local extrema.

3.3 Necessity and Sufficiency

We shall first show that these conditions by necessity imply that the scale-space family L must satisfy the diffusion equation.

Theorem 5 (Non-enhancement scale-space for continuous signals: Necessity) *A continuous non-enhancement scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ of a signal $f \in L^2(\mathbb{R}^N)$ satisfies a parabolic differential equation*

$$\partial_s L = \frac{1}{2}\nabla^T(\Sigma_0 \nabla L) - \delta_0^T \nabla L \tag{27}$$

with initial condition $L(\cdot; 0) = f(\cdot)$ for some positive semidefinite covariance matrix Σ_0 and some vector δ_0.

Proof The proof consists of two parts. The first part has already been presented in Lemma 2, where it was shown that a pre-scale-space family obeys a linear differential equation, where the infinitesimal generator is shift-invariant. Given the regularity requirements of the semi-group, Lemma 2 also implies that the scale-space representation is guaranteed to be smooth as function over the spatial domain and that smooth functions of compact support are in $D(\mathcal{A})$ implying that the partial derivatives $\partial_s L(x; s)$ are well-defined for every $(x; s) \in \mathbb{R}^N \times \mathbb{R}_+$. In the second part, a set of counterexamples $L(x; 0) = f$ at $s = 0$ will be constructed from various simple smooth functions f of compact support to delimit the class of possible operators.

C.1. The extremum point condition (12) in combination with definition 4 means that $\mathcal{A}$ must be a pure differential operator. This can be easily understood by studying the following class of counterexamples: Consider a smooth (C^∞) function $f_1 = f_I: \mathbb{R}^N \to \mathbb{R}$ such that (i) f_I has a maximum point at the origin and the Hessian matrix is negative definite at the origin, and (ii) for some $\varepsilon > 0$ the test function is zero $f_I(x) = 0$ outside a smaller circle around the origin with $|x| \geq \frac{\varepsilon}{2}$. Then, we must have $\mathcal{A}f_1 = \mathcal{A}f_I = C_1 \leq 0$. Fixate this function f_I, the value of ε and the operator $\mathcal{A}$. Consider next the test function

$$f_2 = f_I + f_E, \tag{28}$$

where

$$\nabla f_I(0) = 0, \tag{29}$$

$$\text{sign}\,\mathcal{H}f_I(0) < 0, \tag{30}$$

$$f_I = 0 \quad \text{when } |x| \geq \varepsilon/2, \tag{31}$$

$$f_E = 0 \quad \text{when } |x| \leq \varepsilon \tag{32}$$

while the function f_E may now assume non-zero values outside the circular region $|x| \leq \varepsilon$. Let us initially assume that $\mathcal{A}f_E = C_2 \neq 0$. Then, if we consider a third test function of the form $f_3 = f_I + \beta_1 f_E$, we get

$$\partial_s f_3 = \mathcal{A}f_I + \beta_1 \mathcal{A}f_E = C_1 + \beta_1 C_2. \tag{33}$$

Obviously, the sign of this expression can be made positive and (12) be violated by a suitable choice of β_1. Hence, for any $\varepsilon > 0$ we have to require that $\mathcal{A}f_E$ must be identically zero for all functions that assume non-zero values outside the region $|x| < \varepsilon$. In other words, $\mathcal{A}$ must be a *local operator* and $\mathcal{A}f$ can only exploit information from f at the

central point. This means that for any smooth function $\mathcal{A}f$ must be of the form

$$\mathcal{A}f = \sum_{\xi \in \mathbb{Z}_+^N} a_\xi L_{x^\xi}, \qquad (34)$$

where $\xi = (\xi_1, \xi_2, \ldots, \xi_N)$ is a multi-index, $a_\xi \in \mathbb{R} \; \forall \xi$ and $L_{x^\xi} = L_{x_1^{\xi_1} x_2^{\xi_2} \ldots x_N^{\xi_N}}$.

C.2. The extremum point condition (12) also means that $\mathcal{A}L$ must not contain any term proportional to L or derivatives of order higher than two. This can be seen by considering a test function of the form

$$f_4(x) = (x_1^2 + x_2^2 + \cdots + x_N^2 + \beta_2 x^\eta) \chi(x) \qquad (35)$$

for some $\eta = (\eta_1, \eta_2, \ldots, \eta_N) \in \mathbb{Z}^N$ with $|\eta| = |\eta_1| + |\eta_2| + \cdots + |\eta_N| > 2$ where $\chi(x) \geq 0$ is a smooth function of compact support identically equal to one $\chi(x) = 1$ in the disk $|x| \leq 1$ and identically equal to zero $\chi(x) = 0$ outside the disk $|x| \leq 2$. If $a_\xi \neq 0$ for some $\xi \in \mathbb{Z}^N$, it is clear that we can choose $\eta = \xi$ and by a suitable choice of β_2 we can make the sign of $\mathcal{A}f_4$ arbitrary and hence violate (12). Similarly, by considering a test function of the form

$$f_5(x) = (x_1^2 + x_2^2 + \cdots + x_N^2 + \beta_3) \chi(x) \qquad (36)$$

it follows that a_0 must be zero. Thus, $\mathcal{A}$ can only contain derivatives of order one and two.

C.3. Concerning the remaining first- and second-order terms, let us first note that the non-enhancement condition of local extrema does not impose any constraints on the first-order terms, since the influence of the first-order terms vanishes at critical points.

Regarding the second-order terms, we can next observe that without loss of generality, the influence of the second-order terms can be written

$$\partial_s L = \frac{1}{2} \nabla^T (\Sigma_0 \nabla L) \qquad (37)$$

for some symmetric matrix Σ_0. Let us next show by counterexample that Σ_0 must be positive semi-definite, *i.e.*, if Σ_0 would have a strictly negative eigenvalue, then the non-enhancement property of local extrema would be violated. Such a violation occurs if we can find a positive definite matrix H such that $\partial_s L(0; 0) < 0$ for $L(x; 0) = x^T H x$, *i.e.*, if for a given Σ_0 with both positive and negative eigenvalues, we could find some positive definite H such that

$$\partial_s L = \frac{1}{2} \nabla^T (\Sigma_0 \nabla L) = \nabla^T (\Sigma_0 H x)$$
$$= \text{trace}(\Sigma_0 H) < 0. \qquad (38)$$

To construct such a H, let us assume that Σ_0 has eigenvectors $e_1 \ldots e_N$ with associated eigenvalues $\lambda_1 \ldots \lambda_N$ and choose H to have the same set of eigenvectors while having different eigenvalues $\mu_1 \ldots \mu_N$. Then, from

$$\partial_s L = \text{trace}(\Sigma_0 H) = \sum_{i=1}^N \mu_i \lambda_i, \qquad (39)$$

where all $\mu_i > 0$ and at least one $\lambda_i < 0$, it is clear that we can find a suitable combination of $\mu_i > 0$ such that $\partial_s L < 0$ and we would thus violate the requirement of non-enhancement of local extrema. Hence, to guarantee that the requirement of non-enhancement of local extrema can be valid, we have to require that all the eigenvalues of Σ_0 are positive. $\qquad \square$

Remark Equation (27) can also be written on the more common form

$$\partial_s L = \sum_{i=1}^N \sum_{j=1}^N a^{ij} L_{x_i x_j} - \sum b^i L_{x_i}, \qquad (40)$$

where the coefficients a^{ij} form a positive semi-definite matrix. The reason why we have chosen a parameterization in terms of Σ_0 and δ_0 here is because of the close connections to Gaussian kernels and velocity adaptation that will be developed later in Sects. 4–5.

To prove sufficiency, *i.e.*, the reverse statement of Theorem 5, is straightforward and a basic property of parabolic equations. For completeness, we give the result with an explicit proof.

Theorem 6 (Non-enhancement scale-space for continuous signals: Sufficiency) *Given a semi-definite covariance matrix Σ_0, an arbitrary vector δ_0 and any twice continuously differentiable function $f \in L^2(\mathbb{R}^N)$, the solution of the diffusion equation*

$$\partial_s L = \frac{1}{2} \nabla^T (\Sigma_0 \nabla L) - \delta_0^T \nabla L \qquad (41)$$

with initial condition $L(\cdot; 0) = f$ constitutes a continuous non-enhancement scale-space representation of f. Specifically, L obeys

$$\partial_s L \leq 0 \quad \text{at any non-degenerate local maximum}, \qquad (42)$$

$$\partial_s L \geq 0 \quad \text{at any non-degenerate local minimum}. \qquad (43)$$

Proof The regularity properties of the solution are apparent from the regularity properties of parabolic differential equations. To verify non-enhancement of local extrema, consider any non-degenerate local extremum x_0 of L at scale s_0 with Hessian matrix H_0. Using the fact that $\nabla^T (\Sigma_0 \nabla L) =$

trace($\Sigma_0 H_0$), we have $\partial_s L = \text{trace}(\Sigma_0 H_0)$ at the critical point. If x_0 is a local minimum then H_0 is positive semi-definite. Since the trace of the product of two positive semi-definite matrices is greater or equal to zero, it follows that trace($\Sigma_0 H_0$) ≥ 0. If x_0 is a local maximum, we can apply similar reasoning to $-L$. $\qquad\square$

To conclude, we can take the results in Theorems 5–6 as a generalized definition of the notion of *Gaussian scale-space*.

Definition 7 (Generalized Gaussian scale-space) The (non-enhancement) scale-space families that are defined from the solutions of parabolic differential equations of the form (27) are referred to as generalized Gaussian scale-spaces.

Concerning the regularity properties of the input signal f, it follows from the strong regularizing properties of the Gaussian kernel that the solution of (27) will be smooth for $k > 0$ for any bounded function $f : \mathbb{R}^n \to \mathbb{R}$. Hence, for the purpose of generating a scale-space representation of a real-world signal f, we can relax the condition on f to f being bounded.

4 Interpretations

In previous section, we showed that for a linear and shift-invariant infinitesimal generator $\mathcal{A}$ the requirement of non-enhancement of local extrema in combination with continuity conditions implies that the scale-space representation should satisfy a parabolic differential equation of the form

$$\partial_s L = \frac{1}{2} \nabla^T (\Sigma_0 \nabla L) - \delta_0^T \nabla L \qquad (44)$$

for some positive semi-definite covariance matrix Σ_0 and some translation vector δ_0. If we take a delta function as input, the interpretation of this evolution equation is that at any time moment the solution corresponds to a Gaussian kernel with covariance matrix $\Sigma_s = s \Sigma_0$ centered at $\delta_s = s \delta_0$. Thus, we can interpret the impulse response as a gradually growing elongated Gaussian kernel that moves with velocity δ_0 with respect to the evolution parameter s. In terms of filtering operations, this scale-space can equivalently be constructed by convolution with *affine and velocity-adapted Gaussian kernels*

$$g(x; \ \Sigma_s, \delta_s) = \frac{1}{(2\pi)^{N/2} \sqrt{\det \Sigma_s}} e^{-(x-\delta_s)^T \Sigma_s^{-1} (x-\delta_s)/2},$$
$$(45)$$

which for a given $\Sigma_s = s \Sigma_0$ and a given $\delta_s = s \delta_0$ satisfy the diffusion equation (44). The Fourier transform of this shifted

Gaussian kernel is

$$\hat{g}(\omega; \ \Sigma_s, \delta_s) = \int_{x \in \mathbb{R}^N} g(x; \ \Sigma_s, \delta_s) e^{-i\omega^T x} \, dx$$
$$= e^{i\omega^T \delta_s - \omega^T \Sigma_s \omega/2}. \qquad (46)$$

From the diffusion equation formulation or the Fourier transform, it can be seen that these shifted and shape-adapted kernels satisfy the following generalized semi-group property

$$g(\cdot; \ \Sigma_1, v_1) * g(\cdot; \ \Sigma_2, v_2) = g(\cdot; \ \Sigma_1 + \Sigma_2, v_1 + v_2). \quad (47)$$

Transformation property under linear transformations. This scale-space concept has the attractive property that it is *closed* under affine transformations: If two image patterns f_L and f_R are related by an affine transformation

$$f_L(\xi) = f_R(\eta) \quad \text{where } \eta = A\xi + b, \qquad (48)$$

and if linear scale-space representations of these images are defined by

$$L(\cdot; \ \Sigma_L, \delta_L) = g(\cdot; \ \Sigma_L, \delta_L) * f_L(\cdot), \qquad (49)$$

$$R(\cdot; \ \Sigma_R, \delta_R) = g(\cdot; \ \Sigma_R, \delta_R) * f_R(\cdot), \qquad (50)$$

then L and R are related by

$$L(x; \ \Sigma_L, \delta_L) = R(y; \ \Sigma_R, \delta_R), \qquad (51)$$

where the covariance matrices Σ_L and Σ_R satisfy [51]

$$\Sigma_R = A \Sigma_L A^T, \qquad (52)$$

and the velocity terms δ_L and δ_R in the Gaussian kernels can be traded against coordinate shifts in x and y as long as the following relation is satisfied:

$$y - \delta_R = A(x - \delta_L) + b. \qquad (53)$$

This property is highly useful in connection with visual tasks involving image deformations, such as image matching, flow estimation and shape estimation. The closedness under affine transformation allows for perfect modelling and matching of image data under first-order approximations of image deformations due to motion or the perspective mapping, and has been explored by *e.g.* [5, 6, 36, 37, 51, 52, 55, 56, 61, 69].

4.1 Special Cases

The above mentioned relations provide a general structure for linear scale-space concepts on shift-invariant continuous domains. Specifically, it includes the following special cases:

4.1.1 Rotationally Symmetric Linear Scale-Space

If we require the covariance matrix Σ_0 to be a unit matrix, we obtain the regular (isotropic) Gaussian scale-space [20, 32, 39, 73].

$$L(x; s) = \int_{\xi \in \mathbb{R}^N} f(x - \xi) \, g(\xi; s) \, d\xi \qquad (54)$$

generated by convolutions with rotationally symmetric Gaussian kernels

$$g(x; s) = \frac{1}{(2\pi s)^{N/2}} e^{-(x_1^2 + \cdots + x_N^2)/2s}. \qquad (55)$$

From this scale-space, we can define the *multi-scale N-jet* by applying partial derivatives to the scale-space

$$L_{x^\alpha} = \partial_{x^\alpha} L = \partial_{x_1^{\alpha_1} \ldots x_N^{\alpha_N}} L, \qquad (56)$$

where we have introduced multi-index $\alpha = (\alpha_1, \ldots, \alpha_N)$ to simplify the notation. Due to the linearity of the diffusion equation, all these *scale-space derivatives* L_{x^α} satisfy similar scale-space properties in terms of non-enhancement of local extrema as the original scale-space L. So do also *directional derivatives*. In two dimensions, the M:th order directional derivative in the direction $(\cos \varphi, \sin \varphi)$ is given by

$$\partial_{\varphi^M} L = (\cos \varphi \, \partial_x + \sin \varphi \, \partial_y)^M L$$
$$= \sum_{k=0}^{M} \binom{M}{k} \cos^k \varphi \, \sin^k \varphi \, L_{x^k y^{M-k}}. \qquad (57)$$

With regard to image deformations, the closedness properties of this original scale-space are restricted to translations, rotations and rescalings. This scale-space concept is on the other hand *separable* in any orthonormal coordinate system.

4.1.2 Affine Gaussian Scale-Space

If we relax the condition about rotational symmetry, while keeping a requirement that the corresponding Green's function should be mirror symmetric on every line through the origin (in the sense that the filters h should satisfy $h(-x, -y; s) = h(x, y; s)$ for every $(x, y) \in \mathbb{R}^2$), we obtain the *affine Gaussian scale-space representation*, generated by convolution with non-uniform Gaussian kernels

$$g(x; \Sigma_s) = \frac{1}{(2\pi)^{N/2} \sqrt{\det \Sigma_s}} e^{-x^T \Sigma_s^{-1} x/2}, \qquad (58)$$

where Σ_s is a symmetric positive definite (covariance) matrix. Besides the requirement of rotational symmetry, the affine Gaussian scale-space basically satisfies similar scale-space properties as the linear scale-space. The main difference is that the affine Gaussian scale-space is closed under the full group of non-singular affine transformations.

With regard to image processing and computer vision, this means that image data subjected to affine transformations can be perfectly captured with the extended class of affine scale-space operations. Specifically, for two-dimensional images arising as perspective projections of three-dimensional scenes, this notion of affine image deformations can be used as a first-order linear approximation of non-linear perspective effects. This scale-space concept has been studied by [27, 39, 50] and is highly useful when computing surface shape under local affine distortion [51] and performing affine invariant segmentation [5] and matching [6, 55, 61, 69]. Combined with derivative operations, it can also serve as a natural idealized model for filter banks [24, 64] consisting of elongated directional filters [59].

To parameterize the affine Gaussian kernels, let us in the two-dimensional case consider the covariance matrix determined by two eigenvalues λ_1, λ_2 and one orientation β. Then, the covariance matrix can be written

$$\Sigma' = \begin{pmatrix} \lambda_1 \cos^2 \beta + \lambda_2 \sin^2 \beta & (\lambda_1 - \lambda_2) \cos \beta \sin \beta \\ (\lambda_1 - \lambda_2) \cos \beta \sin \beta & \lambda_1 \sin^2 \beta + \lambda_2 \cos^2 \beta \end{pmatrix}. \qquad (59)$$

Figure 2 shows a few examples of affine Gaussian filter kernels obtained in this way. Directional derivative operators of any order or orientation can then be obtained by combining (59) and (57); see Fig. 3.

When computing directional derivatives from elongated affine Gaussian kernels, it should be noted that it is natural to align the orientations of the directional derivative operators (the angle φ in (57)) with the orientations of the eigendirections of the covariance matrix in the affine Gaussian kernels (the angle β in (59)). This is also the most likely model for biological vision (see Figs. 18 and 19 in Sect. 6).

4.1.3 Gaussian Spatio-Temporal Scale-Space

While the affine Gaussian scale-space generated by (58) has essentially the same invariance properties as the spatio-temporal scale-space representation generated by (45), one motivation for keeping the velocity term v in (45) arises when studying time dependent data. On a temporal domain, the non-zero offset in the Gaussian kernel can be used as a simplified model of the fact that all computations require non-zero computation time and time averages can only be computed from data that have occurred in the past. This requirement of *time causality* implies that any temporal receptive field has to be associated with a non-zero time delay. Moreover, on a spatio-temporal domain, we may want the receptive fields to follow the direction of motion, in such a way that the centres and the shapes of the receptive fields are adapted to the direction of motion; see Fig. 4 for an illustration. Such *velocity adaptation* [43] is useful for reducing the
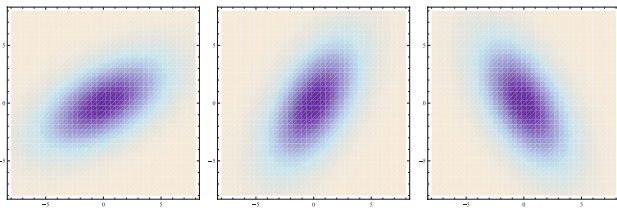
Fig. 2 Examples of affine Gaussian kernels in the two-dimensional case ($\lambda_1 = 16$, $\lambda_2 = 4$, $\beta = \pi/6, \pi/3, 2\pi/3$)
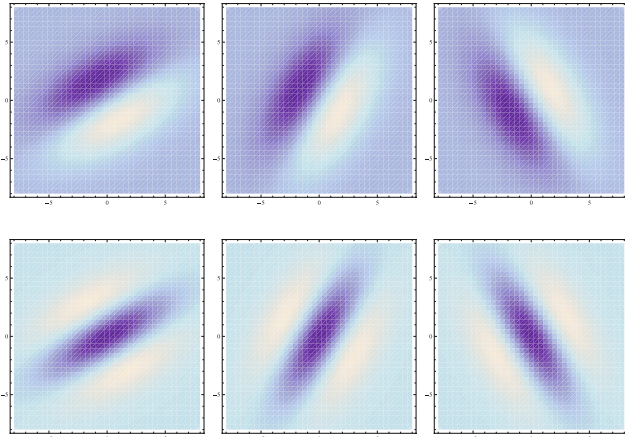

Fig. 3 Elongated filters obtained by applying first- and second order directional derivatives to affine Gaussian kernels ($\lambda_1 = 16$, $\lambda_2 = 4$, $\beta = \pi/6, \pi/3, 2\pi/3$, $\varphi = \beta + \pi/2$)

temporal blur induced by observing objects that move relative to the camera and is a natural mechanism to include in modules for multi-scale motion estimation [22, 56] and for recognizing spatio-temporal activities or events [36, 37]. In particular, invariance to local Galilean transformations can be achieved if the filter parameters can be adapted to the local spatio-temporal image structure [52] (see also Appendix C).

With respect to temporal implementation, however, the filters in this Gaussian filter class do not respect time causality in a strict sense. Although the total mass of the filter coefficients that imply access to the future can be made arbitrarily small, by a suitable choice of time delay associated with the scale parameter in the scale direction, all filters in this filter class have support regions that cover the entire time axis and are not suitable for real-time processing of temporal image data. Nevertheless, they are highly useful as the simplest possible model for studying properties of temporal and spatio-temporal scale-spaces. They are also highly useful for off-line processing. We shall later consider the topic of strict temporal causality in detail in Sect. 5.

Parameterization of shape- and velocity-adapted spatio-temporal filters. In the case with two spatial dimensions and one temporal dimension, which will be referred to as
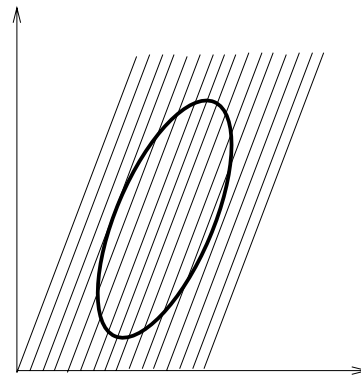
Fig. 4 By adapting the shape and the position of a spatio-temporal smoothing kernel to the direction of motion, we can compute image descriptors that are invariant to constant velocity motion. This property can for example be used for reducing the effect of motion blur when computing image descriptors of moving objects at coarse temporal scales

2+1-dimensional space-time, let us consider a Galilean motion in the image plane

$$\begin{cases} x' = x + v_x t, \\ y' = y + v_y t, \\ t' = t. \end{cases} \tag{60}$$

Then, by transforming the affine covariance matrix in (59) by such a Galilean transformation, we obtain (using (52)) a spatio-temporal covariance matrix of the form

$$\Sigma' = \begin{pmatrix} \lambda_1 \cos^2 \beta + \lambda_2 \sin^2 \beta + v_x^2 \lambda_t & (\lambda_2 - \lambda_1) \cos \beta \sin \beta + v_x v_y \lambda_t & v_x \lambda_t \\ (\lambda_2 - \lambda_1) \cos \beta \sin \beta + v_x v_y \lambda_t & \lambda_1 \sin^2 \beta + \lambda_2 \cos^2 \beta + v_y^2 \lambda_t & v_y \lambda_t \\ v_x \lambda_t & v_y \lambda_t & \lambda_t \end{pmatrix} \tag{61}$$

Velocity-adapted spatio-temporal derivatives are then given by

$$\partial_{\bar{x}} = \partial_x, \qquad \partial_{\bar{y}} = \partial_y, \qquad \partial_{\bar{t}} = v_x\, \partial_x + v_y\, \partial_y + \partial_t. \tag{62}$$

Figures 5–6 show spatio-temporal scale-space kernels generated in this way. Figure 5 shows space-time separable filters corresponding to $v = 0$, while Fig. 6 shows corresponding velocity adapted and non-separable filters for a non-zero velocity $v \neq 0$.

For the specific case with one spatial dimension and one temporal dimension, we obtain

$$\det \Sigma' = \lambda_x \lambda_t = s\tau, \tag{63}$$

$$(X - \delta)^T \Sigma'^{-1} (X - \delta) = \frac{(x - vt)^2}{s} + \frac{(t - \delta_t)^2}{\tau} \tag{64}$$
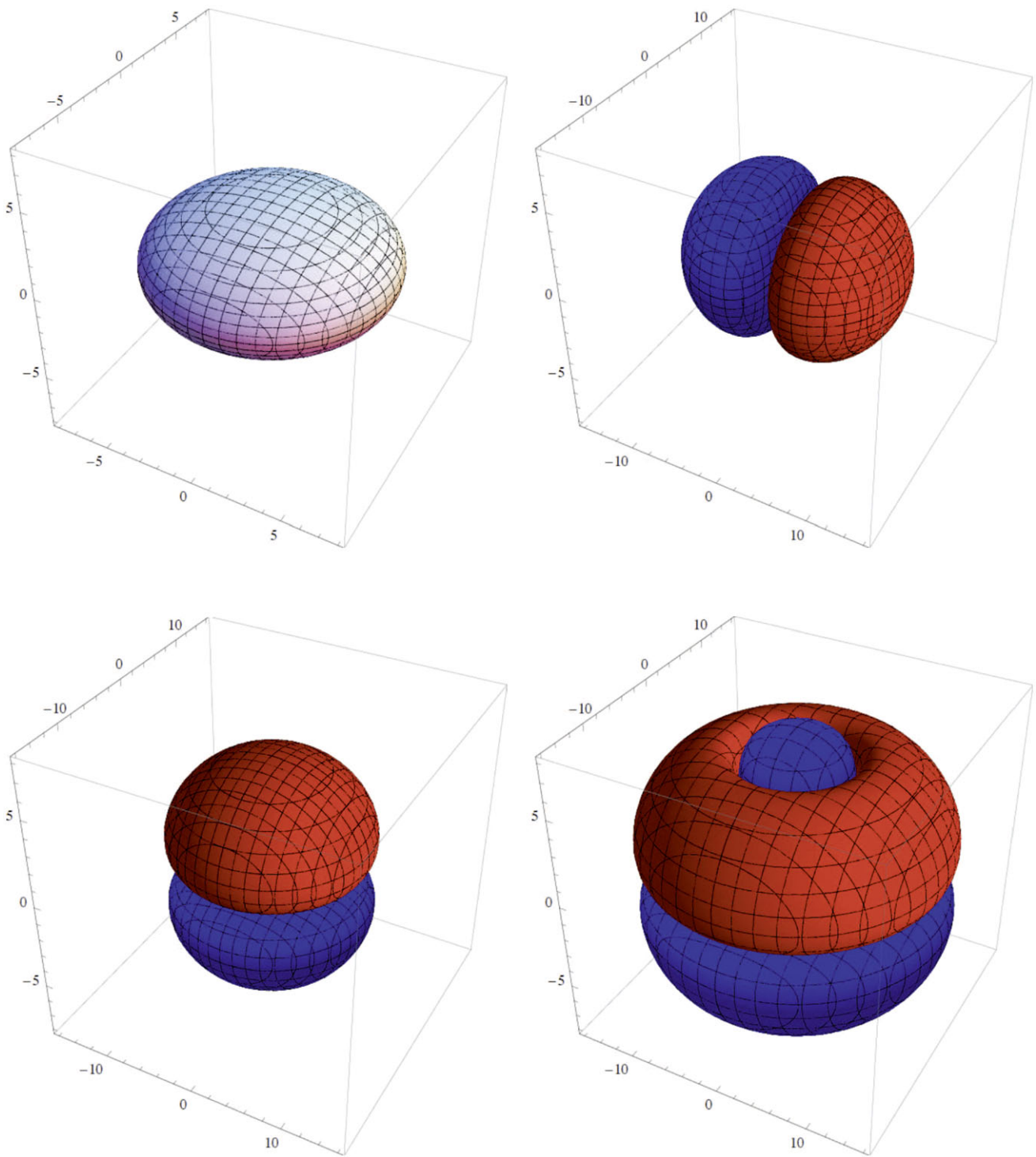
Fig. 5 *Space-time separable Gaussian spatio-temporal scale-space kernels:* (*top left*) Original smoothing kernel $g(x, y, t; \Sigma, v)$; (*top right*) First-order spatial derivative $g_x(x, y, t; \Sigma, v)$; (*bottom left*) First-order temporal derivative $g_t(x, y, t; \Sigma, v)$; (*bottom right*) First-order temporal derivative of the spatial Laplacian $g_{xxt}(x, y, t; \Sigma, v) + g_{yyt}(x, y, t; \Sigma, v)$ ($\lambda_1 = \lambda_2 = 16$, $\lambda_t = 4$, $v_x = v_y = 0$). (*Bottom plane*: space (x, y); *Vertical axis*: time t.)
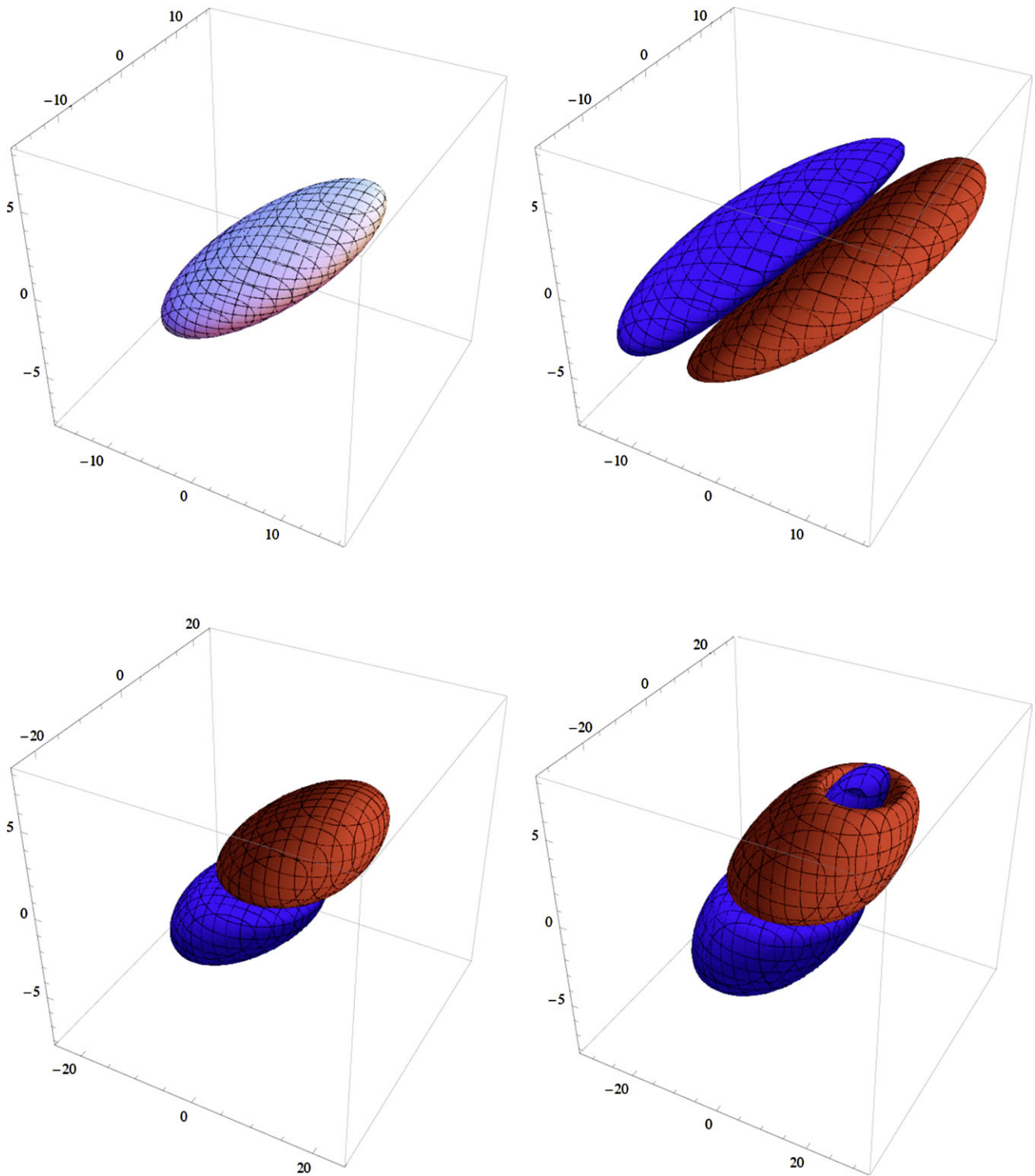
Fig. 6 *Velocity-adapted and non-separable Gaussian spatio-temporal scale-space kernels:* (*top left*) Original smoothing kernel $g(x, y, t; \Sigma, v)$; (*top right*) First-order spatial derivative $g_x(x, y, t; \Sigma, v)$; (*bottom left*) First-order temporal derivative $g_{\bar{t}}(x, y, t; \Sigma, v)$; (*bottom right*) First-order temporal derivative of the spatial Laplacian $g_{xx\bar{t}}(x, y, t; \Sigma, v) + g_{yy\bar{t}}(x, y, t; \Sigma, v)$ ($\lambda_1 = \lambda_2 = 16$, $\lambda_t = 4$, $v_x = 2$, $v_y = 0$). (*Bottom plane*: space (x, y); *Vertical axis*: time t.)

Fig. 7 *Space-time separable kernels $g_{x^{\alpha}t^{\gamma}}(x,t;\ s,\tau,\delta)$ up to order two obtained from the Gaussian spatio-temporal scale-space in the case of a 1+1-D space-time ($s=1$, $\tau=1$, $\delta=2$). (Horizontal axis: space x; Vertical axis: time t.)*
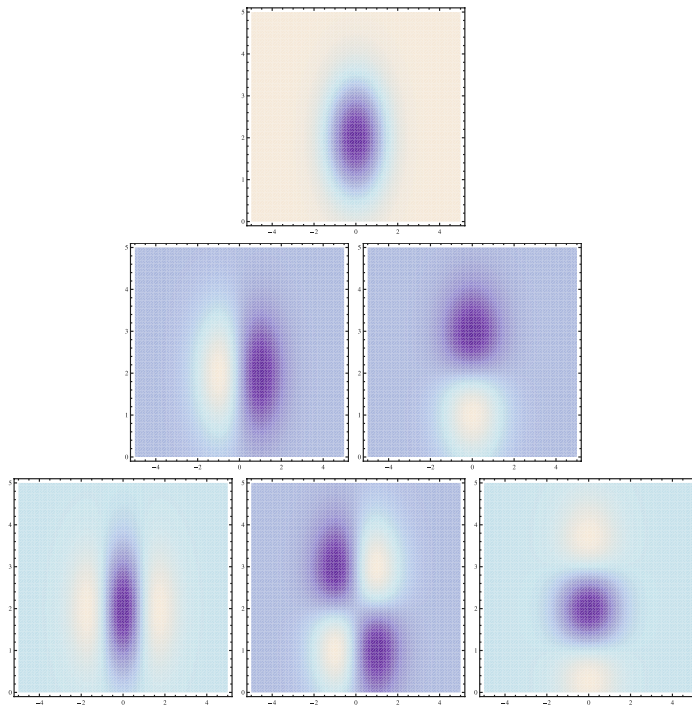
Fig. 8 *Velocity-adapted spatio-temporal kernels $g_{\bar{x}^{\alpha}\bar{t}^{\gamma}}(x,t;\ s,\tau,v,\delta)$ up to order two obtained from the Gaussian spatio-temporal scale-space in the case of a 1+1-D space-time ($s=1$, $\tau=1$, $v=0.75$, $\delta=2$). (Horizontal axis: space x; Vertical axis: time t.)*
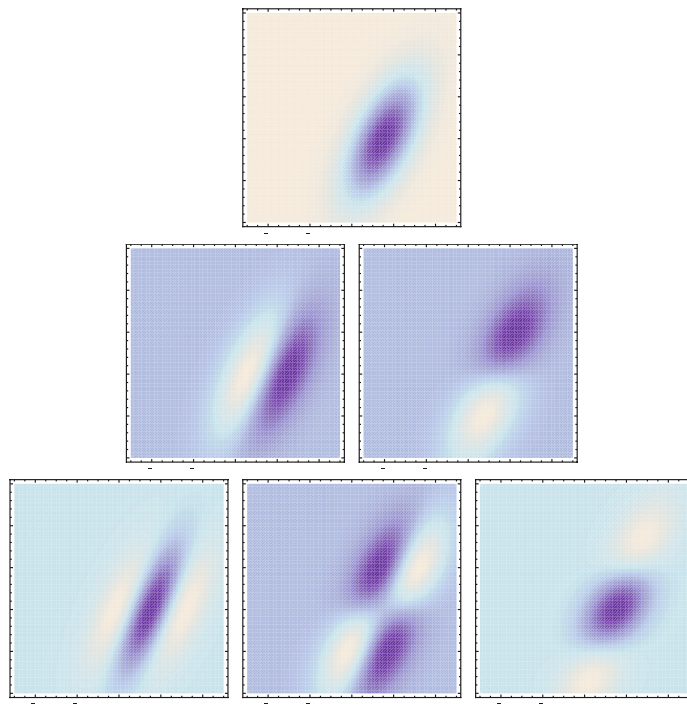
which after insertion into (45) implies that these Gaussian spatio-temporal kernels assume the form

$$g(x,t;\ s,\tau,v,\delta) = \frac{1}{\sqrt{2\pi s}} e^{-\frac{(x-vt)^2}{2s}} \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(t-\delta)^2}{2\tau}}$$

$$= g(x-vt;\ s)\, g(t;\ \tau,\delta). \qquad (65)$$

Figures 7 and 8 show examples of these kernels over a 1+1-D space-time.

When implementing a velocity-adapted spatio-temporal scale-space representation in practice, there are different alternatives to consider. The simplest approach is to use the same velocity vector at all image positions, and is equivalent to global stabilization. More generally, one may also consider different image velocities at different image po-

sitions.[5] In this way, the corresponding velocity-adapted spatio-temporal scale-space representations will for appropriate values of the velocity parameters correspond to filtering along the particle trajectories. Thereby, the system will be able to handle multiple moving objects and will also have the ability to derive a Galilean invariant representation for each object (see next section for details). Alternatively, we may at each image position even consider an *ensemble* of spatio-temporal filters that are tuned to different image velocities—a design with close relations to velocity-tuned receptive fields biological vision (see Sect. 6). Such a parallel treatment of velocity adaption for different image velocities also has the potential to handle transparent motion.

4.1.4 Galilean Invariant Fixed-Point Property of Spatio-Temporal Scale-Space

This spatio-temporal scale-space concept implies that spatio-temporal image data can be smoothed by a family of spatio-temporal filters that correspond to different spatial scales s, temporal scales τ and image velocities v. An underlying intention behind this construction is that the vision system should be able to handle objects that move with different velocities relative to the observer. Specifically, if a particular object moves with image velocity v_0, then the spatio-temporal scale-space representation will for this value of the velocity parameter correspond to filtering along the direction of motion. In practice, however, we cannot expect the velocity of the object to be *a priori* known by the vision system, which is a major motivation for allowing for a *family* of different image velocities in the scale-scale representation. This idea is also in good agreement with findings about velocity-tuned cells in biological vision, that give their strongest responses around a particular stimulus velocity.

If the image velocity v of the filter does not agree with the image velocity v_0 of the object, it is, however, not guaranteed that the corresponding filtered data will be easy to interpret. Therefore one may ask: Is there a way to determine from the output from a particular velocity-adapted filter alone if the output can be regarded as useful or not?

One way of making such judgements of velocity-adapted data can be derived from a *Galilean-invariant fixed-point property* of a notion referred to as *Galilean block diagonalization* of the *spatio-temporal second-moment matrix/structure tensor* [53], which is a generalization of a corresponding affine-invariant fixed-point property of the spatial second-moment matrix/structure tensor [39, Sect. 15.4], [51]. In Appendix C, we describe how such Galilean-invariant fixed points can be constructed theoretically and be detected computationally.

5 Time-Causal Spatio-Temporal Scale-Space

While the above mentioned Gaussian spatio-temporal scale-space model can be successfully used for analyzing off-line data, it is not suitable for a real-time implementation. In a real-time scenario, we have to require all scale-space kernels to be strictly time-causal in the sense that they should not require any access to the future. Fortunately, it is rather straightforward to adapt the above mentioned theories to be truly time-causal.

Time-recursive update rule over time t. Following Lindeberg and Fagerström [49] we would like the scale-space model to be based on time-causal scale-space kernels that lead to a limited internal memory that is successively updated with regard to novel contents. Furthermore, following Fagerström [17] we are seeking a scale-space representation that at any time moment to be computed from a continuous semi-group structure over time. Thus, given any spatio-temporal signal $f(x, t)$ over space-time (x, t), we embed this signal into a scale-space representation L using a complementary spatial scale parameter s and a complementary temporal scale parameter τ that it is to be updated according to the following time-recursive formulation[6]

$$L(x, t_2; s_2, \tau)$$
$$= \int_{\xi \in \mathbb{R}^N} \int_{\zeta \geq 0} T(x - \xi, t_2 - t_1; s_2 - s_1, \tau, \zeta)$$
$$L(\xi, t_1; s_1, \zeta) \, d\zeta \, d\xi, \qquad (66)$$

where T now in combination with L generates a two-parameter cascade structure over both spatial scales s and time t (however NOT over temporal scales τ). In analogy with the spatial scale-space concept, we require the algebraic structure to correspond to convolutions over the spatial domain (x, ξ) and with regard to spatial scales s. We require the algebraic structure to be similar to a cascade structure over time t, using the temporal scales (τ, ζ) as an internal temporal buffer. For reasons that will be apparent later, however, we do not require the updating rule over time to be a true convolution over temporal scales, but a more general integral formulation. Notably, this update rule makes it possible to compute the representation at any coarser scale s_2 and/or any later time moment t_2 from the representation at any finer scale s_1 and any earlier time moment t_1, with the

---

[5]A spatial counterpart of this idea has been developed in Almansa and Lindeberg [2], where the spatial covariance matrix in an affine scale-space representation is allowed to vary in space, to allow for enhancements of local directional image structures in fingerprint images.

[6]Please, note that this formulation is only preliminary. In addition to an update rule on the scale-space representation L, the updating relation must also specify how new information from the input signal f should be incorporated. We will return to this topic in a more precise manner later, in (113). The purpose of this preliminary formulation is to provide a first intuitive start towards the formalism that will follow in terms of differential equations.

arguments of the updating kernel depending only on the differences $s_2 - s_1$ in scale and $t_2 - t_1$ in time. For the update rule over temporal scales as referred to by τ and ζ, however, we do not require a similar structure.

Joint two-parameter semi-group structure over spatial scales s and time t.   Let us now turn to the problem of expressing a spatio-temporal scale-space representation $L : \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ of a spatio-temporal signal $f : \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ defined for all spatial positions x and all positive times $t \geq 0$ and with two scale parameters $(s, \tau) \in \mathbb{R}_+^2$. As combined boundary and initial condition, we take $L(x, t;\ 0, 0) = f(x, t)$, and in terms of transformations from the original spatio-temporal signal f, we assume that the spatio-temporal scale-space can be obtained by a convolution over space x and time t that respects temporal causality over time.

$$L(x, t;\ s, \tau)$$
$$= \int_{u=0}^{t} \int_{\xi \in \mathbb{R}^N} f(\xi, u)\, h(x - \xi, t - u;\ s, \tau)\, d\xi\, du. \quad (67)$$

Initially, we assume that both the original signal f and the convolution kernel h are in the Banach space $X = L^2(\mathbb{R}^N \times \mathbb{R}_+)$. Then, also all scale-space representations L will be in this space. To ensure sufficient regularity, we do also assume that the smoothing functions $T(x, t;\ s, \tau, \zeta)$ in the time recursive formulation (66) and that convolution kernels $h(x, t;\ s, \tau)$ in the transformation (67) from the original spatio-temporal signal are to be absolute integrable, *i.e.,* in $L^1(\mathbb{R}^N \times \mathbb{R}_+)$.

To describe this algebraic structure in terms of operators, let us introduce a two-parameter family of bounded linear operators $\mathcal{T}(s, t)$ from X to X, denoted $\mathcal{T}(s, t) \in O(X)$,

$$L(\cdot, t;\ s, \cdot) = \mathcal{T}(s, t)\, L(\cdot, 0;\ 0, \cdot), \quad (68)$$

where a *two-parameter semi-group* condition on the operator $\mathcal{T}$ over spatial scales s and time t implies that $\mathcal{T}$ must satisfy

$$\begin{cases} \mathcal{T}(s_1, t_1)\, \mathcal{T}(s_2, t_2) = \mathcal{T}(s_1 + s_2, t_1 + t_2), \\ \mathcal{T}(0, 0) = \mathcal{I}. \end{cases} \quad (69)$$

We require this semi-group to be *strongly continuous* (C_0) in the sense that

$$\lim_{(s,t) \downarrow (s_0, t_0)} \|(\mathcal{T}(s, t) - \mathcal{T}(s_0, t_0))(f)\|_2 = 0 \quad (70)$$

should hold for each $f \in X$ and for any $(s_0, t_0) \geq 0$, where the limit operation $(s, t) \downarrow (s_0, t_0)$ should be interpreted so as to hold for all paths $((s - s_0)^2 + (t - t_0)^2) \to 0$ to the origin for which $s \geq s_0$ and $t \geq t_0$.

For this two-parameter semi-group, we define the *infinitesimal generator* as the linear transformation $\mathcal{A} : \mathbb{R}_+^2 \to O(X)$ that satisfies

$$\mathcal{A}(\alpha_1, \alpha_2)\, L = (\mathcal{A}_1, \mathcal{A}_2) \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} L = \alpha_1 \mathcal{A}_1 L + \alpha_2 \mathcal{A}_2 L \quad (71)$$

for all $L \in X$ and all $(a_1, a_2) \in \mathbb{R}_+^2$, where $\mathcal{A}_1$ and $\mathcal{A}_2$ are the infinitesimal generators of the one-parameter semi-groups $\mathcal{T}(s, 0)_{s \geq 0}$ and $\mathcal{T}(0, t)_{t \geq 0}$ respectively, defined in turn from

$$\mathcal{A}_1 L = \lim_{h \downarrow 0} \frac{\mathcal{T}(h, 0)\, L - L}{h} = \partial_s L, \quad (72)$$

$$\mathcal{A}_2 L = \lim_{k \downarrow 0} \frac{\mathcal{T}(0, k)\, L - L}{k} = \partial_t L. \quad (73)$$

Specifically, the *directional derivative* of the semi-group in any direction $u = (\alpha_1, \alpha_2)$ is

$$\mathcal{D}_u L = \lim_{h \downarrow 0} \frac{\mathcal{T}(\alpha_1 h, \alpha_2 h)\, L - L}{h} \quad (74)$$

and for all $u = (\alpha_1, \alpha_2) \in \mathbb{R}_+^2$ and all $f \in X$ we have for $k > 0$ and $t > 0$ that

$$\mathcal{D}_u \mathcal{T}(s, t)\, f = (\alpha_1\, \partial_s + \alpha_2\, \partial_t)\, \mathcal{T}(s, t)\, f$$
$$= (\alpha_1 \mathcal{A}_1 + \alpha_2 \mathcal{A}_2)\, \mathcal{T}(s, t)\, f$$
$$= \mathcal{A}(\alpha_1, \alpha_2)\, \mathcal{T}(s, t)\, f, \quad (75)$$

where the domain of the operator $\mathcal{A}_u = \alpha_1 \mathcal{A}_1 + \alpha_2 \mathcal{A}_2$ is dense in X for all $(\alpha_1, \alpha_2) \in \mathbb{R}_+^2$ [1, pages 405, 407, 409].

To ensure that pointwise derivatives with respect to the scale parameters exist for all points $(x, t;\ s, \tau) \in \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2$ given any smooth function f of compact support as input, we do also require the semi-group should for some $k > (N + 1)/2$ be C_1 continuous with respect to the L^2-based Sobolev norm $\| \cdot \|_{H^k(\mathbb{R}^N \times \mathbb{R}_+)}$ in the sense that for all connected regions $\Omega \in \mathbb{R}_+^2$ that shrink to zero in such a way that the maximum distance $\rho(\Omega)$ between a point in Ω and the origin tends to zero

$$\lim_{\rho(\Omega) \downarrow 0} \left\| \frac{\int_{(s,t) \in \Omega} \mathcal{T}(s, t)\, f\, d\Omega}{\int_{(s,t) \in \Omega} d\Omega} - f \right\|_{H^k(\mathbb{R}^N \times \mathbb{R}_+)} = 0 \quad (76)$$

should be required to hold for all smooth functions $f \in L^1(\mathbb{R}^N \times \mathbb{R}_+) \cap C^\infty(\mathbb{R}^N \times \mathbb{R}_+)$.

This algebraic structure constitutes a natural generalization of the corresponding framework for one-parameter semi-groups in Sect. 3.2; see Definition 11 and Lemma 12 in Appendix E for more formal statements.

Evolution properties over time t and scale s.   From this analysis we can in analogy with Lemma 2 conclude that the

spatio-temporal scale-space representation L should be differentiable with respect to scale s and time t and satisfy the following two differential equations

$$\partial_s L = \mathcal{A}_1 L, \tag{77}$$

$$\partial_t L = \mathcal{A}_2 L \tag{78}$$

for some infinitesimal generators $\mathcal{A}_1$ and $\mathcal{A}_2$. Specifically, the second equation reflects a time-recursive formulation where the temporal scales τ are now treated as an internal temporal memory buffer, which is not explicitly but only implicitly updated by the temporal evolution. In this respect, the relationship between time t and temporal scale τ is reversed compared to the relationship between space x and spatial scale s.

Our next step is to investigate how the notion of non-enhancement of local extrema affects the choice of possible infinitesimal generators $\mathcal{A}_1$ and $\mathcal{A}_2$ and thus the class of possible operators $\mathcal{T}_{s,t}$ with their associated admissible spatio-temporal scale-space representation L. A natural way to formulate the notion of non-enhancement of local extrema in a time-recursive spatio-temporal setting is as follows: If a point $(x_0, t_0; s_0, \tau_0)$ is a local maximum over the spatial coordinates $x \in \mathbb{R}^N$ and the temporal scale τ, then the directional derivative of L in an arbitrary (forward) direction u in (s, t)-space must be negative. Similarly, if the point is a local minimum, then the directional derivative must by positive. In other words:

$$\partial_u L \leq 0 \quad \text{at any non-degenerate local maximum}, \tag{79}$$

$$\partial_u L \geq 0 \quad \text{at any non-degenerate local minimum}. \tag{80}$$

Since the directional derivative $\partial_u L$ in a direction $u = (\alpha_1, \alpha_2)$ in (s, t)-space can be written

$$\partial_u L = \alpha_1 \, \partial_s L + \alpha_2 \, \partial_t L \tag{81}$$

and the sign condition on $\partial_u L$ is required to hold for all non-negative α_1 and α_2, it follows that we have to require that

$$\partial_s L \leq 0 \quad \text{and} \quad \partial_t L \leq 0$$
$$\quad \text{at any non-degenerate local maximum}, \tag{82}$$

$$\partial_s L \geq 0 \quad \text{and} \quad \partial_t L \geq 0$$
$$\quad \text{at any non-degenerate local minimum}. \tag{83}$$

see Definitions 13 and 14 in Appendix E for more detailed statements.

From a similar way of reasoning as in the proof of Theorem 5, we can then conclude that the infinitesimal generators $\mathcal{A}_1$ and $\mathcal{A}_2$ must correspond to linear combinations of first- and second-order derivatives, where the second-order derivatives are determined from positive semi-definite covariance matrices Σ_1 and Σ_2. In other words, the scale-space representation L should satisfy:

$$\partial_s L = \frac{1}{2} \nabla_{x,\tau}^T \left(\Sigma_1 \nabla_{x,\tau} L \right) - \delta_1^T \nabla_{x,\tau} L, \tag{84}$$

$$\partial_t L = \frac{1}{2} \nabla_{x,\tau}^T \left(\Sigma_2 \nabla_{x,\tau} L \right) - \delta_2^T \nabla_{x,\tau} L, \tag{85}$$

where in this setting, the gradient operator $\nabla_{x,\tau}$ contains derivatives with respect to both the spatial coordinates x and the temporal scale τ; see Lemma 15 in Appendix E for a proof.

If we want the spatial scale parameter s to be a pure spatial scale parameter, however, it is natural to require the first evolution equation for $\partial_s L$ to be independent of explicit derivatives with respect to τ; otherwise temporal phenomena would influence the definition of spatial scales. Thus, we reduce the first term in $\mathcal{A}_1$ to

$$\nabla_{x,\tau}^T \Sigma_1 \nabla_{x,\tau} L = \nabla_x^T \Sigma_1 \nabla_x, \tag{86}$$

where the restricted gradient operator ∇_x corresponds to differentiation with respect to the spatial coordinates x only. To avoid an unessential translation with respect to the spatial domain, we can apply similar reflection symmetry arguments as in Sect. 4.1 to set $\delta_1 = 0$.

Similarly, if we want the temporal scale parameter τ to correspond to a pure temporal scale parameter, it is natural to require those second-order terms in the second evolution equation that depend on explicit derivatives with respect to x to be zero. Then, only one second-order term with derivatives with respect to τ remains non-zero, leading to a single term of the form $\frac{1}{2} \partial_{\tau\tau} L$. Concerning the first-order terms, we do not want the temporal evolution to be dependent on the temporal history and thus not dependent on the temporal scale τ. Therefore, we set the first-order term with respect to τ to zero, implying that the first-order term will only contain partial derivatives with respect the spatial coordinates x. What will remain of $\mathcal{A}_2$ will therefore be an operator of the form

$$\mathcal{A}_2 L = \frac{1}{2} \partial_{\tau\tau} L - \delta_2^T \nabla_x L, \tag{87}$$

where we in a moment will rename the x component of δ_2 into v. To conclude, we have shown that given the requirement of non-enhancement of local extrema and complementary regularity and symmetry requirements, the spatio-temporal scale-space representation should satisfy the following evolution equations:

$$\partial_s L = \frac{1}{2} \nabla_x^T (\Sigma \nabla_x L), \tag{88}$$

$$\partial_t L = -v^T \nabla_x L + \frac{1}{2} \partial_{\tau\tau} L. \tag{89}$$

Definition 16 and Theorems 17–18 in Appendix E give precise statements of the corresponding necessity and sufficiency results. Hence, for an original signal of dimensionality $N + 1$, the time-causal scale-space representation will (at least) comprise $N + 3$ dimensions.[7]

In terms of integral expressions, it can be shown[8] that the solutions of these equations with initial condition $L(x, 0; 0, \tau; \Sigma, v) = 0$ and combined boundary and initial condition $L(x, t; 0, 0; \Sigma, v) = f(x, t)$ can be written

$$L(x, t; s, \tau; \Sigma, v)$$
$$= \int_{u=0}^{t} \int_{\xi \in \mathbb{R}^N} f(\xi, u) h(x - \xi, t - u; s, \tau; \Sigma, v) \, d\xi \, du, \tag{90}$$

where the notation with double semi-colons in the list of variables indicates that s and τ are parameters while Σ and v are meta-parameters. The convolution kernel h is in turn given by

$$h(x, t; s, \tau; \Sigma, v) = g_N(x - vt; s; \Sigma) \phi(t; \tau), \tag{91}$$

where

$$g_N(x - vt; s; \Sigma)$$
$$= \frac{1}{(\sqrt{2\pi s})^N \sqrt{\det \Sigma}} e^{-(x-vt)^T \Sigma^{-1}(x-vt)/2s}, \tag{92}$$

$$\phi(t; \tau) = \frac{1}{\sqrt{2\pi} \, t^{3/2}} \tau \, e^{-\tau^2/2t}. \tag{93}$$

This form of time-causal spatio-temporal scale-space has also been derived by Fagerström [17] in the special case when $\Sigma = I$, however, starting from different arguments of scale invariance. The additional degree of freedom in the spatial covariance matrix Σ obtained here has the additional advantage that it allows for non-isotropic smoothing kernels over the spatial domain, which may be useful when dealing with local image deformations over time and when considering motion boundaries.

Figure 9 shows examples of the kernel h with spatio-temporal derivatives computed from it in the space-time separable case with $v = 0$. Figure 10 shows corresponding velocity-adapted kernels for non-zero velocities v with velocity-adapted temporal derivatives computed from a linear combination of temporal derivatives and spatial derivatives

$$\partial_{\bar{t}} = \partial_t + v^T \nabla_x. \tag{94}$$

Before proceeding with detailed analysis of this scale-space it can be mentioned that it will be shown in Sect. 5.1.4 that the spatial extent of these kernels is determined by the spatial covariance matrix Σ, while the temporal extent is proportional to τ^2.

5.1 Properties of the Time-Causal Spatio-Temporal Scale-Space

We can note that there are many structural similarities between this time-causal spatio-temporal scale-space concept and the previously considered Gaussian spatio-temporal scale-space. First of all, due to the linearity, the property of non-enhancement of local extrema carries over to any spatio-temporal derivative. Hence, all spatio-temporal derivatives satisfy corresponding scale-space properties as the original scale-space. Furthermore, with the evolution derivatives $\partial_s L$ and $\partial_t L$ over spatial scales s and over time t given by (88) and (89), it holds that the directional derivative $\partial_u L$ in any direction $u = (\alpha, \beta)$ in (s, t) space (according to (81)) can be written
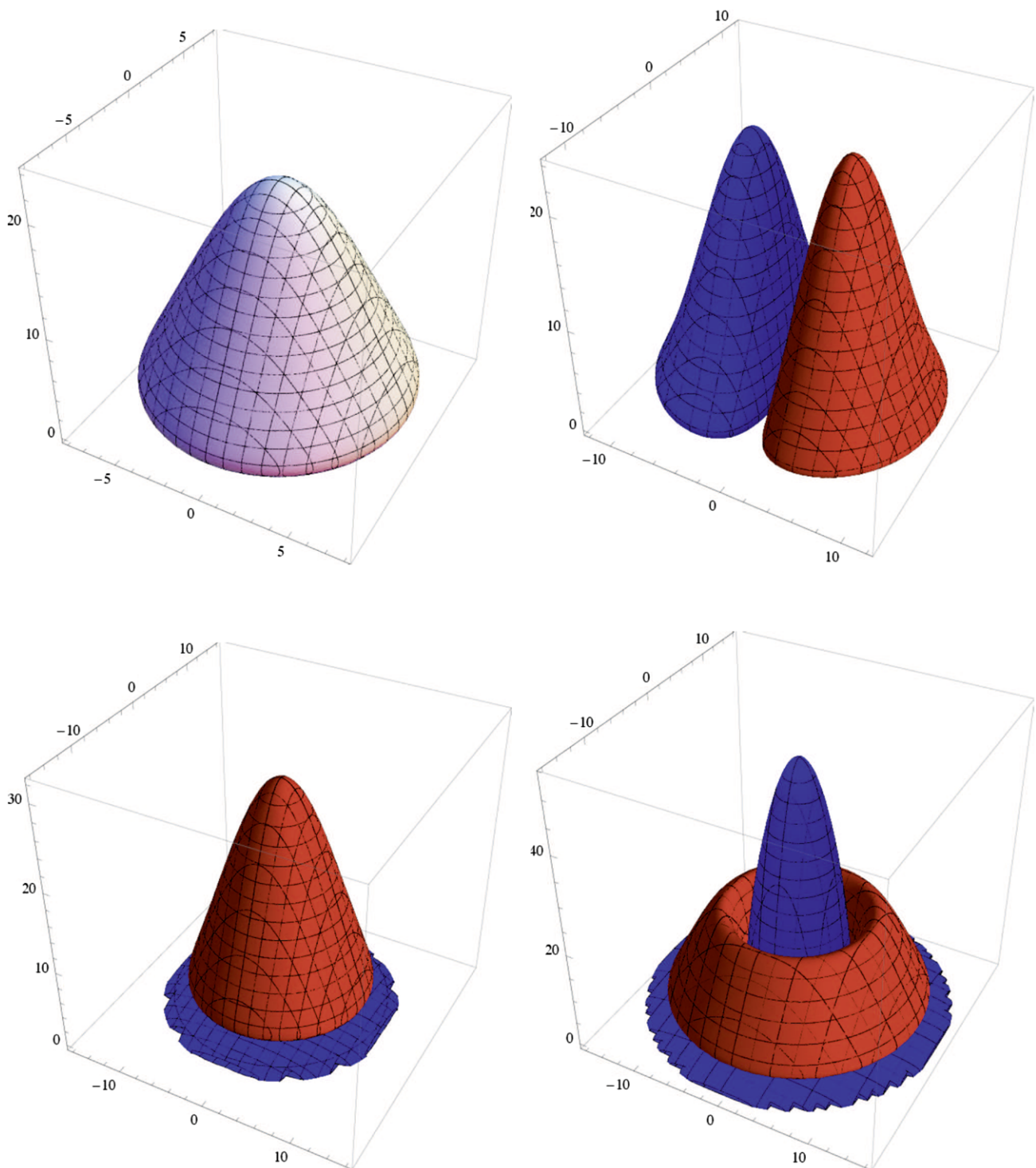
$$\partial_u L = \alpha \, \partial_s L + \beta \, \partial_t L$$
$$= \frac{\alpha}{2} \nabla_x^T \Sigma \nabla_x L + \frac{\beta}{2} \partial_{\tau\tau} L - \beta \, v^T \nabla_x L. \tag{95}$$

Thus, there is a very close similarity between these equations and the differential equations (44) governing the previously considered Gaussian spatio-temporal scale-space concept.

With regard to temporal causality, which is necessary in a real-time setting, it follows from (89) in combination with (94) that velocity-adapted temporal derivatives of L, i.e., $\partial_{\bar{t}}^n L$, can be computed from derivatives over temporal scales (i.e., the internal temporal memory) $\partial_{\tau\tau}^n L$ according to

$$\partial_{\bar{t}} = \frac{1}{2} \partial_{\tau\tau}, \tag{96}$$

thus without need for any additional temporal buffering than the information already contained in the time-causal spatio-temporal scale-space.

---

[7]If the full group of spatial covariance matrices Σ and velocity vectors v is considered as well, the dimensionality of the affine- and velocity-adapted scale-space will be $\dim(x) + \dim(t) + \dim(\Sigma) + \dim(v) + \dim(\tau) = N + 1 + N(N+1)/2 + N + 1 = (N^2 + 5N + 4)/2$. To handle such high-dimensional scale-spaces in practice, some sorts of intelligent search strategies are obviously required, such as combinations of lower-dimensional subgroups. The shape adaptation and velocity adaptation algorithms constitute examples of such simplifying search strategies. With a massively parallel architecture, such as in biological vision, however, one could afford to represent a richer family of affine-adapted and/or velocity-adapted filters than would be possible to handle with a serial computer. We will return to this subject, when we describe relations to biological vision in Sect. 6.

[8]This result follows from the fact that $g_N(x - vt; s; \Sigma, \delta)$ in (45) satisfies the differential equation (44), which (with $\delta = 0$) implies that $g_N(x - vt; s; \Sigma) \phi(t; \tau)$ in (91) satisfies (88). Similarly, since $\phi(t; \tau)$ according to (205) is a solution of (201), it follows that $g_N(x - vt; s; \Sigma) \phi(t; \tau)$ is a solution of (89); see Lindeberg [48].

Fig. 9 *Time-causal and space-time separable spatio-temporal scale-space kernels:* (*top left*) Original smoothing kernel $h(x, y, t; \ \Sigma, v, \tau)$; (*top right*) First-order spatial derivative $h_x(x, y, t; \ \Sigma, v, \tau)$; (*bottom left*) First-order temporal derivative $h_t(x, y, t; \ \Sigma, v, \tau)$; (*bottom right*) First-order temporal derivative of the spatial Laplacian $h_{xxt}(x, y, t; \ \Sigma, v, \tau) + h_{yyt}(x, y, t; \ \Sigma, v, \tau)$ ($\lambda_1 = \lambda_2 = 16$, $\tau = 2$, $v_x = v_y = 0$). (*Bottom plane*: space (x, y); *Vertical axis*: time t.)
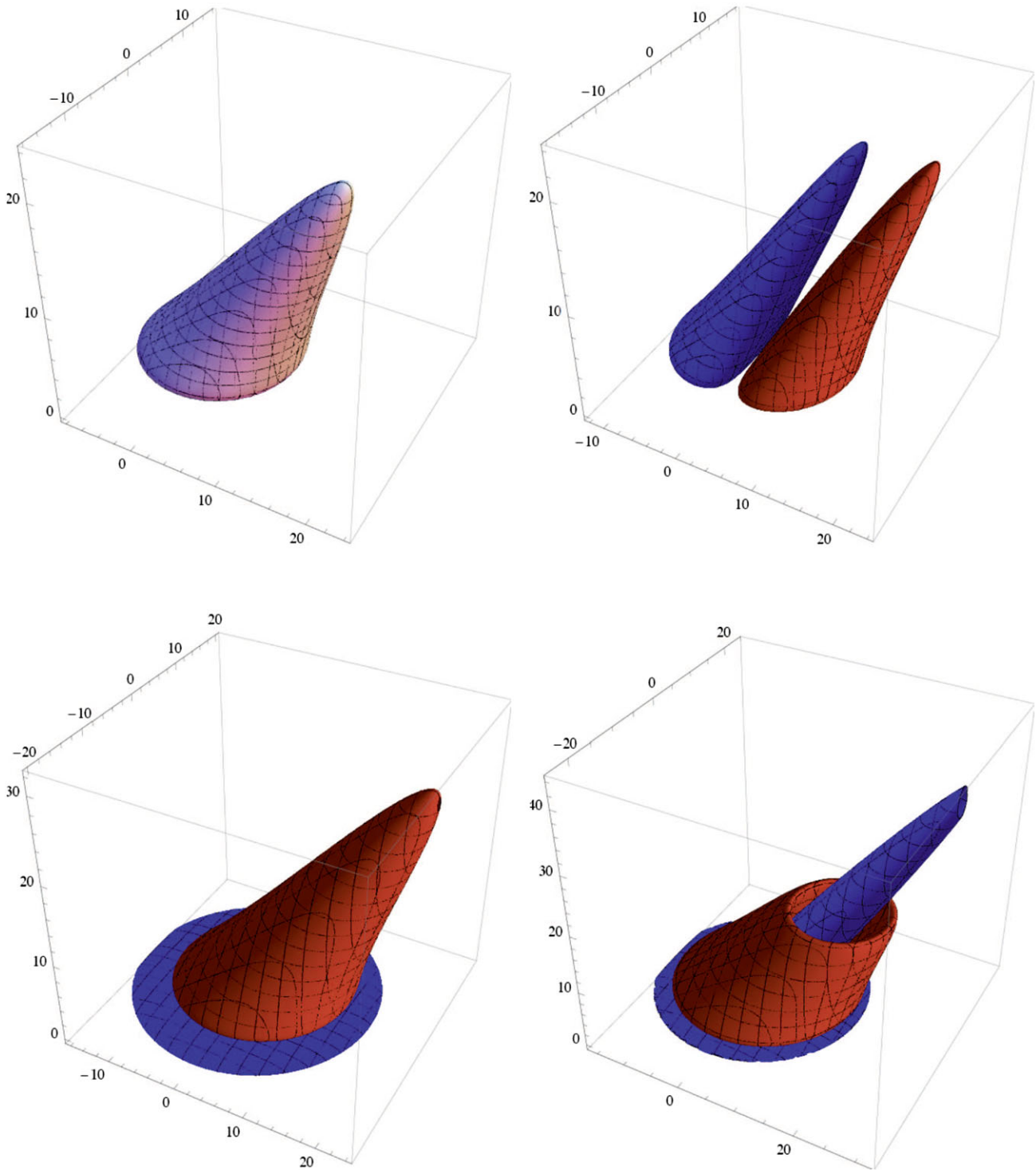
Fig. 10 *Time-causal and velocity-adapted (non-separable) spatio-temporal scale-space kernels:* (*top left*) Original smoothing kernel $h(x, y, t; \; \Sigma, v, \tau)$; (*top right*) First-order spatial derivative $h_x(x, y, t; \; \Sigma, v, \tau)$; (*bottom left*) First-order velocity-adapted temporal derivative $h_{\bar{t}}(x, y, t; \; \Sigma, v, \tau)$; (*bottom right*) First-order velocity-adapted temporal derivative of the spatial Laplacian $h_{xx\bar{t}}(x, y, t; \; \Sigma, v, \tau) + h_{yy\bar{t}}(x, y, t; \; \Sigma, v, \tau)$ ($\lambda_1 = \lambda_2 = 16$, $\tau = 2$, $v_x = 3/4$, $v_y = 0$). (*Bottom plane*: space (x, y); *Vertical axis*: time t.)

5.1.1 Relations to Regular Gaussian Smoothing

We can note that there is also a very close link to regular Gaussian smoothing. By inspection, it can be seen that the time-causal spatio-temporal smoothing can be interpreted as a first-order derivative with respect to temporal scale τ of a one-dimensional Gaussian over temporal scales, *i.e.,*

$$\phi(t; \tau) = -\partial_\tau g(\tau; t), \tag{97}$$

and an N-dimensional Galilean-transformed affine Gaussian kernel

$$g_N(x - vt; \Sigma) = \frac{1}{(\sqrt{2\pi})^M \sqrt{\det \Sigma}} e^{-(x-vt)^T \Sigma^{-1}(x-vt)/2} \tag{98}$$

over space x. For sake of convenience, we will henceforth change to the following notation:

$$L(x, t; \Sigma, v, \tau)$$
$$= \int_{u=0}^{t} \int_{\xi \in \mathbb{R}^N} f(\xi, u) h(x - \xi, t - u; \Sigma, v, \tau) \, d\xi \, du, \tag{99}$$

where

$$h(x, t; \Sigma, v, \tau) = g_N(x - vt; \Sigma) \phi(t; \tau), \tag{100}$$

$$g_N(x; \Sigma) = \frac{1}{(\sqrt{2\pi})^N \sqrt{\det \Sigma}} e^{-x^T \Sigma^{-1} x/2}, \tag{101}$$

$$\phi(t; \tau) = -\partial_\tau g_1(\tau; t) = \frac{1}{\sqrt{2\pi} \, t^{3/2}} \tau \, e^{-\tau^2/2t}. \tag{102}$$

Please, note the shift of the order of the arguments between ϕ and g_1 in (102).

5.1.2 Transformation Properties

This velocity- and shape-adapted spatio-temporal scale-space concept is closed under

– rescalings of the spatial and temporal dimensions,
– Galilean transformations in space-time and
– affine transformations in the spatial domain.

Scaling transformations over space and/or time. To verify the first one of these properties, let us rescale the spatial and temporal domains by scaling factors a and b, *i.e.,* given a spatio-temporal signal $f(x, t)$ introduce a rescaled signal

$$f'(x', t') = f(x, t) \quad \text{with } x' = a x \text{ and } t' = b t, \tag{103}$$

where a and b are non-zero scalar entities. Then, with

$$\Sigma' = a^2 \Sigma, \quad v' = \frac{b}{a} v \quad \text{and} \quad \tau' = \sqrt{b} \, \tau \tag{104}$$

the time-causal spatio-temporal scale-space representations of f and f' are related according to

$$L'(x', t'; \Sigma', v', \tau') = L(x, t; \Sigma, v, \tau) \tag{105}$$

with L given by (99) and L' defined by

$$L'(x', t'; \Sigma', v', \tau')$$
$$= \int_{u'=0}^{t'} \int_{\xi' \in \mathbb{R}^N} f'(\xi', u') \, g_N(x' - v't'; \Sigma')$$
$$\phi(t'; \tau') \, d\xi' \, du'. \tag{106}$$

Galilean transformations in space-time. Concerning the Galilean transformation property, let us next given any spatio-temporal signal $f(x, t)$ and any velocity vector w define a Galilean transformed signal by

$$f''(x'', t'') = f(x, t), \quad \text{where}$$
$$x'' = x - w t \text{ and } t'' = t. \tag{107}$$

Then, provided that

$$\Sigma'' = \Sigma, \quad v'' = v + w \quad \text{and} \quad \tau'' = \tau \tag{108}$$

the spatio-temporal scale-space representations of f and f'' are also equal

$$L''(x'', t''; \Sigma'', v'', \tau'') = L(x, t; \Sigma, v, \tau). \tag{109}$$

Affine transformation over the spatial domain. Finally, concerning the affine transformation property, it follows from (51) and (52) that if we have two spatio-temporal signals f''' and f that are related by $f'''(x''', t''') = f(x, t)$ with $x''' = Ax$ and $t''' = t$ and where A is a non-singular $N \times N$ matrix, then the time-causal spatio-temporal scale-space representations of f and f''' are related according to

$$L'''(x''', t'''; \Sigma''', v''', \tau''') = L(x, t; \Sigma, v, \tau) \tag{110}$$

if

$$\Sigma''' = A \Sigma A^T, \quad v''' = v \quad \text{and} \quad \tau''' = \tau. \tag{111}$$

General implications. To conclude, these results show that this spatio-temporal scale-space satisfies natural transformation properties that allow it to handle

– image data acquired with different spatial and/or temporal sampling rates,
– image structures of different spatial and/or temporal extent,
– objects with different distances from the camera,
– the linear component of relative motions between the camera and objects in the world, and
– the linear component of perspective deformations.

Similar properties hold also for the Gaussian spatio-temporal scale-space in Sect. 4.1.3.

5.1.3 Temporal Cascade-Recursive Formulation

When computing a spatio-temporal scale-space representation at time $t_2 > t_1$, a very attractive property is if this can be done in a *time-recursive* manner, such that it sufficient to use the following sources of information:

– the internal buffer of the spatio-temporal scale-space representation L at time t_1, and
– information about the spatio-temporal input data f during the time interval $[t_1, t_2]$.

This property means that it is sufficient to use the internal states of the spatio-temporal scale-space representation as internal memory, *and we do not need to have any complementary buffer of what else has occurred in the past.*

Such a property can indeed be established for the time-causal scale-space representation, based the fact that the time-causal scale-space kernel $\phi(t; \tau)$ satisfies the following time-recursive cascade smoothing property over a pure temporal domain (derived in (210) and (214) in Appendix D.3)

$$\phi(t_2; \tau) = \int_{\zeta=0}^{\infty} \phi(t_1; \zeta)(g(\tau - \zeta; t_2 - t_1) - g(\tau + \zeta; t_2 - t_1)) \, d\zeta. \quad (112)$$

From this relation it follows that the time-causal spatio-temporal scale-space representation satisfies the following cascade-recursive structure over time t and spatial scales s:

$$L(x, t_2; s_2, \tau)$$
$$= \int_{\xi \in \mathbb{R}^N} \int_{\zeta \geq 0} T(x - \xi, t_2 - t_1; s_2 - s_1, \tau, \zeta)$$
$$L(\xi, t_1; s_1, \zeta) \, d\zeta \, d\xi$$
$$+ \int_{\xi \in \mathbb{R}^N} \int_{u=t_1}^{t_2} B(x - \xi, t_2 - u; s_2, \tau)$$
$$f(\xi, u) \, d\xi \, du, \quad (113)$$

where the kernel T for updating the internal memory representation L is given by

$$T(x, t; s, \tau, \zeta) = g_N(x - vt; s)(g(\tau - \zeta; t) - g(\tau + \zeta; t)) \quad (114)$$

and the kernel B for incorporating new knowledge from the input signal f at the boundary is

$$B(x, t; s, \tau) = g_N(x - vt; s) \phi(t; \tau). \quad (115)$$

Please, note that we have here dropped the arguments for the meta-parameters Σ and v in order to simplify the notation.

5.1.4 Properties of the Time-Causal Smoothing Functions

To describe the evolution properties over temporal scales τ is however somewhat different than for the Gaussian spatio-temporal scale-space. While the integral of h over space-time is finite

$$\int_{t=0}^{\infty} \int_{x \in \mathbb{R}^N} h(x, t; \Sigma, v, \tau) \, dx \, dt = 1, \quad (116)$$

we cannot compute the first- and second-order moments of h over time t, since the corresponding integrals do not converge

$$\int_{t=0}^{\infty} \int_{x \in \mathbb{R}^N} t \, h(x, t; \Sigma, v, \tau) \, dx \, dt \to \infty, \quad (117)$$

$$\int_{t=0}^{\infty} \int_{x \in \mathbb{R}^N} t^2 \, h(x, t; \Sigma, v, \tau) \, dx \, dt \to \infty. \quad (118)$$

Hence, we cannot parameterize the time-causal kernels h in terms of mean vectors and covariance matrices over space-time, as is a natural approach for most other spatio-temporal scale-spaces. Nevertheless, we can compute the position in space-time of the local maximum of $h(x, t; \Sigma, v, \tau)$

$$\begin{pmatrix} \hat{x} \\ \hat{t} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} v \\ 1 \end{pmatrix} \tau^2 \quad (119)$$

and make another definition of the effective temporal extent of the time-causal kernel from the values $t_{t/2,1} < t_{1/2,2}$ for which the one-dimensional purely temporal time-causal kernel $\phi(t; \tau)$ assumes half its maximum value

$$\phi(t_{t/2,1}; \tau) = \phi(t_{t/2,2}; \tau) = \frac{1}{2} \phi(\hat{t}; \tau) \quad (120)$$

with the approximate estimates

$$t_{t/2,1} \approx 0.145 \, \tau^2, \qquad t_{t/2,2} \approx 1.046 \, \tau^2 \quad (121)$$

that lead to the following measure of the *temporal width* of the one-dimensional time-causal scale-space kernel

$$\Delta t = t_{t/2,2} - t_{t/2,1} \approx 0.900 \, \tau^2. \quad (122)$$

We can also define the *temporal delay* according to

$$\delta = \frac{\int_{t=0}^{\infty} t \, \phi^2(t; \tau) \, dt}{\int_{t=0}^{\infty} \phi^2(t; \tau) \, dt} = \tau^2. \quad (123)$$

Both the temporal width and the temporal delay of the time-causal scale-space kernel are hence proportional to the *square* of the temporal scale parameter τ (see Figs. 11–12 for illustrations).

To visualize the temporal response properties of the one-dimensional time-causal kernel $\phi(t; \tau)$, we can also compute the response to a step function $f_{step}(t) = H(t) = 1$ for
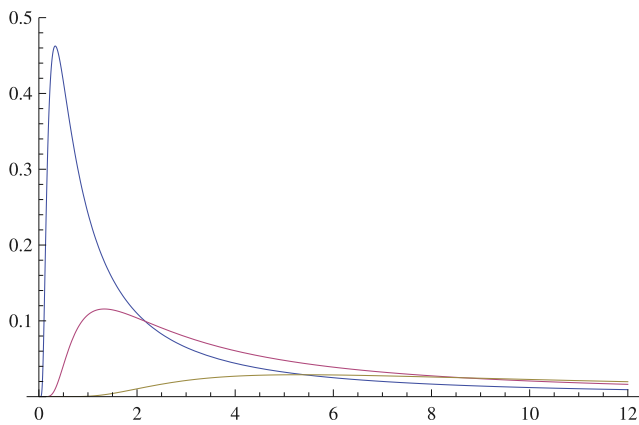
Fig. 11 Graphs of the one-dimensional time-causal scale-space kernels $\phi(t; \tau)$ for $\tau = 1, 2$ and 4
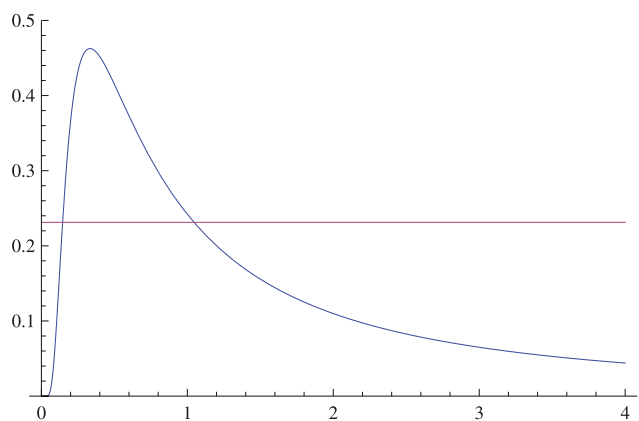
Fig. 12 Illustration of the definition of the *width* of the time-causal kernel from the values $t_{1/2,1} < t_{1/2,2}$ for which the time-causal kernel (here with $\tau = 1$) assumes half its maximum value

$t > 0$ and $f_{step}(t) = H(t) = 0$ for $t < 0$

$$L_{step}(t; \tau) = \text{erfc}\left(\frac{\tau}{\sqrt{2t}}\right) \tag{124}$$

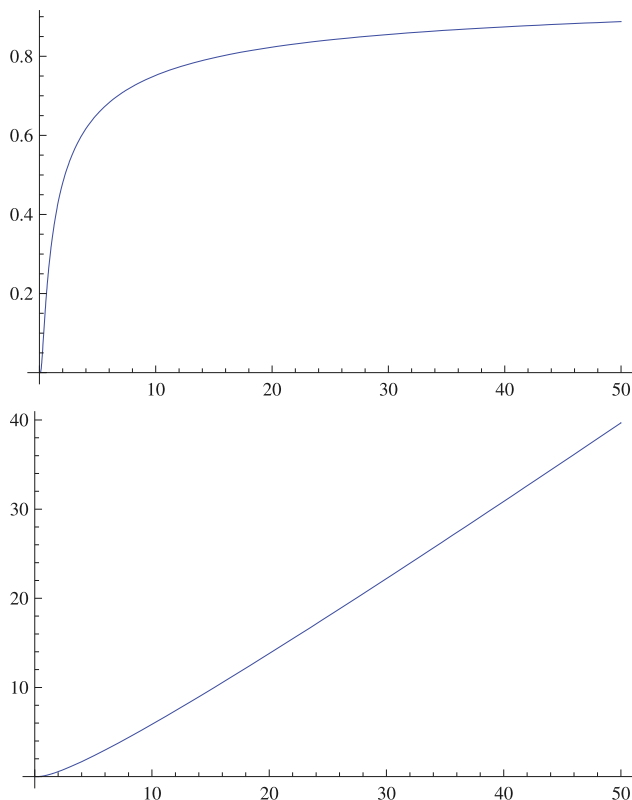and to a linear ramp $f_{ramp}(t) = t$ (see Fig. 13)

$$L_{ramp}(t; \tau) = (t + \tau^2)\,\text{erfc}\left(\frac{\tau}{\sqrt{2t}}\right) - e^{-\frac{\tau^2}{2t}}\sqrt{\frac{2}{\pi}}\,\tau\,\sqrt{t}. \tag{125}$$

We can also compute the spatial mean $\bar{x}$ and the spatial co-variance matrix $C(x, x)$ of the composed spatio-temporal kernel $h(x, t; \Sigma, v, \tau)$ according to

$$\bar{x} = \frac{\int_{x \in \mathbb{R}^N} x\, h(x, t; \Sigma, v, \tau)\, dx}{\int_{x \in \mathbb{R}^N} h(x, t; \Sigma, v, \tau)\, dx} = vt, \tag{126}$$

$$C(x, x) = \frac{\int_{x \in \mathbb{R}^N} x x^T h(x, t; \Sigma, v, \tau)\, dx}{\int_{x \in \mathbb{R}^N} h(x, t; \Sigma, v, \tau)\, dx} - \bar{x}\bar{x}^T = s\,\Sigma. \tag{127}$$

Fig. 13 The response dynamics of the one-dimensional time-causal scale-space kernel $\phi(t; \tau)$ to (*top*) a unit step function and (*bottom*) a linear ramp at temporal scale $\tau = 1$

In other words, (i) the spatial shape of the spatio-temporal kernel $h(x, t; \Sigma, v, \tau)$ is described by the spatial covariance matrix Σ, (ii) the temporal extent is proportional to τ^2 and (iii) the velocity vector v specifies the orientation of the kernel in space-time.

5.1.5 Behaviour over Temporal Scales τ

Under variations of the temporal scale parameter τ, it can be noted that the explicit expression for the time-causal kernel $\phi(t; \tau)$ is self-similar with regard to the ratio $\tau/\sqrt{t}$. This means that the time-causal kernels will be *self-similar over temporal scales*, as reflected in the scaling property in (104)–(105). With the Laplace transform of the purely temporal part of the time-causal kernel defined by

$$(\mathcal{L}\phi)(p; \tau) = \bar{\phi}(p; \tau) = \int_{t=0}^{\infty} \phi(t; \tau)\, e^{-pt} dt$$

$$= \int_{t=0}^{\infty} \frac{1}{\sqrt{2\pi}\, t^{3/2}}\, \tau\, e^{-\tau^2/2t}\, e^{-pt} dt$$

$$= e^{-\sqrt{2p}\,\tau}, \tag{128}$$

it can be seen that the result of multiplying two such Laplace transforms is of the form
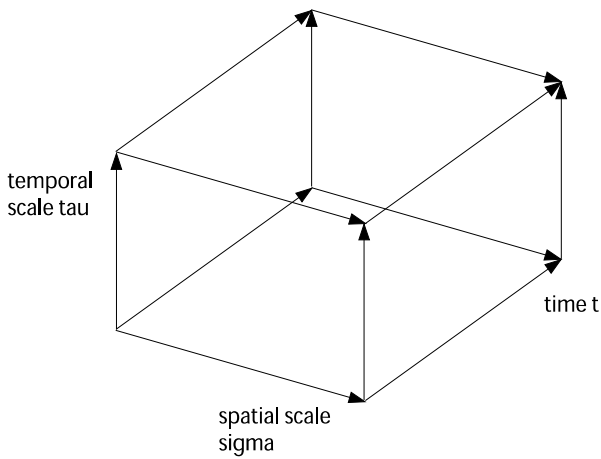
Fig. 14 The spatio-temporal scale-space L representation of an image is a function over space $x \in \mathbb{R}^N$, time $t \in \mathbb{R}_+$, spatial scale $s \in \mathbb{R}_+$ and temporal scale $\tau \in \mathbb{R}_+$, with a spatial covariance matrix $\Sigma \in \mathbb{R}^{N^2}$ and an image velocity vector $v \in \mathbb{R}^N$ as meta parameters. With the temporal scale parameter τ treated as an internal state variable in addition to (the here invisible) space x, we can describe this spatio-temporal scale-space representation as a two-parameter semi-group that obeys non-enhancement of local extrema over spatial scale s and time t, and with a both time-causal and time-recursive update rule over time t. If we on the other hand treat space-time (x, t) as the internal state, we obtain another two-parameter semi-group over spatial scales s and temporal scales τ, which obeys non-enhancement of local extrema with increasing spatial scales s while only scale invariance over temporal scales τ

$$\bar{\phi}(p; \tau_1) \bar{\phi}(p; \tau_2) = e^{-\sqrt{2p}\,\tau_1} e^{-\sqrt{2p}\,\tau_2} = e^{-\sqrt{2p}\,(\tau_1+\tau_2)}$$
$$= \bar{\phi}(p; \tau_1 + \tau_2) \qquad (129)$$

corresponding to the linear semi-group structure of $\phi(t; \tau)$ under additions of the temporal scale parameter τ. Thus, the time-causal kernels also form a semi-group over temporal scales with regard to one-sided and finite support convolution operations

$$\phi(\cdot; \tau_1) \tilde{*} \phi(\cdot; \tau_2) = \phi(\cdot; \tau_1 + \tau_2) \qquad (130)$$

of the form

$$(f \tilde{*} g)(t) = \int_{u=0}^{t} f(u)\, g(t-u)\, du$$
$$= \int_{u=0}^{t} f(t-u)\, g(u)\, du. \qquad (131)$$

This property carries over to a corresponding semi-group property of the time-causal spatio-temporal kernels under convolutions over space-time

$$(h(\cdot, \cdot; \Sigma_1, \tau_1) * h(\cdot, \cdot; \Sigma_2, \tau_2))$$
$$= h(\cdot, \cdot; \Sigma_1 + \Sigma_2, \tau_1 + \tau_2) \qquad (132)$$

which means that the spatio-temporal derivatives obey the following cascade smoothing property

$$L(\cdot, \cdot; \Sigma_2, \tau_2)$$
$$= h(\cdot, \cdot; \Sigma_2 - \Sigma_1, \tau_2 - \tau_1) * L(\cdot, \cdot; \Sigma_1, \tau_1). \qquad (133)$$

Along the direction of temporal scales τ, however, this semi-group does not obey non-enhancement of local extrema with increasing temporal scales, only scale invariance (see Fig. 14).

5.2 Temporal Derivatives with Respect to Transformed Time

When computing temporal derivatives of time-causal spatio-temporal kernels, the temporal derivative operator reduces to temporal derivatives of the one-dimensional time kernel $\phi(t; \tau)$. For space-time separable kernels with $v = 0$ we have

$$\partial_{x^\alpha t^\gamma} h(x, t; s, \tau) = \partial_{x^\alpha t^\gamma} (g(x; s)\, \phi(t; \tau))$$
$$= g_{x^\alpha}(x; s)\, \phi_{t^\gamma}(t; \tau) \qquad (134)$$

while for velocity-adapted kernels the corresponding velocity-adapted spatio-temporal derivatives are given by

$$\partial_{x^\alpha \bar{t}^\gamma} h(x, t; s, \tau, v) = \partial_{x^\alpha \bar{t}^\gamma} (g(x - vt; s)\, \phi(t; \tau))$$
$$= g_{x^\alpha}(x - vt; s)\, \phi_{t^\gamma}(t; \tau). \qquad (135)$$

Hence, we can reveal many of the temporal response properties of the composed spatio-temporal kernels $h(x, t; s, \tau)$ and $h(x, t; s, \tau, v)$ by studying the temporal derivatives of the one-dimensional time-causal kernel $\phi(t; \tau)$.

In this subsection, we shall follow this notion and extend the regular temporal derivative operator ∂_t by a transformed derivative operator $\partial_{t'}$ with respect to a transformed time. The motivation for this extension is that the time-causal kernels are highly asymmetric over time, which means that it may be more natural to consider temporal derivatives with respect to a transformed time axis

$$t' = \varphi(t), \qquad (136)$$

where φ should be a monotonically increasing function, we can in particular consider a self-similar logarithmic transformation
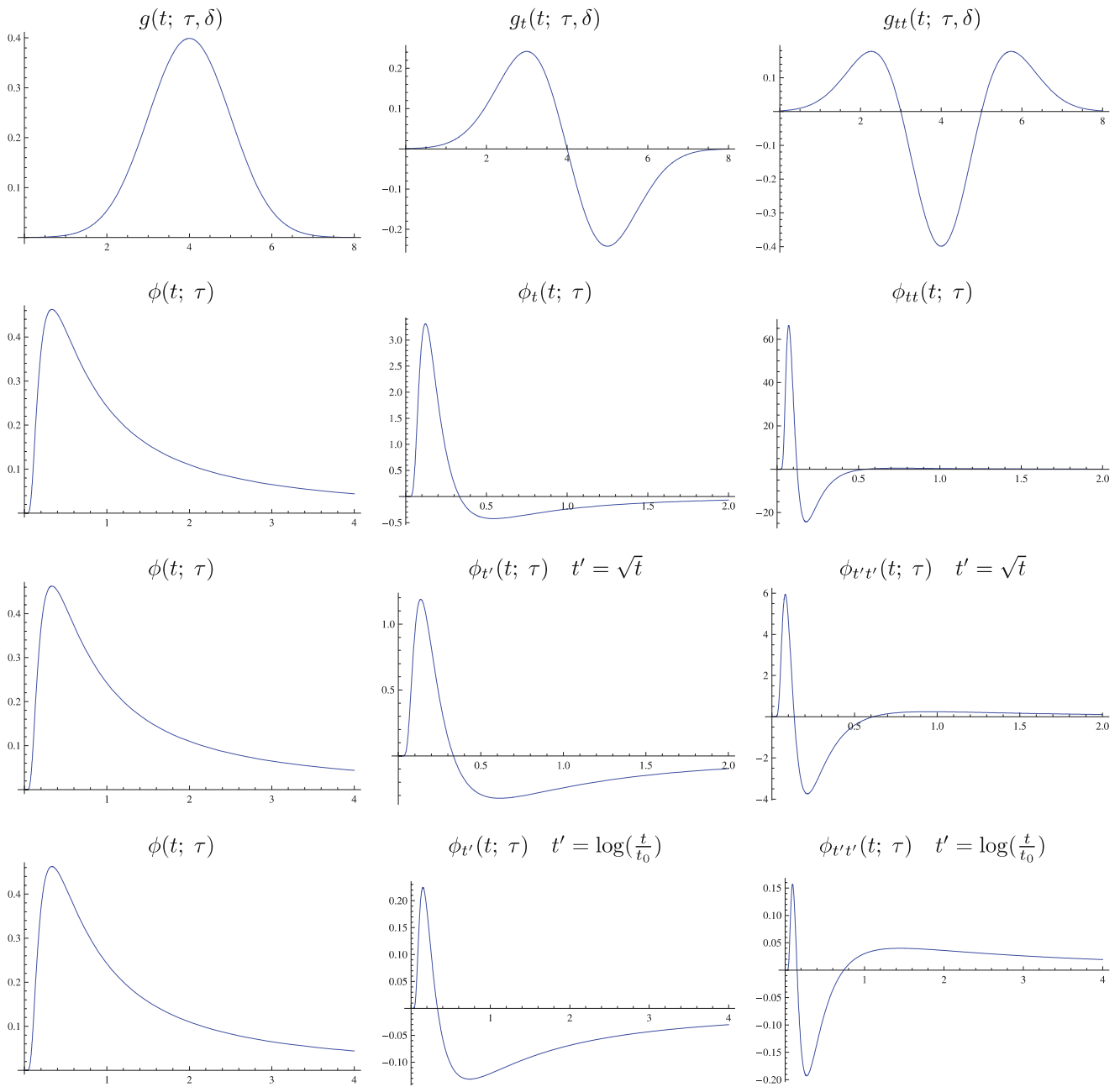
$$t' = \log\left(\frac{t}{t_0}\right) \qquad (137)$$

or a self-similar power law

$$t' = t^\alpha. \qquad (138)$$

The derivative operation $\partial_{t'}$ with respect to logarithmic time will then related to the regular temporal derivative operator ∂_t according to

Fig. 15 Temporal smoothing kernels that can be used in the modelling of space-time separable spatio-temporal receptive fields, with their first- and second-order derivatives displayed as well: (*top row*) Time-shifted Gaussian kernel $g(t; \tau, \delta) = 1/\sqrt{2\pi\tau} \exp(-(t-\delta)^2/2\tau)$ with $\tau = 1$ and $\delta = 4$; (*second row from top*) The time-causal kernel $\phi(t; \tau) = 1/\sqrt{2\pi t^3} \, \tau \exp(-\tau^2/2t)$ with $\tau = 1$; (*third row from* *top*) The time-causal kernel $\phi(t; \tau) = 1/\sqrt{2\pi t^3} \, \tau \exp(-\tau^2/2t)$ considered in the previous row, but with derivatives computed with respect to square root time $t' = \sqrt{t}$; (*bottom row*) The time-causal kernel $\phi(t; \tau) = 1/\sqrt{2\pi t^3} \, \tau \exp(-\tau^2/2t)$ considered in the previous row, but with derivatives computed with respect to logarithmic time $t' = \log(t/t_0)$ with $t_0 = 2$

$$\partial_{t'} = \frac{1}{\varphi'(t)} \partial_t \tag{139}$$

which for the logarithmic time transformation implies

$$\partial_{t'} = \frac{t}{t_0} \partial_t \tag{140}$$

and for the power law $t' = t^\alpha$

$$\partial_{t'} = \frac{t^{1-\alpha}}{\alpha} \partial_t. \tag{141}$$

The third and fourth rows in Fig. 15 show the result of computing first- and second-order temporal derivatives from the

Fig. 16 *Space-time separable kernels* $h_{x^\alpha t^\gamma}(x, t; s, \tau, v)$ up to order two obtained from the *time-causal spatio-temporal scale-space* in the case of a 1+1-D space-time ($s = 1$, $\tau = 2$) and with temporal derivatives computed with respect to a self-similar transformation of the temporal axis $\partial_{t'} \sim t^\kappa \partial_t$, here with $\kappa = 1/2$. (*Horizontal axis*: space; *Vertical axis*: time.)
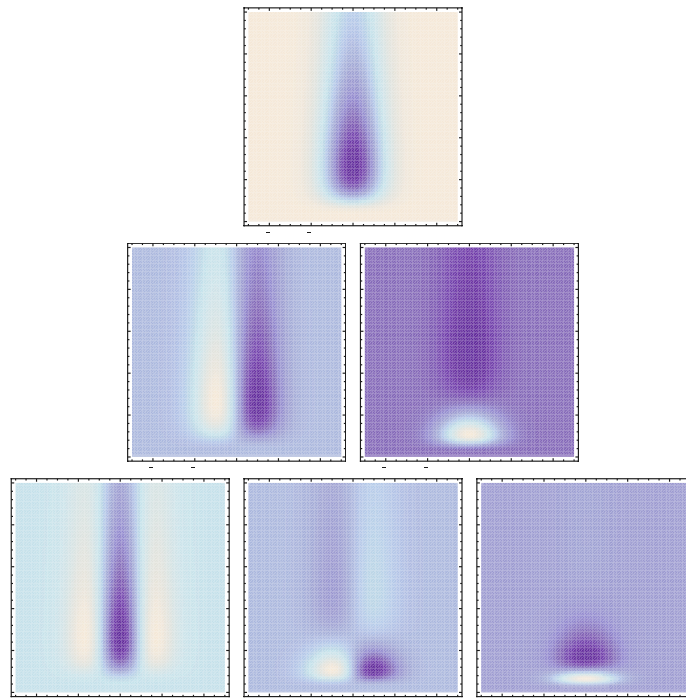
time-causal smoothing kernel in this way. For comparison, we also show derivatives of the Gaussian temporal kernel and regular temporal derivatives of the time-causal kernel.

As we can see, the change of temporal variable by a self-similar monotonically increasing transformation does not change the position of the zero-crossing for the first-order derivative; it only leads a multiplication by linear time-dependent scalar factor. For the second-order derivative, however, the behaviour is qualitatively different. The regular first-order derivative of the time-causal kernel has two peaks and one interior zero-crossing, while the second-order derivative with respect to transformed time has three clearly visible peaks and two internal zero-crossings, as a second-order scale-space derivative kernel should have.

For these reasons, we will henceforth consider this generalized notion of temporal derivatives with respect to transformed time when studying temporal derivative responses of highly asymmetric smoothing kernels. Indeed, for both the logarithmic transformation and for the power law the transformed derivative operator is of the form

$$\partial_{t'} \sim t^\kappa \partial_t, \tag{142}$$

and we will refer to this operator as *temporal derivatives with respect to self-similarly transformed time*. It can be shown that this definition is compatible with spatio-temporal scale invariance for scale selection based on local extrema over temporal scales of scale-normalized derivatives (manuscript in preparation). Specifically, the value $\kappa = 1/2$ can be motivated both from theoretical considerations and agreement with biological receptive fields. Figure 16 and 17

show corresponding spatio-temporal derivative operators for a 1+1-D space-time.

6 Relations to Biological Vision

In a comprehensive review, DeAngelis et al. [11] present an overview of temporal response properties of receptive fields in the central visual pathways. Foremost, the authors point out the limitations of defining receptive fields in the spatial domain only, and emphasize the need to characterize receptive fields in the *joint* space-time domain, in order to describe how a neuron processes the visual image. Then, for basic cell types in the LGN and the striate cortex, they essentially describe the spatio-temporal response characteristics as follows:

LGN neurons: The neurons in the LGN have approximately circular center-surround organization in the spatial domain (see Fig. 18(a)) and most receptive fields are separable in space-time. There are two main classes of temporal responses for such cells: In a "non-lagged cell" the first temporal lobe is the largest one (Fig. 20(a)) whereas for a "lagged cell" the second lobe dominates (Fig. 20(b)). Such temporal response properties are typical for first- and second-order temporal derivatives of a time-causal temporal scale-space representation (see Fig. 15).[9] The spatial re-

---

[9]For the first-order temporal derivative of a time-causal temporal scale-space kernel, the first peak is usually strongest, whereas for certain

Fig. 17 *Velocity-adapted spatio-temporal kernels* $h_{\bar{x}^\alpha \bar{t}^{\prime\gamma}}(x, t;\ s, \tau, v)$ up to order two obtained from the *time-causal spatio-temporal scale-space* in the case of a 1+1-D space-time ($s = 1$, $\tau = 2$, $v = 0.75$) and with temporal derivatives computed with respect to self-similar transformation of the temporal axis $\partial_{t'} \sim t^\kappa \partial_t$, here with $\kappa = 1/2$. (*Horizontal axis*: space; *Vertical axis*: time.)
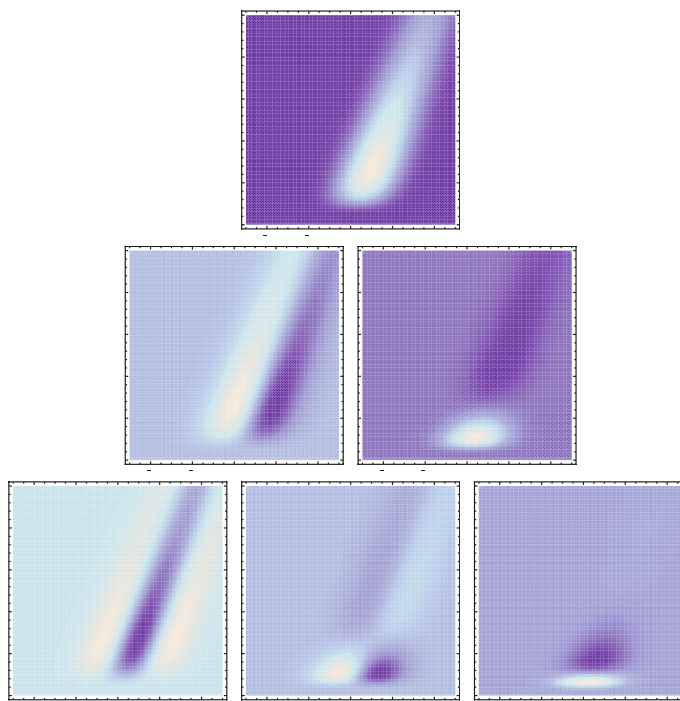
sponse, on the other hand, shows a high similarity to a Laplacian of a Gaussian.

Within the above mentioned spatio-temporal scale-space theory, we can model the qualitative shape of these circular center-surround receptive fields in the LGN as:

$$h_{LGN}(x, y, t;\ s, \tau)$$
$$= \pm (\partial_{xx} + \partial_{yy})\, g(x, y;\ s)\, \partial_{t'^n}\, h(t;\ \tau), \qquad (143)$$

where

- the sign determines whether the cell is of type "on-center-off-surround" or "off-center-on-surround",
- the parameter $n = 1, 2$ describes the order of differentiation with respect to time,
- $g(x, y;\ s)$ is an isotropic smoothing kernel in the spatial domain generated by the rotationally symmetric Gaussian scale-space concept with spatial scale parameter s,
- $h(t;\ \tau)$ is a temporal smoothing kernel over time t with temporal scale parameter τ,
- the operator $\partial_{t'}$ denotes differentiation with respect to a possibly transformed temporal axis according to a self-similar transformation $\partial_{t'} \sim t^\kappa \partial_t$, where $\kappa = 0$ corresponds to regular temporal derivatives, $\kappa = 1$ corresponds to the computation of temporal derivatives with respect to logarithmic time, and $\kappa \in\,]0, 1[$ corresponds to tempo-

classes of time-causal temporal smoothing kernel, the second peak may be the most dominant for second-order temporal derivatives (see Fig. 15).
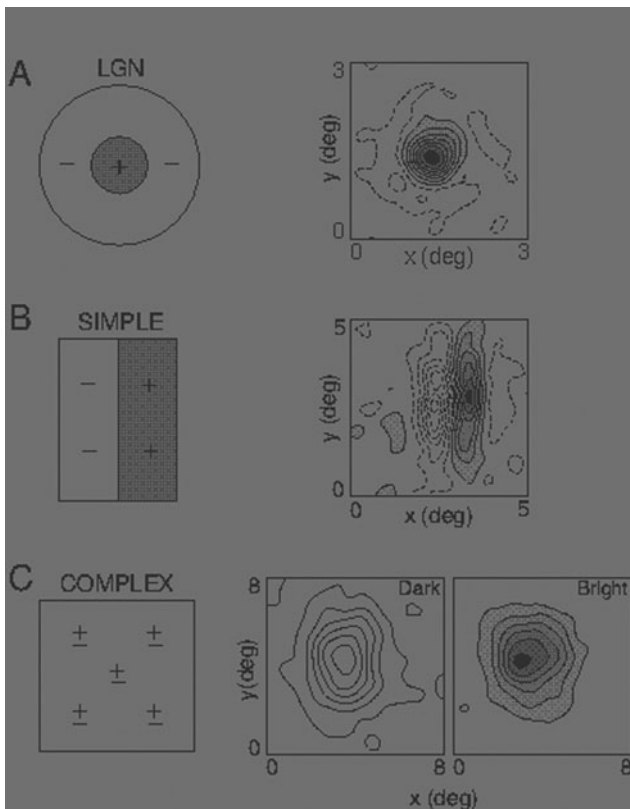
ral derivatives with respect to a power law transformation $t' = t^\alpha$ with $\kappa = 1 - \alpha$.

Figure 19(a) shows an illustration of the spatial response properties of such a receptive field. Figure 15 shows examples of different kernels that can be used for modelling the temporal smoothing component for this class of space-time separable filters.

Note: In all illustrations in Sect. 6, where spatial and spatio-temporal derivative expressions are aligned to biological data, the unit for the spatial scale parameter s corresponds to [degrees^2] of visual angle, the units for the temporal scale parameter τ in the Gaussian spatio-temporal scale-space representation is [milliseconds^2], while the units for the temporal scale parameter τ in the time-causal spatio-temporal scale-space representation is [$\sqrt{\text{milliseconds}}$]. For image velocities v of velocity-adapted filters, the units are [degrees/millisecond].

Simple cells: For simple cells in the striate cortex, the receptive fields are oriented in the spatial domain (see Fig. 18(b)). The *spatial component* of such cells can be modelled by directional derivatives of affine Gaussian kernels according to (58), (59) and (57); see Fig. 19(b):

$$h_{space}(x, y;\ s)$$
$$= (\cos\varphi\, \partial_x + \sin\varphi\, \partial_y)^m \left(\frac{1}{2\pi \sqrt{\det \Sigma_s}}\, e^{-x^T \Sigma_s^{-1} x / 2} \right). \qquad (144)$$

Fig. 19 (*left*) The Laplacian of an isotropic two-dimensional Gaussian smoothing kernel over a spatial domain $\nabla^2 g(x, y; \ s) = (x^2 + y^2 - 2s)/(2\pi s^3) \exp(-(x^2 + y^2)/2s)$ with here $s = 0.4$ can be used as a model for the circular center-surround responses in the LGN illustrated in Fig. 18(a). (*right*) First-order directional derivatives of anisotropic affine Gaussian kernels, here aligned to the coordinate directions $\partial_x g(x, y; \ \Sigma) = \partial_x g(x, y; \ \lambda_x, \lambda_y) = -\frac{x}{\lambda_x} 1/(2\pi \sqrt{\lambda_x \lambda_y}) \exp(-x^2/2\lambda_x - y^2/2\lambda_y)$ and here with $\lambda_x = 0.2$ and $\lambda_y = 2$, can be used as a model for simple cells with a strong directional preference as illustrated in Fig. 18(b)

Figures 21, 23, and 25 show a few examples of separable and inseparable kernels obtained in this way for a 1+1-dimensional space-time, based on the general models

$$h_{Gaussian}(x, t; \ s, \tau, v, \delta) = \partial_{\bar{x}}^m \partial_{\bar{t}}^n g(x, t; \ s, \tau, v, \delta), \quad (145)$$

$$h_{time\text{-}causal}(x, t; \ s, \tau, v) = \partial_{\bar{x}}^m \partial_{\bar{t}}^n h(x, t; \ s, \tau, v) \quad (146)$$

with space-time tilted spatio-temporal derivative operators $\partial_{\bar{x}} = \partial_x$ and $\partial_{\bar{t}} = v\, \partial_x + \partial_t$ and with the transformed self-similar temporal derivative operator according to $\partial_{t'} \sim t^\kappa \, \partial_t$.

Motion selectivity. Concerning motion selectivity, DeAngelis et al. [11] report that most cortical neurons are quite sensitive to stimulus velocity and the speed tuning is more narrow than for LGN cells. Simple cells with inseparable receptive fields have directional preference while cells with space-time separable receptive fields do not. Moreover, the preferred direction of motion corresponds to the orientation of the filter in space-time.

This structure is nicely compatible with velocity adaptation, as described in Sects. 4.1.3–4.1.4, 5 and Appendix C. Within the above mentioned terminology, separable receptive fields correspond to spatio-temporal scale-space kernels without velocity adaptation, while inseparable receptive fields correspond to kernels that are explicitly adapted to non-zero velocities.

The directional preference of the cells in the spatial domain can, in turn, be controlled by the covariance matrix of the affine Gaussian scale-space concept as outlined in Sect. 4.1.2. We obtain receptive fields without directional preference in the spatial domain if we set the covariance
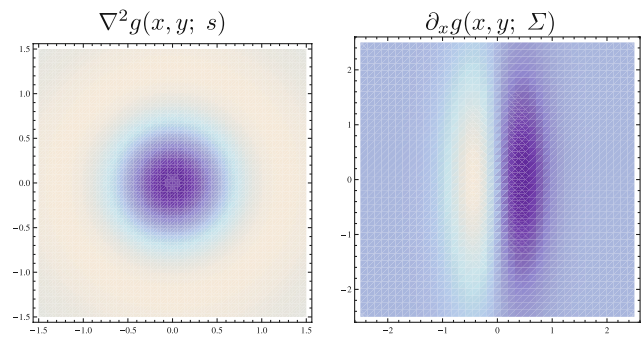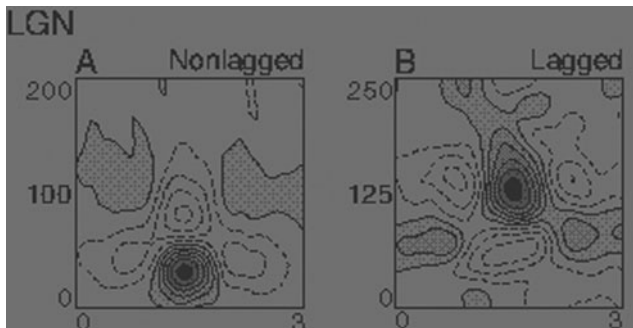

Fig. 18 Examples of receptive field profiles in the spatial domain as reported by DeAngelis et al. [11]. (**a**) Receptive fields in the LGN have approximately circular center-surround responses in the spatial domain. In terms of Gaussian derivatives, this spatial response profile can be modelled by the Laplacian of the Gaussian $\nabla^2 g(x; \ t)$ (see Fig. 19(a)). (**b**) Simple cells in the cerebral cortex do usually have strong directional preference in the spatial domain. In terms of Gaussian derivatives, this spatial response can be modelled as a directional derivative of an elongated affine Gaussian kernel (see Fig. 19(b)). (**c**) Complex cells are non-linear and do not obey the superposition principle

In the *joint space-time domain*, the spatio-temporal response properties range from separable (Fig. 22) to strongly inseparable (Fig. 24), where a majority exhibit marked space-time inseparability. The temporal profile is reported to be typically biphasic, although some cells are reported to have monophasic or triphasic responses.

In terms of temporal derivatives, a biphasic behaviour arises from first-order derivatives, a monophasic behaviour from zero-order derivatives and a triphasic behaviour from second-order derivatives. Concerning the oriented spatial response characteristics, there is a high similarity with directional derivatives of Gaussian kernels (Young [75]).

In fact, for all these linear receptive fields, spatio-temporal filters with qualitatively similar response characteristics can be generated by applying Cartesian or directional partial derivative operators of low orders to the spatio-temporal filters obtained from the spatio-temporal scale-space framework outlined in Sects. 4 and 5.

matrix $\Sigma = s\,I$ proportional to the unit matrix, and space-time separable receptive fields if we in addition choose the velocity adaptation vector v equal to zero. Assuming that the influence of Σ and v can be neglected (*e.g.* by setting Σ proportional to the unit matrix and v to zero), the filter shape will then be determined solely by the spatial scale s and the temporal scale λ. Conversely, we can construct inseparable kernels with strong directional preference by ap-

propriate combinations of the covariance matrix Σ and the velocity adaptation vector v.

The above mentioned fact that a majority of the cells are inseparable in space-time is indeed nicely compatible with a description in terms of a multi-parameter scale-space as outlined in Sect. 4. If the vision system is to give a reasonable coverage of a set of filter parameters Σ and v, then the set of filters corresponding to space-time separable receptive fields (corresponding to the filter parameters $v = 0$) will be much smaller than the set of filters allowing for non-zero values of the mixed parameters over space and time.

Complex cells.   Besides the above mentioned linear receptive fields, there is a large number of early non-linear receptive fields that do not obey the superposition principle and which are referred to as complex cells. The response profile of such a cell in the spatial domain is typically of the form illustrated in Fig. 18(c).

In their study of spatio-temporal receptive field properties, DeAngelis et al. [11] also report a large number of complex cells with non-linear response profiles in the joint space-time domain; see Fig. 26 for an example. Within the framework of the presented spatio-temporal scale-space concept, it is interesting to note that non-linear receptive fields with qualitatively similar properties can be constructed by squaring first- and second-order derivative responses and summing up these components [34]. Provided that the filters are appropriately normalized, we


Fig. 20 Examples of space-time separable receptive field profiles in the LGN as reported by DeAngelis et al. [11]. There are two main categories of such cells; (**a**) for a non-lagged cell, the first temporal lobe dominates, while (**b**) for a lagged cell the second temporal lobe is strongest. In terms of the spatio-temporal receptive field model presented in this paper, non-lagged cells can be modelled by first-order temporal derivatives, while the shape of lagged cells resembles second-order temporal derivatives (see Fig. 21). (*Horizontal dimension*: space x; *Vertical dimension*: time t.)

Fig. 21 Space-time separable kernels obtained from the spatio-temporal scale-space concepts: (*upper left*) Gaussian spatio-temporal kernel $g_{xxt}(x, t;\ s, \tau, \delta) = g_{xx}(x;\ s)\,g_t(t;\ \tau, \delta)$ with $s = 0.4$, $\tau = 30^2$, $\delta = 60$; (*upper right*) Gaussian spatio-temporal kernel $g_{xxtt}(x, t;\ s, \tau, \delta) = g_{xx}(x;\ s)\,g_{tt}(t;\ \tau, \delta)$ with $s = 0.3$, $\tau = 35^2$, $\delta = 120$; (*lower left*) Time-causal spatio-temporal kernel $h_{xxt'}h(x, t;\ s, \tau) = g_{xx}(x;\ s)\,\phi_{t'}(t;\ \tau, \delta)$ with $s = 0.4$, $\tau = 17$; (*lower right*) Time-causal spatio-temporal kernel $h_{xxt't'}h(x, t;\ s, \tau) = g_{xx}(x;\ s)\,\phi_{t't'}(t;\ \tau, \delta)$ with $s = 0.4$, $\tau = 25$. For the time-causal kernels, the temporal derivatives have been computed using the transformed temporal derivative operator $\partial_{t'} \sim t^\kappa\,\partial_t$, here with $\kappa = 1/2$. Compare the qualitative shapes of these kernels with the kernels in with Fig. 20. (*Horizontal dimension*: space x; *Vertical dimension*: time t.)
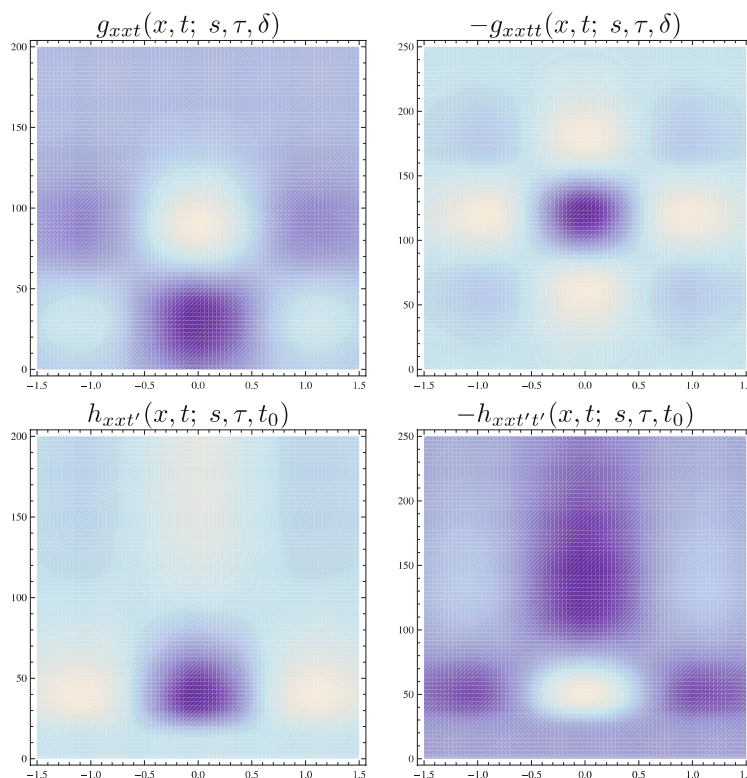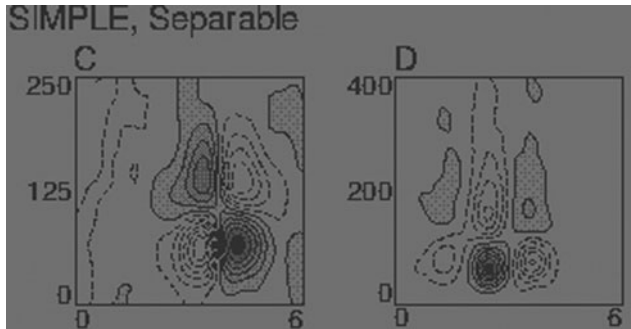
can then construct a quasi quadrature measure over a one-dimensional either spatial or temporal domain as [42]

$$QL = L_\xi^2 + C\,L_{\xi\xi}^2 = s L_x^2 + C\,s^2 L_{xx}^2, \qquad (147)$$

where $\partial_\xi = \sqrt{s}\,\partial_x$ denotes normalized derivatives with respect to normalized coordinates $\xi = x/\sqrt{s}$ [44] and where the constant C can be determined either to minimize the amount of ripples in the operator response ($C = 2/3 \approx 0.667$) or from scale selection properties ($C = e/4 \approx 0.670$). In the case of operating on a 1+1-D space-time with a Gaussian spatio-temporal scale-space, and with normalized


Fig. 22 Examples of space-time separable receptive field profiles in the striate cortex as reported by DeAngelis et al. [11]: (**a**) a non-lagged cell reminiscent of a first-order temporal derivative in time and a first-order derivative in space (compare with Fig. 23(a)) (**b**) a non-lagged cell reminiscent of a first-order temporal derivative in time and a second-order derivative in space (compare with Fig. 23(b)). (*Horizontal dimension*: space x; *Vertical dimension*: time t.)

derivatives over scale-normalized time $\lambda = t/\sqrt{\tau}$ according to $\partial_\lambda = \sqrt{\tau}\,\partial_t$ or more generally $\partial_\lambda = \tau^{\gamma/2}\,\partial_t$, we can then consider the following generalizations of the quasi quadrature measure:[10]

---

[10] In the first spatio-temporal quasi quadrature entity $Q_1 L$, the square of the first-order derivative L_ξ^2 in the corresponding one-dimensional measure QL has been replaced by the squared gradient magnitude $L_\xi^2 + L_\lambda^2$ in space-time, while the square of the second-order derivative $L_{\xi\xi}^2$ has been replaced by the Frobenius norm $L_{\xi\xi}^2 + 2L_{\xi\lambda}^2 + L_{\lambda\lambda}^2$ of the Hessian matrix computed over a scale-normalized space-time. For this entity to respond, it is sufficient that there are significant variations in the signal over *either* space or time. The second spatio-temporal quasi quadrature measure $Q_2 L$ has been defined as the product of corresponding quasi quadrature measures over pure space $L_\xi^2 + C\,L_{\xi\xi}^2$ and pure time $L_\lambda^2 + C\,L_{\lambda\lambda}^2$. Therefore, this entity will only generate responses when there simultaneously occur variations in the signal over *both* space and time. Hence, the operator $Q_2 L$ will be much more selective than $Q_1 L$. Finally, the third entity $Q_3 L$ has been defined from spatio-temporal derivative operators with high degree of qualitative similarity to biological receptive fields, and with a similar spirit of summing up squares of Gaussian derivative operator responses that correspond to first- and second-order derivatives with respect to space and time. Since all the primitives in $Q_3 L$ contain derivatives with respect to both space and time, this entity will only generate significant responses if there are significant variations over *both* space and time. For all of these quasi quadrature entities, we can compute the derivatives either from a space-time separable spatio-temporal scale-space or a velocity-adapted scale-space. In the latter case, the entire operator will therefore be tuned to a particular stimulus velocity. The illustrations in Fig. 27 have, however, been computed with space-time separable derivative operators.

Fig. 23 Space-time separable kernels obtained from the spatio-temporal scale-space concepts: (*upper left*) Gaussian spatio-temporal kernel $g_{xt}(x,t;\ s,\tau,\delta) = g_x(x;\ s)\,g_t(t;\ \tau,\delta)$ with $s = 0.3$, $\tau = 40^2$, $\delta = 100$; (*upper right*) Gaussian spatio-temporal kernel $g_{xxt}(x,t;\ s,\tau,\delta) = g_{xx}(x;\ s)\,g_t(t;\ \tau,\delta)$ with $s = 0.3$, $\tau = 60^2$, $\delta = 150$; (*lower left*) Time-causal spatio-temporal kernel $h_{xt'}(x,t;\ s,\tau) = g_x(x;\ s)\,\phi_{t'}(t;\ \tau,\delta)$ with $s = 0.4$, $\tau = 17$; (*lower right*) Time-causal spatio-temporal kernel $h_{xxt'}(x,t;\ s,\tau) = g_{xx}(x;\ s)\,\phi_{t'}(t;\ \tau,\delta)$ with $s = 0.4$, $\tau = 22$. For the time-causal kernels, the temporal derivatives have been computed using the transformed temporal derivative operator $\partial_{t'} \sim t^\kappa \partial_t$, here with $\kappa = 1/2$. Compare the qualitative shapes of these kernels with the kernels in Fig. 22. (*Horizontal dimension*: space x; *Vertical dimension*: time t.)
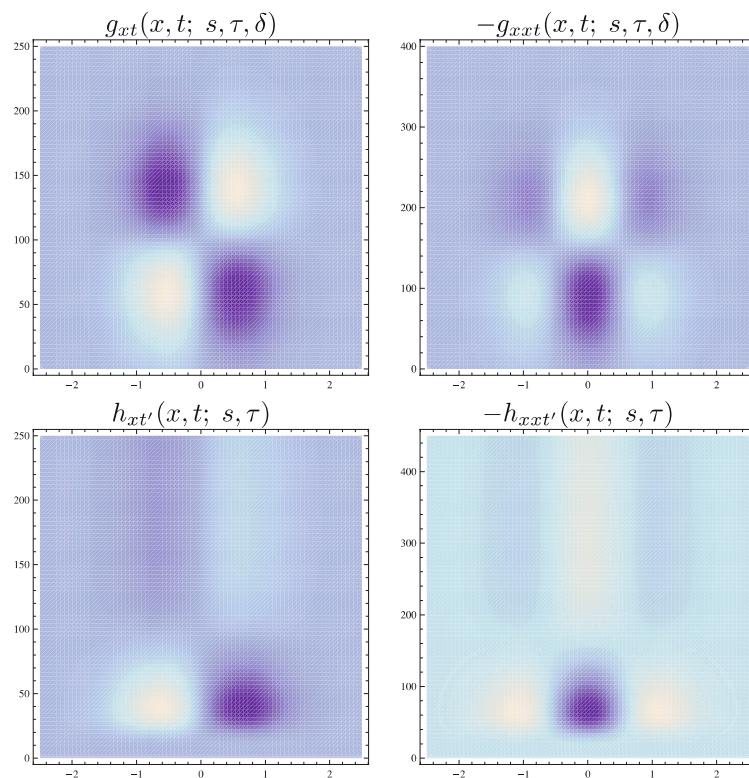
$$\mathcal{Q}_1 L = L_\xi^2 + L_\lambda^2 + C\,(L_{\xi\xi}^2 + 2L_{\xi\lambda}^2 + L_{\lambda\lambda}^2)$$
$$= sL_x^2 + \tau L_t^2 + C\,(s^2 L_{xx}^2 + 2s\tau L_{xt}^2 + \tau^2 L_{tt}^2), \quad (148)$$

$$(\mathcal{Q}_2 L)^2 = (L_\xi^2 + C\,L_{\xi\xi}^2)(L_\lambda^2 + C\,L_{\lambda\lambda}^2)$$
$$= (sL_x^2 + C\,s^2 L_{xx}^2)(\tau L_t^2 + C\,\tau^2 L_{tt}^2), \quad (149)$$

$$\mathcal{Q}_3 L = L_{\xi\lambda}^2 + C\,L_{\xi\xi\lambda}^2 + C\,L_{\xi\lambda\lambda}^2 + C^2\,L_{\xi\xi\lambda\lambda}^2$$
$$= s\tau L_{xt}^2 + C\,s^2\tau L_{xxt}^2 + C\,s\tau^2 L_{xtt}^2 + C^2\,s^2\tau^2 L_{xxtt}^2. \quad (150)$$

Fig. 24 Examples of non-separable receptive field profiles in the striate cortex as reported by DeAngelis et al. [11]: (**a**) a receptive field reminiscent of a second-order derivative in tilted space-time (compare with the *left column* in Fig. 25) (**b**) a receptive field reminiscent of a third-order derivative in tilted space-time (compare with the *right column* in Fig. 25). (*Horizontal dimension*: space x; *Vertical dimension*: time t.)

For the time-causal scale-space, we can express corresponding scale-normalized operators as
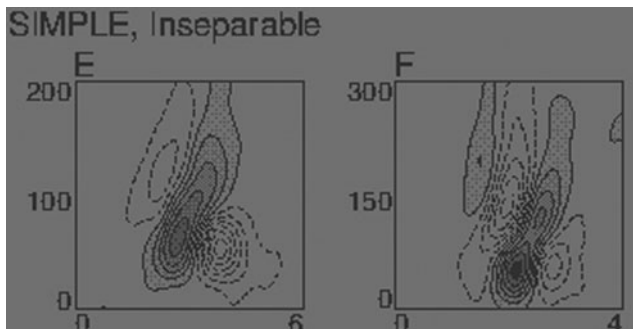
$$\mathcal{Q}_1 L = L_\xi^2 + L_\lambda^2 + C\,(L_{\xi\xi}^2 + 2L_{\xi\lambda}^2 + L_{\lambda\lambda}^2)$$
$$= sL_x^2 + \tau L_t^2 + C\,(s^2 L_{xx}^2 + 2s\tau L_{xt'}^2 + \tau^2 L_{t't'}^2), \quad (151)$$

$$(\mathcal{Q}_2 L)^2 = (L_\xi^2 + C\,L_{\xi\xi}^2)(L_\lambda^2 + C\,L_{\lambda\lambda}^2)$$
$$= (sL_x^2 + C\,s^2 L_{xx}^2)(\tau L_{t'}^2 + C\,\tau^2 L_{t't'}^2), \quad (152)$$

$$\mathcal{Q}_3 L = L_{\xi\lambda}^2 + C\,L_{\xi\xi\lambda}^2 + C\,L_{\xi\lambda\lambda}^2 + C^2\,L_{\xi\xi\lambda\lambda}^2$$
$$= s\tau L_{xt'}^2 + C\,s^2\tau L_{xxt'}^2$$
$$+ C\,s\tau^2 L_{xt't'}^2 + C^2\,s^2\tau^2 L_{xxt't'}^2, \quad (153)$$

where the temporal derivatives $\partial_{t'}$ with respect to self-similarly transformed time are related to derivatives with respect to regular time according to $\partial_{t'} \sim t^\kappa \partial_t$.[11]

---

[11]In analogy with the computation of multi-scale second-moment descriptors Gårding and Lindeberg [25], we can apply a second convolution stage determined by integration scale parameters to the computation of the local pointwise derivative descriptors $\mathcal{Q}_i$ in order to suppress local ripples. For the quasi quadrature entities derived from the Gaussian spatio-temporal scale-space, we do of course choose a non-causal Gaussian spatio-temporal kernel, while we for the corresponding entities derived from the time-causal spatio-temporal scale-space choose a time-causal spatio-temporal kernel for the second-stage integration smoothing.

Fig. 25 Non-separable spatio-temporal receptive fields obtained by applying velocity-adapted second- and third-order derivative operations in space-time to spatio-temporal smoothing kernels generated by the spatio-temporal scale-space concept. (*upper left*) Gaussian spatio-temporal kernel $g_{xx}(x,t;\,s,\tau,v,\delta)$ with $s=0.5$, $\tau=50^2$, $v=0.006$, $\delta=100$; (*upper right*) Gaussian spatio-temporal kernel $g_{xxx}(x,t;\,s,\tau,v,\delta)$ with $s=0.5$, $\tau=60^2$, $v=0.006$, $\delta=130$; (*lower left*) Time-causal spatio-temporal kernel $h_{xx}(x,t;\,s,\tau,v)$ with $s=0.4$, $\tau=15$, $v=0.006$; (*lower right*) Time-causal spatio-temporal kernel $h_{xxx}(x,t;\,s,\tau,v)$ with $s=0.4$, $\tau=15$, $v=0.006$. Compare the qualitative shapes of these kernels with the kernels in Fig. 24. (*Horizontal dimension*: space x; *Vertical dimension*: time t.)
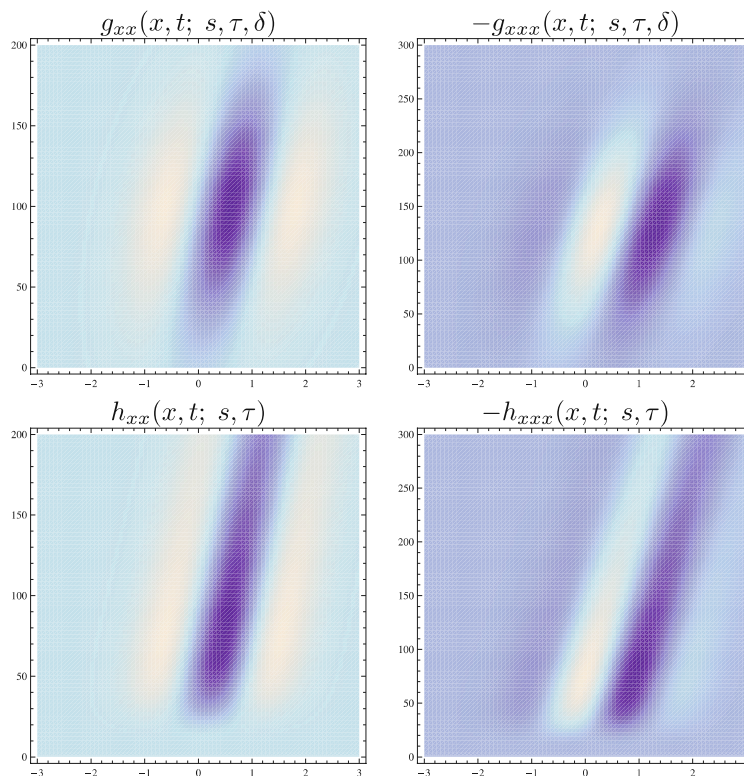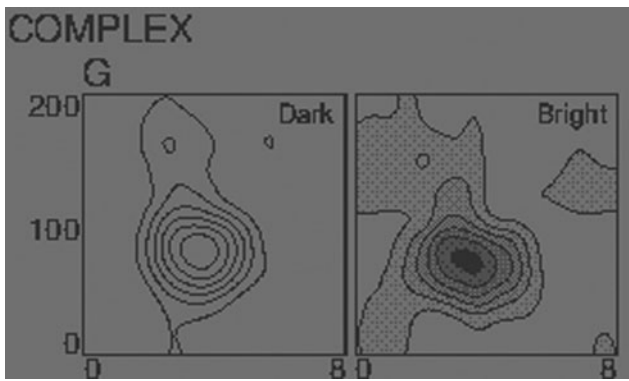
Fig. 26 Response profile of a complex cell in the joint space-time domain as reported by DeAngelis et al. [11]. Within the framework of the spatio-temporal scale-space framework presented in this paper, such a response property can be obtained by a quasi-quadrature combination of first- and second-order receptive fields; see Fig. 27. (*Horizontal dimension*: space x; *Vertical dimension*: time t.)

Figure 27 shows the result of computing the response of these quasi quadrature measures to a delta function over a 1+1-D space-time. Note that this type of computational structure is nicely compatible with results by Valois et al. [70], who show that first- and second-order receptive fields typically occur in pairs that can be modelled as approximate Hilbert pairs.

7 Summary and Discussion

We have presented a generalized theory for Gaussian scale-space representation of spatial and/or spatio-temporal data. Starting from a general condition about non-creation of spurious structures with increasing scales formalized in terms of non-enhancement of local extrema, a complete characterization has been given of the semi-groups of convolution transformations that obey this requirement on different types of image domains. The resulting theory comprises the existing continuous scale-space theory on symmetric spatial domains, with extensions to non-symmetric spatial domains as well as spatio-temporal domains. Specifically, we have shown that this combination of scale-space axioms makes it possible to axiomatically derive the notions of:

– rotationally symmetric Gaussian scale-space on isotropic spatial domains,
– affine Gaussian scale-space on anisotropic spatial domains,
– Gaussian spatio-temporal scale-space on non-causal spatio-temporal domains, and
– time-causal spatio-temporal scale-space on time-causal spatio-temporal domains.

A main message is that a much richer structure of affine as well as spatio-temporal filters can be obtained if we start
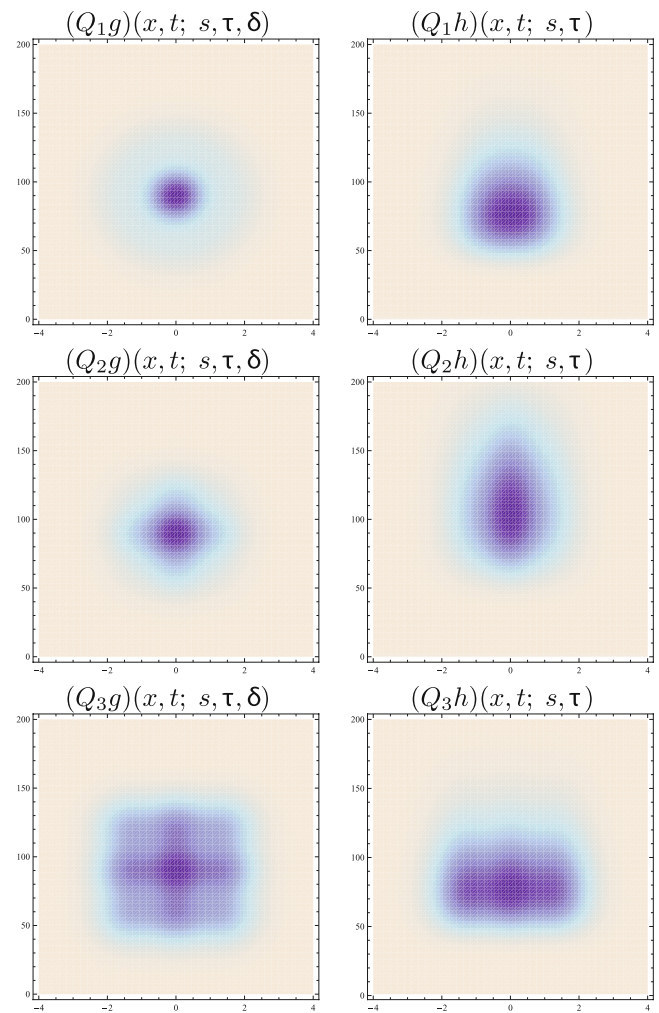
Fig. 27 The response of different spatio-temporal quasi quadrature measures to a delta function. (*left*) Computed for a spatio-temporal Gaussian $g(x, t; s, \tau, \delta)$ according to (*top*) $Q_1 g = s g_x^2 + \tau g_t^2 + C (s^2 g_{xx}^2 + 2s\tau g_{xt}^2 + \tau^2 g_{tt}^2)$; (*middle*) $Q_2^2 g = (s g_x^2 + s^2 g_{xx}^2)(\tau g_t^2 + C \tau^2 g_{tt}^2)$; (*bottom*) $Q_3 g = (s\tau g_{xt}^2 + C s^2\tau g_{xxt}^2 + C s\tau^2 g_{xtt}^2 + C^2 s^2\tau^2 g_{xxtt}^2)$ with $s = 1.2$, $\tau = 25^2$, $\delta = 90$, $C = e/4$. (*right*) Computed for the time-causal kernel $h(x, t; s, \tau)$ according to (*top*) $Q_1 h = s h_x^2 + \tau h_t^2 + C (s^2 h_{xx}^2 + 2s h_{xt}^2 + \tau^2 h_{tt}^2)$; (*middle*) $Q_2^2 h = (s h_x^2 + s^2 h_{xx}^2)(\tau h_t^2 + C \tau^2 h_{tt}^2)$; (*bottom*) $Q_3 h = (s\tau h_{xt}^2 + C s^2\tau h_{xxt}^2 + C s\tau^2 h_{xtt}^2 + C^2 s^2\tau^2 h_{xxtt}^2)$ with $s = 1.2$, $\tau = 25^2$, $\delta = 90$, $C = e/4$. (*Horizontal dimension*: space x; *Vertical dimension*: time t.)

from a reformulation of Koenderink's causality requirement into non-enhancement of local extrema, and then relax the requirement of spatial symmetry that was prevalent in the earliest scale-space formulations as well as most follow-up works.

In companion works, such affine and spatio-temporal scale-spaces have been shown to be highly useful for different tasks in computer vision, by allowing the vision system to take into explicit account as well as to compensate for the

following type of image transformations that arise when a vision system observes a real world:

– affine transformations arising from the first-order linearized component of the perspective mapping, and
– Galilean transformations arising because of relative motions between the observer and objects in the world.

Indeed, by considering more general covariance matrices for anisotropic handling of different dimensions and as well as spatial and/or spatio-temporal derivative operators applied to corresponding filters, a much richer family of filter shapes can be generated than from rotationally symmetric Gaussian filters. All these generalized derivative filters do also obey non-enhancement of local extrema as well as a transfer of the semi-group property into a cascade smoothing property. We have also showed that the resulting spatial as well as spatio-temporal derivative operations have high similarities to receptive fields recorded from biological vision. The treatment does hence show that a very rich and general set of visual front-end operations can be obtained from a unified and generalized Gaussian scale-space theory.

Of course, we do not exclude the possibilities of considering other types of non-Gaussian scale-space theories, such as the self-similar scale-space families arising from (3) or its affine generalization $\hat{h}(\omega; s) = e^{-\alpha|B\omega|^p}$, where B is a non-singular $N \times N$ matrix. In this context we would, however, like to stress that the generalized Gaussian scale-space theory presented in this paper constitutes a particularly convenient class with most attractive properties. For example, compared to the Poisson kernel in (4), the Gaussian smoothing filter decreases much faster towards infinity and faster than any polynomial, which implies a very strong regularizing property for any scale-space derivative. Compared to the α-scale-spaces, the Gaussian scale-spaces have classical infinitesimal generators, straightforward closed-form expressions in the spatial domain and obey non-enhancement of local extrema. The Gaussian scale-spaces are also maximally uncommitted in the sense that their smoothing kernels have maximum entropy.

Concerning more technical contributions, we have also analysed the time-causal spatio-temporal scale-space in more detail, regarding the temporal cascade structure it satisfies over time and as well as specific properties of the corresponding time-causal spatio-temporal derivative kernels, which differ from the properties of the more commonly used Gaussian spatio-temporal derivatives. We have moreover shown how temporal derivatives with respect to self-similarly transformed time can be defined, resulting in the formulation of a novel analogue to normalized derivatives for time-causal temporal or spatio-temporal scale-spaces.

We propose that this generalized Gaussian scale-space framework constitutes both a natural, theoretically well-founded and general basis to consider (i) when designing visual front-end operations for computer vision systems and (ii) when modelling some of the earliest processing stages in biological vision.

Appendix A: Continuity Requirements on the Semi-group with Respect to Scale

When deriving the necessity of the form of the scale-space representation, we use a condition about non-enhancement of local extrema expressed in terms of a sign condition on the derivative $\partial_s L(x_0; s_0)$ at local extrema x_0 of the scale-space representation L with respect to the scale parameter s at any scale s_0. While the notion of strong C_0 continuity of the semi-group $\mathcal{T}$ in (17) implies that $L(\cdot; s) = \mathcal{T}(s) f$ satisfies the differential equation $\partial_s L = \mathcal{A}L$ and that the derivative exists almost everywhere for a dense subset of functions f in the Banach space X over which the semi-group is defined, this structure does not necessarily implies that the partial derivative $\partial_s L(x_0; s_0)$ is well-defined pointwise for every $(x_0; s_0) \in \mathbb{R}^N \times \mathbb{R}_+$. In this appendix, we will express a sufficient condition on the semi-group $\mathcal{T}$ such that the derivative of the scale-space representation L with respect to the scale parameter is well-defined for smooth functions f of compact support. We start by expressing a basic lemma that will make it possible for us to define the infinitesimal generator for a specific set of functions $f^r \in X$. This result will then be used for approximating general functions $f \in X$ using Sobolev norms.

Lemma 8 (Explicit form of infinitesimal generator for a subset of functions $f^r \in X$) *Let X be a Banach space, and let $\mathcal{T}(s)$ be a semi-group of operators from X to X. For any $x \in X$, let*

$$f^r = \mathcal{C}(r) f = \frac{1}{r} \int_{s=0}^{r} \mathcal{T}(s) f \, ds. \tag{154}$$

Then, provided that the semi-group $\mathcal{T}$ satisfies the C_0 continuity requirement that

$$\lim_{r \downarrow 0} \|\mathcal{C}(r) f - f\|_X = 0 \tag{155}$$

should hold for every $f \in X$, it follows that for every $r > 0$, we have that $\mathcal{C}(r) f \in D(\mathcal{A})$ for every $f \in X$ and the result

of applying the infinitesimal generator of the semi-group to f^r is

$$\mathcal{A}f^r = \lim_{h \to 0} \frac{\mathcal{T}(h) - I}{h} f^r = \frac{\mathcal{T}(r)f - f}{r}. \tag{156}$$

Proof Consider

$$\frac{\mathcal{T}(h)f^r - f^r}{h} = \frac{1}{rh}\left(\mathcal{T}(h)\int_{s=0}^{r} \mathcal{T}(s)f \, ds \right.$$
$$\left. - \int_{s=0}^{r} \mathcal{T}(s)f \, ds\right)$$
$$= \frac{1}{rh}\int_{s=0}^{r} \mathcal{T}(s+h)f \, ds - \int_{s=0}^{r} \mathcal{T}(s)f \, ds. \tag{157}$$

Then, by a change of variables in the first integral we have

$$\frac{\mathcal{T}(h)f^r - f^r}{h} = \frac{1}{rh}\left(\int_{s=h}^{r+h} \mathcal{T}(s)f \, ds - \int_{s=0}^{r} \mathcal{T}(s)f \, ds\right) \tag{158}$$

which by an inspection of the intervals over which the integration is performed can be written

$$\frac{\mathcal{T}(h)f^r - f^r}{h} = \frac{1}{r}\left(\frac{1}{h}\int_{s=r}^{r+h} \mathcal{T}(s)f \, ds \right.$$
$$\left. - \frac{1}{h}\int_{s=0}^{h} \mathcal{T}(s)f \, ds\right). \tag{159}$$

Due to the assumption (155) in combination with (159) it therefore follows that

$$\frac{\mathcal{T}(h)f^r - f^r}{h} \to \frac{\mathcal{T}(r)f - f}{r} \tag{160}$$

when $h \to 0$. Hence, we know for sure that $f^r \in D(\mathcal{A})$ with

$$\mathcal{A}f^r = \frac{\mathcal{T}(r)f - f}{r}. \tag{161}$$

(Please, note that there is no limit operator in this expression.) Specifically, it holds that

$$\left(\frac{\mathcal{T}(h) - I}{h} - A\right)f^r$$
$$= \frac{1}{r}\left(\frac{1}{h}\int_{s=0}^{h} \mathcal{T}(s+r)f \, ds - \frac{1}{h}\int_{s=0}^{h} \mathcal{T}(s)f \, ds\right)$$
$$- \frac{\mathcal{T}(r)f - f}{r}. \tag{162}$$

$\square$

Our aim is then to use an equality with an L^2-based Sobolev norm[12] to estimate the maximum norm of the deviation between the difference approximation $(\mathcal{T}(h) - I)/h$ and the infinitesimal generator in (162).

Lemma 9 (Maximum norm estimate from L^2-based Sobolev norm) *Given any function $u \in L^1(\mathbb{R}^N) \cap L^2(\mathbb{R}^N)$, define its Fourier transform by*

$$\hat{u}(\omega) = \int_{x \in \mathbb{R}^N} u(x)\, e^{-i\omega x}\, dx \tag{163}$$

and let for any $k > N/2$ (where k is not required to be an integer)

$$\|u\|_{H^k(\mathbb{R}^N)} = \left(\int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2\right)^k |\hat{u}(\omega)|^2 \, d\omega\right)^{1/2} \tag{164}$$

denote the L^2-based Sobolev norm of u of order k. Then, we can estimate the maximum norm of u from its L^2-based Sobolev norm by

$$\|u\|_{L^\infty(\mathbb{R}^N)} \le C \|u\|_{H^k(\mathbb{R}^N)}, \tag{165}$$

where the constant C does only depend on k and N and not on u.

Proof Consider

$$|u(x)| = \left|\int_{\omega \in \mathbb{R}^N} e^{-i\omega x} \hat{u}(\omega)\, d\omega\right| \le \int_{\omega \in \mathbb{R}^N} |\hat{u}(x)|\, d\omega$$
$$= \int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2\right)^{-k/2}\left(1 + |\omega|^2\right)^{k/2}|\hat{u}(\omega)|\, d\omega \tag{166}$$

which by the Schwartz inequality and the definition of the L^2-based Sobolev norm $\|u\|_{H^k(\mathbb{R}^N)}$ can be overestimated by

$$|u(x)| \le \left(\int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2\right)^{-k}\, d\omega\right)^{1/2}$$
$$\left(\int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2\right)^{k} |\hat{u}(\omega)|^2\, d\omega\right)^{1/2}$$
$$= \left(\int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2\right)^{-k}\, d\omega\right)^{1/2} \|u\|_{H^k(\mathbb{R}^N)} \tag{167}$$

provided that the integral

$$\int_{\omega \in \mathbb{R}^N} \left(1 + |\omega|^2\right)^{-k}\, d\omega = \int_{R=0}^{\infty} \frac{C'(N)\, R^{N-1}}{(1 + R^2)^k}\, dR = C^2 \tag{168}$$

---

[12]See *e.g.* [23, Chap. 6] for an introduction to Fourier based Sobolev norms as used here.

converges. By the use of N-dimensional spherical co-ordinates, where $C'(N)$ denotes the area of an $N-1$-dimensional unit hypersphere, it is clear that the latter integral converges provided that $N-1-2k < -1$, *i.e.,* provided that $k > N/2$. $\qquad\square$

Using this Sobolev inequality, we will estimate

$$\left\| \left(\frac{\mathcal{T}(h) - I}{h} - \mathcal{A} \right) f \right\|_{L^\infty(\mathbb{R}^N)}$$
$$\leq C \left\| \left(\frac{\mathcal{T}(h) - I}{h} - \mathcal{A} \right) f \right\|_{H^k(\mathbb{R}^N)} \qquad (169)$$

and prove differentiability with respect to s and that the infinitesimal generator $\mathcal{A}$ is well-defined for smooth functions f of compact support given certain regularity requirements on the semi-group $\mathcal{T}$ by showing that

$$\lim_{h\downarrow 0} \left\| \left(\frac{\mathcal{T}(h) - I}{h} - \mathcal{A} \right) f \right\|_{H^k(\mathbb{R}^N)} = 0. \qquad (170)$$

Lemma 10 (Existence of infinitesimal generator for smooth functions with bounded support) *Let $\mathcal{T}$ be a linear semi-group corresponding to convolution kernels $T(\cdot;\, s) \in L^1(\mathbb{R}^N)$ that satisfies the continuity requirement that for some value of $k > N/2$*

$$\lim_{h\downarrow 0} \left\| \int_{s=0}^h \frac{(\mathcal{T}(s) - I)}{h} f\, ds \right\|_{H^k(\mathbb{R}^N)} = 0 \qquad (171)$$

should hold for all smooth functions $f \in L^1(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$. Then, smooth ($C^\infty$) functions f with bounded support are in the domain $D(\mathcal{A})$ and for such functions the derivative $\partial_s L(x;\, t) = (\mathcal{A}L)(x;\, s)$ exists for every $(x;\, s) \in \mathbb{R}^N \times \mathbb{R}_+$.

Proof Given the estimate (169), let us for any smooth function f with bounded support consider

$$\left\| \left(\frac{\mathcal{T}(h) - I}{h} - \mathcal{A} \right) f \right\|_{H^k(\mathbb{R}^N)}$$
$$\leq \left\| \left(\frac{\mathcal{T}(h) - I}{h} - A \right) (f - f^r) \right\|_{H^k(\mathbb{R}^N)}$$
$$+ \left\| \left(\frac{\mathcal{T}(h) - I}{h} - A \right) f^r \right\|_{H^k(\mathbb{R}^N)} \qquad (172)$$

with f^r according to (154), where we from equations (160) and (161) in combination with the assumption (171) know that

$$\lim_{h\downarrow 0} \left\| \left(\frac{\mathcal{T}(h) - I}{h} - A \right) f^r \right\|_{H^k(\mathbb{R}^N)} = 0 \qquad (173)$$

for every $r > 0$. From (162) it follows that

$$\left(\frac{\mathcal{T}(h) - I}{h} - A \right) (f - f^r)$$
$$= \frac{1}{h} \int_{s=0}^h \frac{(\mathcal{T}(s+r) - \mathcal{T}(s))(f - f^r)}{r}\, ds$$
$$\quad - \frac{(\mathcal{T}(r) - I)(f - f^r)}{r}$$
$$= \frac{1}{h} \int_{s=0}^h \frac{\mathcal{T}(s)(\mathcal{T}(r) - I)(f - f^r)}{r}\, ds$$
$$\quad - \frac{(\mathcal{T}(r) - I)(f - f^r)}{r}$$
$$= \frac{1}{h} \int_{s=0}^h \frac{(\mathcal{T}(s) - I)(\mathcal{T}(r) - I)(f - f^r)}{r}\, ds. \qquad (174)$$

Returning to the definition of f^r, we have that

$$f^r - f = \frac{1}{r} \int_{s'=0}^r \mathcal{T}(s') f\, ds' - f = \int_{s'=0}^r \frac{\mathcal{T}(s') - I}{r} f\, ds' \qquad (175)$$

and we can write

$$\left(\frac{\mathcal{T}(h) - I}{h} - A \right) (f^r - f)$$
$$= \frac{1}{h} \int_{s=0}^h \int_{s'=0}^r \frac{(\mathcal{T}(s) - I)(\mathcal{T}(r) - I)(\mathcal{T}(s') - I)}{r^2} f\, ds\, ds'. \qquad (176)$$

A sufficient condition for

$$\lim_{h\downarrow 0} \left\| \frac{1}{h} \left(\frac{\mathcal{T}(h) - I}{h} - \mathcal{A} \right) f \right\|_{L^\infty(\mathbb{R}^N)} = 0 \qquad (177)$$

in (169) and for f to be in $D(\mathcal{A})$ is therefore that

$$\lim_{h\downarrow 0} \left\| \frac{1}{h} \int_{s=0}^h \int_{s'=0}^r \frac{(\mathcal{T}(s) - I)(\mathcal{T}(r) - I)(\mathcal{T}(s') - I)}{r^2} f\, ds\, ds' \right\|_{H^k(\mathbb{R}^N)}$$
$$= 0 \qquad (178)$$

should hold for some $k > N/2$, some $r > 0$ and all functions f of compact support in combination with (171); also for all functions f of compact support. By rewriting (178) into

$$\lim_{h\downarrow 0} \left\| \frac{1}{h} \left(\int_{s=0}^h (\mathcal{T}(s) - I)\, ds \right) \right.$$
$$\left. \left(\int_{s'=0}^r \frac{(\mathcal{T}(r) - I)(\mathcal{T}(s') - I)}{r^2}\, ds' \right) f \right\|_{H^k(\mathbb{R}^N)} = 0 \qquad (179)$$

by noting that $\mathcal{T}(s)$ is a bounded operator and $r > 0$ is not required to tend to zero, then we can see that the operator

$$\mathcal{U} = \int_{s'=0}^{r} \frac{(\mathcal{T}(r) - I)(\mathcal{T}(s') - I)}{r^2} \, ds' \qquad (180)$$

will also be a bounded operator. Since the operator $\mathcal{T}$ corresponds to a convolution kernel $T(\cdot; s) \in L^1(\mathbb{R}^N)$, the operator $\mathcal{U}$ does also preserves smoothness. Given that f has bounded support, it follows that $\mathcal{U}f$ will be in $L^1(\mathbb{R}^N)$. Hence, a sufficient condition for (179) to hold can also be expressed as the requirement that

$$\lim_{h \downarrow 0} \left\| \frac{1}{h} \left(\int_{s=0}^{h} (\mathcal{T}(s) - I) \, ds \right) f' \right\|_{H^k(\mathbb{R}^N)} = 0 \qquad (181)$$

should for some $k > N/2$ hold for all smooth functions $f' \in L^1(\mathbb{R}^N) \cap C^\infty(\mathbb{R}^N)$. The latter condition is similar to the notion of C_1 continuity of the semi-group, although here being expressed in terms of the Sobolev norm $\| \cdot \|_{H^k(\mathbb{R}^N)}$. For a general Banach space, the notion of C_1 continuity is a stronger condition than C_0 continuity [29, page 322]. □

Appendix B: Non-enhancement of Local Extrema vs. the Maximum Principle

This appendix describes relationships between non-enhancement of local extrema as formulated in Definition 3 and the maximum principle.

The *strong* maximum principle for parabolic or elliptic differential equations states that if a function assumes its maximum in the interior of the domain, then the function must be a constant [15, pages 330–333, 375–377]. The *weak* maximum principle on the other hand says that the maximum of the function is to be found on the boundary, but may also occur in the interior [15, pages 327–329, 368–370]. Corresponding results can be expressed for minima.

For our purpose of defining a scale-space representation, however, we cannot *a priori* assume that the scale-space should be generated by a parabolic or elliptic differential equation. Hence, it is of general interest to perform a study without *a priori* preconceptions regarding the form of the evolution equation. Moreover it should be noted that the maximum principle refers to *global* properties of the function, while non-enhancement of local extrema refers to *local* properties. For a general evolution equation, one may hence conceive situations where the global maximum of a function has to occur at the boundary, while some other local maximum point (which is not the local maximum) may nevertheless be enhanced (see Fig. 28). Therefore, one may express maximum principles for differential equations that do not obey non-enhancement of local extrema, such as the simple differential equation

$$\partial_t L = -L \qquad (182)$$
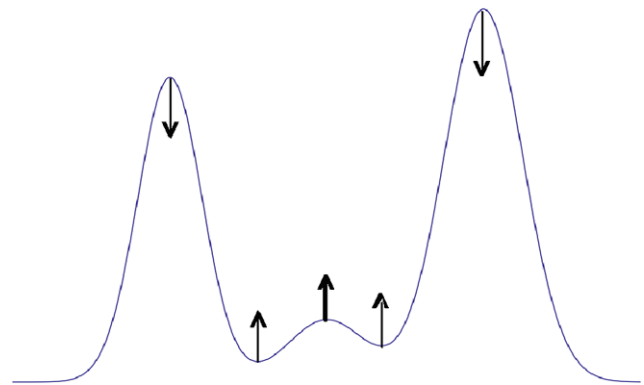
Fig. 28 For a general evolution process, one may conceive cases where the global maximum (minimum) always has to decrease (increase), while local extrema may nevertheless be enhanced. This figure gives a schematic sketch of the sign of derivative with respect to the evolution parameter s for an evolution equation of the form (187), where the central local maximum marked indicated by a bold arrow will violate non-enhancement of local extrema provided that the support region of the weight function w is sufficiently wide

with initial condition $L(x; 0) = f(x)$. From the solution $L(x; t) = e^{-t} f(x)$ it is obvious that at a negative maximum over x will always increase with t and the value at a positive minimum over x will always decrease with t. This example hence shows that *the assumption of non-enhancement of local extrema leads to a different set of smoothing processes than the maximum principle would lead to*.

B.1 A Formal Connection Given the Assumption of a Local Process

Still, however, there is an at least formal connection between non-enhancement of local extrema and the maximum principle. If we *assume* that the scale-space should be governed by a partial differential equation with a linear operator $\mathcal{A}L$ that corresponds to a *local operator* (in terms of derivatives at the central point only), then we can use similar arguments as in step C.2 in the proof of Theorem 5 to show that $\mathcal{A}L$ must not contain derivatives of order higher than two. If we use a test function of the form

$$f_4(x) = (-x_1^2 - x_2^2 - \cdots - x_N^2 + \beta_2 x^\eta) \, \chi(Kx), \qquad (183)$$

where $\chi(x)$ is defined in a similar way as in the text following (35), then we can choose β_2 and K in such a way that $f_4(x) \leq 0 \, \forall x$ and $\mathcal{A}L(0) \geq 0$. Thus, the maximum principle would be violated, because the maximum of the solution would be greater than zero and would not occur at the boundary of the domain.

Given the complementary assumption of a local process, the requirement of the maximum principle to hold does, however, not yield an offset a_0 equal to zero in the evolution equation as the requirement of non-enhancement of local extrema leads to. If we allow ourselves to reparameterize the

scale parameter in the evolution equation by the following change of variables

$$\tilde{L} = e^{-a_0 s} L \tag{184}$$

and if L is a solution of

$$\partial_s L = \sum_{0 < |\eta| \le 2} a_\eta \partial_{x^\eta} L + a_0 L \tag{185}$$

as would be obtained from the maximum principle, then the transformed representation $\tilde{L}$ is a solution of

$$\partial_s \tilde{L} = \sum_{0 < |\eta| \le 2} a_\eta \partial_{x^\eta} \tilde{L} \tag{186}$$

which agrees with the form of evolution equations obtained from non-enhancement of local extrema. In this respect, there is a close formal relationship between non-enhancement of local extrema and the maximum principle. Given the requirement of the maximum principle to hold, it should, however, be emphasized that non-enhancement of local extrema would not be guaranteed to hold in the original domain, only in *some* transformed domain.

B.2 A More Fundamental Difference in Terms of Local vs. Non-local Processes

In the above mentioned analysis, we made a complementary assumption that the evolution equation should be determined by a *local* evolution equation. In our previous treatment in Sect. 3, such a local form of the evolution process was established from the assumption of non-enhancement of local extrema in step C.1 in the necessity proof underlying Theorem 5. A very notable difference between the maximum principle and non-enhancement of local extrema, however, is that the maximum principle allows for *non-local evolution processes*, where the minimum and maximum values have to occur at the boundary, whereas non-enhancement of local extrema is violated for such evolution processes. Consider, for example, an evolution equation of the form

$$\partial_s L(x; \; s) = \int_{\xi=-\infty}^{\infty} w(\xi) \left(L(x - \xi; \; s) - L(x; \; s) \right) d\xi, \tag{187}$$

where w is a non-negative smooth C^∞ function with bounded L^1 norm. From the averaging interpretation of this process, it is evident that if x_0 is a global maximum (minimum) of the mapping $x \to L(x; \; s_0)$, then $L(x_0 - \xi; \; s) - L(x_0; \; s_0)$ will always be negative (positive), which means that $\partial_s L(x_0; \; s_0)$ is guaranteed to be negative (positive). Hence, this evolution process satisfies a maximum (minimum) principle in the sense that the global maximum (minimum) of the mapping $x \to L(x; \; s_0)$ is always

guaranteed to decrease (increase). This property in turn means that the global maximum (minimum) of the mapping $(x, s) \to L(x; \; s)$ has to occur at the boundary of the (x, s) domain. There is, however, no guarantee concerning the sign of $\partial_s L(x; \; s)$ at local extrema of the mapping $x \to L(x; \; s_0)$, and an evolution process of this form may therefore violate non-enhancement of local extrema (see Fig. 28 for an illustration).

Thus, *an evolution process that obeys the maximum principle does not even have to be local*, in the sense that the right-hand side in the evolution equation (187) may depend on values of L at several points, while the result in Theorem 5 shows that a smoothing process that obeys non-enhancement of local extrema always has to be local and be determined by a second-order parabolic operator.

B.3 Fourth-Order Evolution Equations

Maximum principles can also be stated for certain types of fourth-order elliptic equations (Dunninger [14], Zhang and Zhang [79]), such as

$$(\nabla^2)^2 L + L = 0, \tag{188}$$

which are, however, not within the class of non-enhancement scale-spaces that we shall consider. In connection with the problem of choosing between candidate smoothing processes, it may also be highly relevant to ask if, for example, a fourth-order evolution equation of the form[13]

$$\partial_s L = -(\nabla^2)^2 L \tag{189}$$

could be regarded as a possible model for generating a scale-space? The solutions of this equation may, however, have non-intuitive properties in the sense that positivity may be violated and local ripples may be amplified [7, page 366]. The latter fourth-order process can be ruled out by both the maximum principle and non-enhancement of local extrema.

Appendix C: Galilean-Invariant Fixed Points in Gaussian Spatio-Temporal Scale-Space

The subject of this appendix is to describe *one* way to make appropriate selection of velocity parameters for interpreting the results in a multi-parameter spatio-temporal scale-space representation in a general situation when there is no *a priori* information of the motion of observed objects relative to the observer. The method that will be proposed is based on the definition of *Galilean invariant fixed points* constructed by normalizing local spatio-temporal image patterns by a

---

[13]The minus sign is needed here to make this equation dissipative in L^2 [7, page 371].

method referred to as *Galilean block diagonalization* and by comparing and adapting the velocity parameter based on local spatio-temporal image measurements.

C.1 Basic Definitions

Consider a spatio-temporal image sequence $f(p) = f(x, t)$ defined over $\mathbb{R}^N \times \mathbb{R}$ with spatio-temporal scale-space representation $L(p; \Sigma)$ given by $L(\cdot; \Sigma, \delta) = g(\cdot; \Sigma, \delta) * f(\cdot)$, where g denotes the spatio-temporal Gaussian kernel

$$g(x; \Sigma, \delta) = \frac{1}{(2\pi)^{(N+1)/2}\sqrt{\det \Sigma}} e^{-(p-\delta)^T \Sigma^{-1}(p-\delta)/2}, \tag{190}$$

with a spatio-temporal covariance matrix Σ of the form (61) and with time delay δ.

Given any velocity vector u, define a Galilean transformed image f'' by $f'(p'') = f(p)$, where $p'' = G(u)p$ and $G(u)$ denotes a Galilean transformation with image velocity u. Moreover, define the spatio-temporal scale-space representation L'' of f'' according to $L''(\cdot; \Sigma'', \delta'') = g(\cdot; \Sigma'', \delta'') * f''(\cdot)$. Then, from the transformation property of the Gaussian scale-space under Galilean transformations, which in turn originates from the corresponding transformation property under affine transformations (51), it follows that $L''(p'' \; \Sigma'', \delta'') = L(p; \Sigma, \delta)$ provided that the covariance matrices satisfy $\Sigma'' = G \Sigma G^T$ and that the time delays are equal $\delta'' = \delta$.

C.2 Spatio-Temporal Second-Moment Matrix/Structure Tensor

Let us next define *velocity-adapted spatio-temporal second-moment matrices (structure tensors)* of L and L'' according to

$$\mu(p; \Sigma_1, \Sigma_2, \delta_1 + \delta_2)$$
$$= \int_{q\in\mathbb{R}^{(N+1)}} (\nabla L(q; \Sigma_1, \delta_1))(\nabla L(q; \Sigma_1, \delta_1))^T$$
$$g(p - q; \Sigma_2, \delta_2) \, dq, \tag{191}$$

$$\mu''(p''; \Sigma_1'', \Sigma_2'', \delta_1'' + \delta_2'')$$
$$= \int_{q''\in\mathbb{R}^{(N+1)}} (\nabla L''(q''; \Sigma_1'', \delta_1''))(\nabla L''(q''; \Sigma_1'', \delta_1''))^T$$
$$g(p'' - q''; \Sigma_2'', \delta_2'') \, dq'', \tag{192}$$

where δ_1 and δ_2 denote the time constant in the first and second stages of scale-space smoothing, respectively. Then, from the general transformation property of second-moment matrices under affine transformations [39, Sect. 15.3], [51],

if follows that the second-moment matrices μ and μ'' are related according to

$$\mu'' = G^{-T}(u) \, \mu \, G^{-1}(u). \tag{193}$$

C.3 Galilean Block Diagonalization

Our next step is to introduce the notion of *Galilean block diagonalization*, which corresponds to finding the unique Galilean transformation that transforms the spatio-temporal second-moment matrix to block diagonal form with all mixed purely spatio-temporal components being zero $\mu'_{x_1 t} = \mu'_{x_2 t} = \cdots = \mu'_{x_N t} = 0$ [52]

$$\mu' = \begin{pmatrix} \mu'_{x_1 x_1} & \mu'_{x_1 x_2} & \cdots & \mu'_{x_1 x_N} & 0 \\ \mu'_{x_1 x_2} & \mu'_{x_2 x_2} & \cdots & \mu'_{x_2 x_N} & 0 \\ \vdots & \vdots & \ddots & & 0 \\ \mu'_{x_1 x_N} & \mu'_{x_2 x_N} & \cdots & \mu'_{x_N x_N} & 0 \\ 0 & 0 & 0 & 0 & \mu'_{tt} \end{pmatrix}. \tag{194}$$

Such a block diagonalization can be obtained if the velocity vector u satisfies

$$\begin{pmatrix} \mu'_{x_1 x_1} & \mu'_{x_1 x_2} & \cdots & \mu'_{x_1 x_N} \\ \mu'_{x_1 x_2} & \mu'_{x_2 x_2} & \cdots & \mu'_{x_2 x_N} \\ \vdots & \vdots & \ddots & \\ \mu'_{x_1 x_N} & \mu'_{x_2 x_N} & \cdots & \mu'_{x_N x_N} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}$$
$$= - \begin{pmatrix} \mu_{x_1 t} \\ \mu_{x_2 t} \\ \vdots \\ \mu_{x_N t} \end{pmatrix} \tag{195}$$

with the solution

$$u = -\{\mu_{xx}\}^{-1}\{\mu_{xt}\} \tag{196}$$

i.e., structurally similar equations as are used for computing optic flow according to the method by Lukas and Kanade [54]. This is a very general approach for normalizing local spatio-temporal image patterns, which also applies to spatio-temporal patterns that cannot be modelled by a Galilean transformation of an otherwise temporally stationary spatial pattern.

C.4 Galilean Invariant Property of Galilean Diagonalization

In view of the above mentioned definition of the notion of Galilean block diagonalization of the spatio-temporal second-moment matrix, let us next return to the previously stated general transformation property (193) of the Gaussian spatio-temporal scale-space under Galilean transformations. Given a certain spatio-temporal pattern L around a point p

in space-time, let us assume that we have a Galilean transformation $G(u)$ that transforms μ'' into block diagonal form

$$\mu'' = G^{-T}(u)\,\mu\,G^{-1}(u). \tag{197}$$

Let us moreover assume that the original spatio-temporal image pattern L is transformed by some unknown Galilean transformation with velocity vector v into a transformed scale-space representation L'. From (193), it then follows that the spatio-temporal second-moment matrix μ for this transformed image pattern can be written $\mu' = G^{-T}(v)\,\mu\,G^{-1}(v)$. This transformed spatio-temporal second-moment matrix can in turn be brought to a Galilean block diagonalized form by a velocity vector w such that $\mu'' = G^{-T}(w)\,\mu'\,G^{-1}(w)$. By combining the last two expressions, we thus obtain

$$\mu'' = G^{-T}(w)\,G^{-T}(v)\,\mu\,G^{-1}(v)\,G^{-1}(w). \tag{198}$$

Since Galilean transformation matrices satisfy $G^{-1}(v) = G(-v)$ as well as $G(v + w) = G(v)\,G(w)$, it follows that

$$\mu'' = G^{-T}(v + w)\,\mu\,G^{-1}(v + w) \tag{199}$$

and we have that the Galilean transformation $G(w) = G(u - v)$ will bring the second-moment matrix μ' of the transformed pattern into block diagonal form. Thus, the property of Galilean block diagonalization is preserved under Galilean transformations. Specifically, the velocity vector associated with the Galilean transformation, that brings a second-moment matrix into block diagonal form, is additive under superimposed Galilean transformations.

Therefore, if we normalize local space-time structures using a local Galilean transformations determined from the requirement that the second-moment matrix should be block diagonal, it follows that the result after normalization will always be the same, irrespective of any superimposed Galilean transformation. From this view-point, the notion of Galilean block diagonalization leads to a canonical Galilean invariant way of normalizing local spatio-temporal image structures.[14]

---

[14]Note that although a similar result could be expected from the view-point of optic flow computations according to the method by Lukas and Kanade [54], we have in this proof not made any assumption that the local spatio-temporal image structures within the support region of the window function should represent a local translational model. (The optic flow estimation method by Lukas and Kanade is derived from such an assumption.) Therefore this result applies to arbitrary types of space-time structures and spatio-temporal events. The only assumption we have made above is that the purely spatial component of the second-moment matrix is non-singular, *i.e.,* that $\det\{\mu_{xx}\} \neq 0$. If this assumption is violated, then the velocity vector u in the Galilean transformation $G(u)$ that diagonalizes the spatio-temporal second-moment matrix μ is not uniquely determined, and we have a situation with a local aperture problem. This indeterminacy will, however, not effect the Galilean normalization, since the indeterminacy will not effect the transformed pattern.

C.5 Fixed-Point Property: Adaptation of the Velocity Vectors to the Local Spatio-Temporal Image Structure

A pre-requisite for carrying out the proofs underlying the transformation properties above is that the spatio-temporal covariance matrices used for computing the second-moment matrices are related according to $\Sigma'' = G(u)\,\Sigma\,G^T(u)$ and $\Sigma' = G(v)\,\Sigma\,G^T(v)$. Thus, perfect Galilean invariance can only be expected if the shapes of the spatio-temporal smoothing kernels are coupled. Otherwise, the transformation properties will only be approximate. From this view-point, a scale-space concept that allows for velocity adaptation for *any* image velocity can therefore be motivated from the desire of achieving true Galilean invariance.

We can formally test if true Galilean invariance has been achieved by checking if the velocity estimate $\hat{u}$ according to (196) agrees with the velocity v of the velocity-adapted spatio-temporal filters used for computing the estimate. Then, the image measurements are in agreements with the assumptions used for computing them. In practice, an operational criterion of the form

$$\|\hat{u} - v\| \leq \varepsilon \tag{200}$$

can therefore be expected to sort out stable spatio-temporal image descriptors from unstable ones, with the interpretation that if this condition is satisfied, then the velocity parameter in the scale-space representation could be regarded as approximately matching an average velocity estimate for the local spatio-temporal image pattern [53].

In an actual implementation based on a limited set of filter parameters, we can also iteratively adapt the velocity parameter to previous image measurements and stop the iterations when the increments are below a threshold. Provided that this *iterative velocity adaption* procedure converges, the resulting image descriptors will then be Galilean invariant.

An underlying assumption for the iterative velocity adaptation approach to be applicable is that the velocity estimates are stationary over time intervals longer than combined temporal delay $\delta_1 + \delta_2$ corresponding to the time constants of the first- and second-layer scale-space filters. If the velocity estimates vary strongly on a much shorter time scale, the results of iterative velocity adaptation may, however, be unpredictable.

Let us finally remark that although the analysis in this appendix concerns the Gaussian spatio-temporal scale-space, a corresponding treatment can be performed for the time-causal spatio-temporal scale-space. Moreover, besides a Galilean invariant representation based on a *single* velocity parameter, there is also a potential in exploring representations based *multiple* values of the velocity parameter in a corresponding manner as multiple orientations of spatial receptive fields may be used in an advantageous manner in spatial vision.

Appendix D: Solutions for the Time-Causal Spatio-Temporal Scale-Space

In this section, we will study the solution of diffusion equations on semi-infinite domains, with emphasis on how time-dependent boundary conditions influence the temperature distribution in the interior. This type of physical model arises in time-causal scale-spaces on temporal and spatio-temporal domains. To simplify the treatment, let us begin with a simplified case with one temporal dimension and no spatial dimensions.

Please note that this treatment is given for introductory and tutorial purposes. For a detailed account of the solutions of the diffusion equation in different types of solids, please refer to Carslaw and Jaeger [9]. A difference compared to classical heat diffusion, however, is that we will here later consider diffusion over two temporal dimensions; one for the spatial scale parameter s and one for the ordinary time t.

D.1 Pure Temporal Time-Causal Scale-Space

Consider the solution of the diffusion equation

$$\partial_t L = \frac{1}{2} L_{\tau\tau} \tag{201}$$

on a semi-infinite solid with initial condition $L(0; \tau) = L_0(\tau)$ and a time-dependent boundary condition $L(t; 0) = f(t)$. This equation describes how an initial heat distribution $L_0(\tau)$ evolves over time t in an infinite solid with $\tau > 0$ while also being strongly influenced by a time varying temperature $f(t)$ at the boundary $\tau = 0$. With regard to a temporal scale-space representation, the temperature distribution at a non-infinitesimal distance τ from the boundary $\tau = 0$ will be interpreted as the time-causal scale-space representation of a temporal signal $f(t)$ at coarser time scales τ.

In [9, Sect. 14.2] it is shown that the solution of this one-dimensional equation can be written

$$L(t; \tau) = \int_{\zeta=0}^{\infty} L_0(\zeta)(g(\tau-\zeta; t) - g(\tau+\zeta; t))\, d\zeta$$
$$+ \int_{u=0}^{t} f(u)\,\phi(t-u; \tau)\, du \tag{202}$$

where

$$g(\tau; t) = \frac{1}{\sqrt{2\pi t}} e^{-\tau^2/2t} \quad \text{and}$$
$$\phi(t; \tau) = \frac{1}{\sqrt{2\pi}\, t^{3/2}}\, \tau\, e^{-\tau^2/2t}. \tag{203}$$

As can be seen, this solution consists of two terms. The first term describes how the initial heat distribution $L_0(\tau)$ evolves over time t in such a way that the influence of this heat distribution will be kept to zero at the boundary $\tau = 0$.

For this reason of keeping the boundary influence to zero, an artificial distribution of negative heat sinks $-L_0(-\tau)$ has been introduced on the negative τ axis, manifested in terms of the addition of a negative Gaussian in the convolution expression. The second term describes how the time dependent boundary condition $f(t)$ at $\tau = 0$ spreads into the heat conducting medium. Specifically, the kernel $\phi(t-u; \tau)$ arising in this term describes how an amount of heat $f(u)$ at the boundary at time $t-u$ spreads over time t into the medium at a penetration depth τ from the boundary. Due to the linearity of the diffusion equation, these two solutions can be superimposed, and the resulting solution will satisfy both the initial condition and the boundary condition. While one could in general initiate a temporal or spatio-temporal scale-space concept over time t with an initial distribution over temporal scales τ, we will henceforth simplify the treatment by setting the initial condition to zero and focus on the latter boundary component of the solution only. Thus, with initial condition $L(0; \tau) = L_0(\tau) = 0$ we will consider

$$L(t; \tau) = \int_{u=0}^{t} f(u)\,\phi(t-u; \tau)\, du \tag{204}$$

with

$$\phi(t; \tau) = \frac{1}{\sqrt{2\pi}\, t^{3/2}}\, \tau\, e^{-\tau^2/2t} \tag{205}$$

as the solution for the boundary-dependent part of the diffusion equation with the time-dependent boundary condition $L(t; 0) = f(t)$.

By verification, it can be shown that $\phi(t; \tau)$ satisfies the diffusion equation. More easily, by observing that $\phi(t; \tau) = -\partial_\tau g(\tau; t)$ we can also immediately see that $\phi(t; \tau)$ has to satisfy the diffusion equation, since the Gaussian $g(\tau; t)$ is a solution of the diffusion equation and the operator ∂_τ used for computing ϕ from g is a linear operator.

D.2 Fulfillment of the Boundary Condition

To verify that L according to (204) satisfies the boundary condition $L(t; 0) = f(t)$, we can start from the explicit integral expression

$$L(t; \tau) = \frac{1}{\sqrt{2\pi}} \int_{u=0}^{t} f(u)\, \frac{\tau}{(t-u)^{3/2}}\, e^{-\tau^2/2(t-u)}\, du \tag{206}$$

and perform a change of variables by [9]

$$\mu = \frac{\tau}{\sqrt{2(t-u)}} \quad \text{with } u = t - \frac{\tau^2}{2\mu^2} \text{ and } du = \frac{\tau^2}{\mu^3}\, d\mu \tag{207}$$

which transforms the integral expression into

$$L(t; \tau) = \frac{2}{\sqrt{\pi}} \int_{\mu=\tau/\sqrt{2t}}^{\infty} f\left(t - \frac{\tau^2}{2\mu^2}\right) e^{-\mu^2}\, d\mu. \tag{208}$$

If we would allow ourselves to take limits inside the integral, we would obtain

$$
\begin{aligned}
\lim_{\tau \downarrow 0} L(t; \tau) &= \lim_{\tau \downarrow 0} \frac{2}{\sqrt{\pi}} \int_{\mu=\tau/\sqrt{2t}}^{\infty} f\left(t - \frac{\tau^2}{2\mu^2}\right) e^{-\mu^2} d\mu \\
&= \frac{2}{\sqrt{\pi}} \int_{\mu=0}^{\infty} \left(\lim_{\tau \downarrow 0} f\left(t - \frac{\tau^2}{2\mu^2}\right)\right) e^{-\mu^2} d\mu \\
&= \frac{2}{\sqrt{\pi}} \int_{\mu=0}^{\infty} f(t) e^{-\mu^2} d\mu \\
&= f(t) \frac{2}{\sqrt{\pi}} \int_{\mu=0}^{\infty} e^{-\mu^2} d\mu = f(t)
\end{aligned}
\tag{209}
$$

and the fulfillment of the boundary condition would be proven. A more detailed proof showing that this conclusion can indeed be made valid is given in [48].

D.3 Temporal Cascade-Recursive Structure Over Time

In addition to the cascade smoothing property over temporal scales τ, the time-causal scale-space also obeys a recursive structure over time t. If we start from the general expression for the solution of the diffusion equation over a semi-infinite solid in (202), and let the initial condition be the temporal scale-space representation at time t_1, *i.e.*, $L_0(\zeta) = L_0(\tau) = L(t_1; \tau)$, then we can write the temporal scale-space representation $L(t_2; \tau)$ at time t_2, *i.e.*, after a time increment $t_2 - t_1$ as

$$
\begin{aligned}
L(t_2; \tau) &= \int_{\zeta=0}^{\infty} L(t_1; \zeta)(g(\tau - \zeta; t_2 - t_1) \\
&\qquad - g(\tau + \zeta; t_2 - t_1)) d\zeta \\
&\quad + \int_{u=t_1}^{t_2} f(u) \phi(t_2 - u; \tau) du.
\end{aligned}
\tag{210}
$$

If we at time $t = t_2$ consider $t = t_1$ as a divider of the history, we can from this expression explicitly see how $L(t_1; \zeta)$, when considered as a representation over all temporal scales ζ works as an internal memory of what has happened before $t = t_1$. This information is updated by integration with two shifted Gaussians, one of these negative, each with variance $t_2 - t_1$. The novel contribution of information between $t = t_1$ and t_2 is incorporated into the temporal scale-space representation L by convolving the input signal f with the time-causal scale-space kernel $\phi(t; \tau)$. If we in turn combine this expression with the definition of $L(t_i; \tau)$

$$
L(t_i; \tau) = \int_{u=0}^{t_i} f(u) \phi(t_i - u; \tau) du
\tag{211}
$$

and divide the integral for computing $L(t_i; \tau)$ over the intervals $[0, t_1]$ and $[t_1, t_2]$ we obtain

$$
\int_{u=0}^{t_1} f(u) \phi(t_2 - u; \tau) du + \int_{u=t_1}^{t_2} f(u) \phi(t_2 - u; \tau) du
$$

$$
\begin{aligned}
&= \int_{\zeta=0}^{\infty} \left(\int_{u=0}^{t_1} f(u) \phi(t_1 - u; \zeta) du\right) \\
&\qquad (g(\tau - \zeta; t_2 - t_1) - g(\tau + \zeta; t_2 - t_1)) d\zeta \\
&\quad + \int_{u=t_1}^{t_2} f(u) \phi(t_2 - u; \tau) du.
\end{aligned}
\tag{212}
$$

By changing the order of integration in the middle integral and then identifying first the integrals over the interval $[0, t_1]$

$$
\begin{aligned}
&\int_{u=0}^{t_1} f(u) \phi(t_2 - u; \tau) du \\
&= \int_{u=0}^{t_1} f(u) \left(\int_{\zeta=0}^{\infty} \phi(t_1 - u; \zeta) (g(\tau - \zeta; t_2 - t_1) \right. \\
&\qquad \left. - g(\tau + \zeta; t_2 - t_1)) d\zeta\right) du
\end{aligned}
\tag{213}
$$

and then also identifying the arguments of these integrals with the complementary motivation that this relation should hold for all sufficiently regular f, we can after replacing the arguments $t_i - u$ by t_i state that

$$
\begin{aligned}
\phi(t_2; \tau) &= \int_{\zeta=0}^{\infty} \phi(t_1; \zeta) (g(\tau - \zeta; t_2 - t_1) \\
&\qquad - g(\tau + \zeta; t_2 - t_1)) d\zeta.
\end{aligned}
\tag{214}
$$

This is the time-recursive cascade smoothing property of the time-causal scale-space kernel $\phi(t; \tau)$ over time t.

Appendix E: Formal Statements Concerning the Time-Causal Spatio-Temporal Scale-Space

This appendix contains formal statements for the time-causal spatio-temporal scale-space analogous to Definition 1, Lemma 2, Definitions 3, 4 and Theorem 5 in Sect. 3.2 regarding the generalized Gaussian scale-space; please refer to Sect. 5 for an overview.

Definition 11 (Continuous time-causal pre-scale-space representation) Let $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+)$ be a continuous spatio-temporal signal and let $\mathcal{T}(s, t)$ with $(s, t) \in \mathbb{R}_+^2$ be a strongly continuous two-parameter semi-group of linear and shift-invariant operators from $L^2(\mathbb{R}^N \times \mathbb{R}_+)$ to $L^2(\mathbb{R}^N \times \mathbb{R}_+)$ according to

$$
\begin{cases}
\mathcal{T}(s_1, t_1)\, \mathcal{T}(s_2, t_2) = \mathcal{T}(s_1 + s_2, t_1 + t_2), \\
\mathcal{T}(0, 0) = \mathcal{I}
\end{cases}
\tag{215}
$$

and (70), where the semi-group corresponds to a spatio-temporal convolution operation of the form (67) with the convolution kernels $h(x, t; s, \tau) \in L^1(\mathbb{R}^N \times \mathbb{R})$ and the smoothing functions $T(x, t; s, \tau, \zeta)$ in the time-recursive

update rule (66) also being in $L^1(\mathbb{R}^N \times \mathbb{R})$. The semi-group is also required to be C_1 continuous with respect to the L^2-based Sobolev norm $\| \cdot \|_{H^k(\mathbb{R}^N \times \mathbb{R}_+)}$ in the sense that for all connected regions $\Omega \in \mathbb{R}_+^2$ near the origin that shrink to zero in such a way that the maximum distance $\rho(\Omega)$ between a point in Ω and the origin tends to zero

$$\lim_{\rho(\Omega) \downarrow 0} \left\| \frac{\int_{(s,t) \in \Omega} \mathcal{T}(s,t) f \, d\Omega}{\int_{(s,t) \in \Omega} d\Omega} - f \right\|_{H^k(\mathbb{R}^N \times \mathbb{R}_+)} = 0 \qquad (216)$$

should for some $k > (N + 1)/2$ be required to hold for all smooth functions $f \in L^1(\mathbb{R}^N \times \mathbb{R}_+) \cap C^\infty(\mathbb{R}^N \times \mathbb{R}_+)$. Then, the two-parameter family of signals $L: \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ given by

$$L(\cdot, t; \, s, \cdot) = \mathcal{T}(s,t) \, L(\cdot, 0; \, 0, \cdot) \qquad (217)$$

with boundary condition $L(x, t; \, 0, 0) = f(x, t)$ and combined initial and boundary condition $L(x, 0; \, 0, \tau) = L_0(x; \, \tau)$ is said to be a continuous time-causal pre-scale-space representation of f generated by $\mathcal{T}(s,t)$.

Lemma 12 (A continuous time-causal pre-scale-space representation is differentiable) *Let $L: \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ be the continuous time-causal pre-scale-space representation of a smooth spatio-temporal signal $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+) \cap C^\infty(\mathbb{R}^N \times \mathbb{R}_+)$. Then, L satisfies a two-parameter differential equation with regard to the parameters of the semi-group with the directional derivative of the semi-group in direction $u = (\alpha_1, \alpha_2)$ given by*

$$\mathcal{D}_u L = (\alpha_1 \, \mathcal{A}_1 + \alpha_2 \, \mathcal{A}_2) L \qquad (218)$$

for some linear and shift-invariant operators $\mathcal{A}_1$ and $\mathcal{A}_2$ from $L^2(\mathbb{R}^N \times \mathbb{R}_+)$ to $L^2(\mathbb{R}^N \times \mathbb{R}_+)$.

Proof Follows from results in [1, page 407] reviewed in connection with (71)–(75) in Sect. 5. □

Definition 13 (Time-causal pre-scale-space property: Non-enhancement of local extrema) A time-causal continuous pre-scale-space representation $L: \mathbb{R}^N \times \times \mathbb{R}_+ \mathbb{R}_+^2 \to \mathbb{R}$ of a smooth signal $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+) \cap C^\infty(\mathbb{R}^N \times \mathbb{R}_+)$ is said to possess continuous non-enhancement pre-scale-space properties, or equivalently not to enhance local extrema, if for any time $t_0 \in \mathbb{R}_+$ and any scale $s_0 \in \mathbb{R}_+$ it holds that if $(x_0, \tau_0) \in \mathbb{R}^N \times \mathbb{R}_+$ is a critical point for the mapping $(x, \tau) \mapsto L(x, t_0; \, s_0, \tau)$ and if the Hessian matrix $\mathcal{H}_{x,\tau}$ with respect to both space x and temporal scales τ at this point is non-degenerate, then the (semi-group) directional derivative of L in any direction $u = (\alpha_1, \alpha_2)$ in (s, t) space at this point has the same sign as the Hessian matrix, *i.e.,*

$$\text{sign } \partial_u L = \text{sign trace } \mathcal{H}_{x,\tau} L. \qquad (219)$$

Remark The definition of local extrema underlying this definition, *i.e.,* local extrema of the mapping $(x, \tau) \mapsto L(x, t_0; \, s_0, \tau)$ for every scale $s_0 \in \mathbb{R}_+$, means that these local extrema can be detected from the always available internal buffer over $(x; \, s, \tau)$ at any time moment t_0 and do not imply any explicit references to the future $t > t_0$ or the past $t < t_0$. In this respect, the notion of non-enhancement of local extrema for the time-causal spatio-temporal scale-space differs from the notion of non-enhancement of local extrema for the Gaussian spatio-temporal scale-space, where non-enhancement of local extrema is expressed in terms of local extrema of the mapping $(x, t) \mapsto L(x, t; \, s_0, \tau_0)$.

Definition 14 (Time-causal continuous non-enhancement scale-space representation) Let $\mathcal{T}(s, t)$ be a strongly continuous two-parameter semi-group of linear and shift-invariant operators from $L^2(\mathbb{R}^N \times \mathbb{R}_+)$ to $L^2(\mathbb{R}^N \times \mathbb{R}_+)$. Given a spatio-temporal signal $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+)$, the pre-scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ of f is said to be a time-causal continuous scale-space representation of f if and only if it for every smooth function $f' \in L^2(\mathbb{R}^N \times \mathbb{R}_+) \cap C^\infty(\mathbb{R}^N \times \mathbb{R}_+)$ of compact support and for every initial condition $L(x, 0; \, 0, \tau) = L_0(x; \, \tau)$ it holds that the time-causal pre-scale-space representation $L': \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ of f' generated by $\mathcal{T}(s, t)$ obeys non-enhancement of local extrema.

Lemma 15 (Time-causal non-enhancement scale-space for continuous signals: Necessity I) *A continuous time-causal non-enhancement scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ of a spatio-temporal signal $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+)$ satisfies the following system of parabolic differential equations*

$$\partial_s L = \frac{1}{2} \nabla_{x,\tau}^T \left(\Sigma_1 \nabla_{x,\tau} L \right) - \delta_1^T \nabla_{x,\tau} L, \qquad (220)$$

$$\partial_t L = \frac{1}{2} \nabla_{x,\tau}^T \left(\Sigma_2 \nabla_{x,\tau} L \right) - \delta_2^T \nabla_{x,\tau} L, \qquad (221)$$

with boundary condition $L(x, t \, 0, 0) = f(x, t)$ for some positive semi-definite (spatial) covariance matrices Σ_1 and Σ_2 and some velocity vectors δ_1 and δ_2.

Proof A proof that the evolution property over spatial scales s has to be of the form in (220) can be performed by constructing a similar set of counterexamples from functions $f_i(x, t) = f_i(x)$ $i = 1..5$ with $f_i(x)$ given by the five functions used for constructing counterexamples in the proof of Theorem 5.

A proof that the evolution property over time t has to be of the form in (221) can be performed by constructing a similar set of counterexamples determined by $L_0(x; \, \tau) = f_i(\tau)$ with $f_i(\tau)$ determined by the five counterexamples in the proof of Theorem 5.

The existence of derivatives with respect to space x as used in the first part of the proof follows from the assumption of the convolution kernels $h(x, t; s, \tau)$ in (67) being in $L^1(\mathbb{R}^N \times \mathbb{R})$, which implies that we for smooth functions f with compact support have

$$\partial_{x_j}(h(\cdot; s, \tau) * f(\cdot)) = h(\cdot; s, \tau) * (\partial_{x_j} f)(\cdot)). \qquad (222)$$

The existence of derivatives with respect to temporal scale τ as used in the second part of the proof follows from the assumption of the smoothing functions $T(\cdot, t; s, \tau, \cdot)$ in the time recursive formulation (66) being in $L^1(\mathbb{R}^N \times \mathbb{R})$. $\qquad \square$

Out of this family of differential equations, it is not necessarily the case that all differential equations give rise to reasonable evolution processes with regard to interpretations in terms of spatial scales and temporal scales. More generally, given a two-parameter semi-group, one may also conceive different ways of parameterizing the same semi-group. In addition to the above mentioned structural conditions, we would, however, also would like to have the ability to interpret the first parameter s in the two-parameter semi-group as a spatial scale parameter and the second parameter t as regular time t. We say that a parameterization is on *standard form* if such an interpretation is possible.

Definition 16 (Standard form for the spatio-temporal scale-space) A system of partial differential equations describing a spatio-temporal scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ of a spatio-temporal signal $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+)$ is said to be on standard form if

 (i) the parameter s can be interpreted as a spatial scale parameter in the sense that the evolution equation over s does not explicitly depend on derivatives with respect to temporal scale τ,
 (ii) the evolution over spatial scales s does not imply any translation in space in the sense that the coefficients of the first-order derivatives with respect to space x and temporal scale τ are zero,
 (iii) the evolution equation over spatial scales s is stationary over time in the sense that the coefficients do not depend on time t,
 (iv) the evolution equation over time t should not be affected by the spatial scales in the sense that the coefficients of the second-order derivatives with respect to space x must be zero,
 (v) the evolution equation over time t does not involve temporal scale dependent translation in time in the sense that the coefficient of the first-order derivative with respect to temporal scale τ is zero.

Given these conditions, the evolution equations reduce to a simpler form:

Theorem 17 (Time-causal non-enhancement scale-space for continuous signals: Necessity II) *A continuous time-causal non-enhancement scale-space representation $L: \mathbb{R}^N \times \mathbb{R}_+ \times \mathbb{R}_+^2 \to \mathbb{R}$ of a spatio-temporal signal $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+)$ on standard form satisfies the following system of parabolic differential equations*

$$\partial_s L = \frac{1}{2} \nabla_x^T (\Sigma_0 \nabla_x L), \qquad (223)$$

$$\partial_t L = -v^T \nabla_x L + \frac{1}{2} \partial_{\tau\tau} L \qquad (224)$$

with initial condition $L(x, t\, 0, 0) = f(x, t)$ for some positive semi-definite (spatial) covariance matrix Σ_0 and some velocity vector v.

Proof Lemma 15 implies that the spatio-temporal scale-space representation has to satisfy the evolution equations (220) and (221).

Property (i) implies that $\nabla_{x,\tau}^T \Sigma_1 \nabla_{x,\tau} L$ reduces to $\nabla_{x,\tau}^T \Sigma_1 \nabla_{x,\tau} L$. From property (ii) it follows that $\delta_1 = 0$. From property (iii) we have that Σ_1 does not depend on time t.

From property (iv) we obtain that $\nabla_{x,\tau}^T (\Sigma_2 \nabla_{x,\tau} L)$ reduces to $\partial_{\tau\tau} L$. Property (v) implies that $\delta_2^T \nabla_{x,\tau} L$ reduces to $\delta_2^T \nabla_x L$ and we can rename the x component of δ_2 into v. $\qquad \square$

Theorem 18 (Time-causal non-enhancement scale-space for continuous signals: Sufficiency) *Given a semi-definite (spatial) covariance matrix Σ_0, an arbitrary vector v and any twice continuously differentiable function $f \in L^2(\mathbb{R}^N \times \mathbb{R}_+)$, the solution of the diffusion equation*

$$\partial_s L = \frac{1}{2} \nabla_x^T (\Sigma_0 \nabla_x L), \qquad (225)$$

$$\partial_t L = -v^T \nabla_x L + \frac{1}{2} \partial_{\tau\tau} L \qquad (226)$$

with boundary condition $L(x, t\, 0, 0) = f(x, t)$ does for every combined initial and boundary condition $L(x, 0; 0, \tau) = L_0(x; \tau)$ constitute a time-causal continuous non-enhancement scale-space representation of f. Specifically, for any (forward) direction u in (s, t) space L obeys

$$\partial_u L \leq 0 \quad \text{at any non-degenerate local maximum,} \qquad (227)$$

$$\partial_u L \geq 0 \quad \text{at any non-degenerate local minimum.} \qquad (228)$$

Proof If (x_0, τ_0) is a local extremum of the mapping $(x, \tau) \mapsto L(x, t_0; s_0, \tau)$, then x_0 is also a local extremum of the mapping $x \mapsto L(x, t_0; s_0, \tau)$. From Theorem 6 it then follows that $\partial_s L \leq 0$ if the point is a local maximum and $\partial_s L \geq 0$ if the point is a local minimum.

Similarly, if (x_0, τ_0) is a local extremum of the mapping $(x, \tau) \mapsto L(x, t_0; s_0, \tau)$, then τ_0 is also a local extremum of

the mapping $\tau \mapsto L(x, t_0; s_0, \tau)$. From Theorem 6 it then follows that $\partial_t L \leq 0$ if the point is a local maximum and $\partial_t L \geq 0$ if the point is a local minimum.

Since $\partial_u L = \alpha_1 \partial_s L + \alpha_2 \partial_t L$, where $\alpha_1, \alpha_2 \geq 0$ for a forward direction in (s, t) space, we have that sign $\partial_u L =$ sign trace $\mathcal{H}_{x,\tau} L$ and the result follows. □

References

1. Al-Sharif, S., Khalil, R.: On the generator of two parameter semi-groups. Appl. Math. Comput. **156**(2), 403–414 (2004)
2. Almansa, A., Lindeberg, T.: Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale-selection. IEEE Trans. Image Process. **9**(12), 2027–2042 (2000)
3. Alvarez, L., Guichard, F., Lions, P.-L., Morel, J.-M.: Axioms and fundamental equations of image processing. Arch. Ration. Mech. Anal. **123**(3), 199–257 (1993)
4. Babaud, J., Witkin, A.P., Baudin, M., Duda, R.O.: Uniqueness of the Gaussian kernel for scale-space filtering. IEEE Trans. Pattern Anal. Mach. Intell. **8**(1), 26–33 (1986)
5. Ballester, C., Gonzalez, M.: Affine invariant texture segmentation and shape from texture by variational methods. J. Math. Imaging Vis. **9**, 141–171 (1998)
6. Baumberg, A.: Reliable feature matching across widely separated views. In: Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, pp. I:1774–1781. Hilton Head, SC (2000)
7. Broadbridge, P.: Entropy diagnostics for fourth order partial differential equations in conservation form. Entropy **10**, 365–379 (2008). doi:10.3390/e10030365
8. Burt, P.J.: Fast filter transforms for image processing. Comput. Vis. Graph. Image Process. **16**, 20–51 (1981)
9. Carslaw, H.S., Jaeger, J.C.: Conduction of Heat in Solids. Clarendon Press, Oxford (1959)
10. Crowley, J.L.: *A representation for visual information*. PhD thesis, Carnegie-Mellon University, Robotics Institute, Pittsburgh, Pennsylvania (1981)
11. DeAngelis, G.C., Ohzawa, I., Freeman, R.D.: Receptive field dynamics in the central visual pathways. Trends Neurosci. **18**(10), 451–457 (1995)
12. Duits, R., Florack, L., de Graaf, J., ter Haar Romeny, B.: On the axioms of scale space theory. J. Math. Imaging Vis. **22**, 267–298 (2004)
13. Duits, R., Felsberg, M., Florack, L., Platel, B.: α-scale-spaces on a bounded domain. In: Griffin, L., Lillholm, M. (eds.) Proc. Scale-Space Methods in Computer Vision: Scale-Space'03, Isle of Skye, Scotland, Jun. 2003. Lecture Notes in Computer Science, vol. 2695, pp. 494–510. Springer, Berlin (2003)
14. Dunninger, D.R.: Maximum principles for solutions of some fourth-order elliptic equation. J. Math. Anal. Appl. **37**, 665–668 (1972)
15. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19. American Mathematical Society, Providence (1998)
16. Fagerström, D.: Temporal scale-spaces. Int. J. Comput. Vis. **2–3**, 97–106 (2005)
17. Fagerström, D.: Spatio-temporal scale-spaces. In: Gallari, F., Murli, A., Paragios, N. (eds.) Proc. 1st International Conference on Scale-Space Theories and Variational Methods in Computer Vision. Lecture Notes in Computer Science, vol. 4485, pp. 326–337. Springer, Berlin (2007)
18. Felsberg, M., Sommer, G.: The monogenic scale-space: A unifying approach to phase-based image processing in scale-space. J. Math. Imaging Vis. **21**, 5–26 (2004)
19. Fleet, D.J., Langley, K.: Recursive filters for optical flow. IEEE Trans. Pattern Anal. Mach. Intell. **17**(1), 61–67 (1995)
20. Florack, L.M.J.: Image Structure. Series in Mathematical Imaging and Vision. Springer, Dordrecht (1997)
21. Florack, L.M.J., ter Haar Romeny, B.M., Koenderink, J.J., Viergever, M.A.: Scale and the differential structure of images. Image Vis. Comput. **10**(6), 376–388 (1992)
22. Florack, L.M.J., Niessen, W., Nielsen, M.: The intrinsic structure of optic flow incorporating measurement duality. Int. J. Comput. Vis. **27**(3), 263–286 (1998)
23. Folland, G.B.: Introduction to Partial Differential Equations, 2nd edn. Princeton University Press, Princeton (1995)
24. Freeman, W.T., Adelson, E.H.: The design and use of steerable filters. IEEE Trans. Pattern Anal. Mach. Intell. **13**(9), 891–906 (1991)
25. Gårding, J., Lindeberg, T.: Direct computation of shape cues using scale-adapted spatial derivative operators. Int. J. Comput. Vis. **17**(2), 163–191 (1996)
26. Goldstein, J.A.: Semigroups of Linear Operators and Applications. Oxford Mathematical Monographs. Oxford Science, Oxford (1985)
27. Griffin, L.: Critical point events in affine scale space. In: Sporring, J., Nielsen, M., Florack, L., Johansen, P. (eds.) Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory, Copenhagen, Denmark, May 1996, pp. 165–180. Springer, Berlin (1996)
28. Guichard, F.: A morphological, affine, and Galilean invariant scale-space for movies. IEEE Trans. Image Process. **7**(3), 444–456 (1998)
29. Hille, E., Phillips, R.S.: Functional Analysis and Semi-Groups. American Mathematical Society Colloquium Publications, vol. 31 (1957)
30. Hummel, R.A., Moniot, R.: Reconstructions from zero-crossings in scale-space. IEEE Trans. Acoust. Speech Signal Process. **37**(12), 2111–2130 (1989)
31. Iijima, T.: Observation theory of two-dimensional visual patterns. Technical report, Papers of Technical Group on Automata and Automatic Control, IECE, Japan (1962)
32. Koenderink, J.J.: The structure of images. Biol. Cybern. **50**, 363–370 (1984)
33. Koenderink, J.J.: Scale-time. Biol. Cybern. **58**, 159–162 (1988)
34. Koenderink, J.J., van Doorn, A.J.: Receptive field families. Biol. Cybern. **63**, 291–298 (1990)
35. Koenderink, J.J., van Doorn, A.J.: Generic neighborhood operators. IEEE Trans. Pattern Anal. Mach. Intell. **14**(6), 597–605 (1992)
36. Laptev, I., Lindeberg, T.: Velocity-adapted spatio-temporal receptive fields for direct recognition of activities. Image Vis. Comput. **22**(2), 105–116 (2004)
37. Laptev, I., Caputo, B., Schuldt, C., Lindeberg, T.: Local velocity-adapted motion events for spatio-temporal recognition. Comput. Vis. Image Underst. **108**, 207–229 (2007)
38. Lindeberg, T.: Scale-space for discrete signals. IEEE Trans. Pattern Anal. Mach. Intell. **12**(3), 234–254 (1990)
39. Lindeberg, T.: Scale-Space Theory in Computer Vision. The Kluwer International Series in Engineering and Computer Science. Springer, Dordrecht (1994)
40. Lindeberg, T.: Scale-space theory: A basic tool for analysing structures at different scales. J. Appl. Stat. **21**(2), 225–270 (1994). Also available from http://www.csc.kth.se/~tony/abstracts/Lin94-SI-abstract.html
41. Lindeberg, T.: On the axiomatic foundations of linear scale-space. In: Sporring, J., Nielsen, M., Florack, L., Johansen, P. (eds.) Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory, Copenhagen, Denmark, May, 1996. Springer, Berlin (1996)

42. Lindeberg, T.: On automatic selection of temporal scales in time-casual scale-space. In: Sommer, G., Koenderink, J.J. (eds.) Proc. AFPAC'97: Algebraic Frames for the Perception-Action Cycle, Kiel, Germany, Sep., 1997. Lecture Notes in Computer Science, vol. 1315, pp. 94–113. Springer, Berlin (1997)

43. Lindeberg, T.: Linear spatio-temporal scale-space. In: ter Haar Romeny, B.M., Florack, L.M.J., Koenderink, J.J., Viergever, M.A. (eds.) Scale-Space Theory in Computer Vision: Proc. First Int. Conf. Scale-Space'97, Utrecht, The Netherlands, July, 1997. Lecture Notes in Computer Science, vol. 1252, pp. 113–127. Springer, Berlin (1997). Extended version available as technical report ISRN KTH NA/P–01/22–SE from KTH

44. Lindeberg, T.: Feature detection with automatic scale selection. Int. J. Comput. Vis. **30**(2), 77–116 (1998)

45. Lindeberg, T.: Linear spatio-temporal scale-space. Report, ISRN KTH/NA/P–01/22–SE, Dept. of Numerical Analysis and Computing Science, KTH, Nov. (2001)

46. Lindeberg, T.: Time-recursive velocity-adapted spatio-temporal scale-space filters. In: Johansen, P. (ed.) Proc. 7th European Conference on Computer Vision, Copenhagen, Denmark, May 2002. Lecture Notes in Computer Science, vol. 2350, pp. 52–67. Springer, Berlin (2002)

47. Lindeberg, T.: Scale-space. In: Wah, B. (ed.) Encyclopedia of Computer Science and Engineering, pp. 2495–2504. Wiley, Hoboken (2008). doi:10.1002/9780470050118.ecse609. Also available from http://www.nada.kth.se/~tony/abstracts/Lin08-EncCompSci.html

48. Lindeberg, T.: Generalized Gaussian scale-space axiomatics comprising linear scale-space, affine scale-space and spatio-temporal scale-space. Technical report TRITA-CSC-CV 2010:3, ISSN 1653-7092, School of Computer Science and Communication, KTH (Royal Institute of Technology), Stockholm, Sweden, 2010. Also available from http://www.csc.kth.se/~tony/abstracts/Lin10-GenGaussScSp.html

49. Lindeberg, T., Fagerström, D.: Scale-space with causal time direction. In: Proc. 4th European Conf. on Computer Vision, Cambridge, UK, Apr, 1996, vol. 1064, pp. 229–240. Springer, Berlin (1996)

50. Lindeberg, T., Gårding, J.: Shape-adapted smoothing in estimation of 3-D depth cues from affine distortions of local 2-D structure. In: Eklundh, J.-O. (ed.) Proc. 3rd European Conf. on Computer Vision, Stockholm, Sweden, May 1994. Lecture Notes in Computer Science, vol. 800, pp. 389–400. Springer, Berlin (1994)

51. Lindeberg, T., Gårding, J.: Shape-adapted smoothing in estimation of 3-D depth cues from affine distortions of local 2-D structure. Image Vis. Comput. **15**, 415–434 (1997)

52. Lindeberg, T., Akbarzadeh, A., Laptev, I.: Galilean-corrected spatio-temporal interest operators. In: International Conference on Pattern Recognition, Cambridge, pp. I:57–62. (2004). Extended version available as technical report ISRN KTH NA/P–04/05–S from KTH

53. Lindeberg, T., Akbarzadeh, A., Laptev, I.: Galilean-corrected spatio-temporal interest operators. Technical Report ISRN KTH/NA/P–04/05–SE, Dept. of Numerical Analysis and Computing Science, KTH, Mar. 2004

54. Lukas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Image Understanding Workshop (1981)

55. Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. Int. J. Comput. Vis. **60**(1), 63–86 (2004)

56. Nagel, H., Gehrke, A.: Spatiotemporal adaptive filtering for estimation and segmentation of optical flow fields. In: Proc. 5th European Conf. on Computer Vision, Freiburg, Germany, Jun. 1998, pp. 86–102. Springer, Berlin (1998)

57. Pauwels, E.J., Fiddelaers, P., Moons, T., van Gool, L.J.: An extended class of scale-invariant and recursive scale-space filters. IEEE Trans. Pattern Anal. Mach. Intell. **17**(7), 691–701 (1995)

58. Pazy, A.: Semi-groups of Linear Operators and Applications to Partical Differential Equations. Applied Mathematical Sciences. Springer, Berlin (1983)

59. Perona, P.: Steerable-scalable kernels for edge detection and junction analysis. Image Vis. Comput. **10**, 663–672 (1992)

60. Sato, K.-I.: Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1999)

61. Schaffalitzky, F., Zisserman, A.: Viewpoint invariant texture matching and wide baseline stereo. In: Proc. 8th Int. Conf. on Computer Vision, Vancouver, Canada, July 2001, pp. II:636–643. (2001)

62. Schoenberg, I.J.: On Pòlya frequency functions. II. Variation-diminishing integral operators of the convolution type. Acta Sci. Math. (Szeged) **12**, 97–106 (1950)

63. Schoenberg, I.J.: On smoothing operations and their generating functions. Bull. Am. Math. Soc. **59**, 199–230 (1953)

64. Simoncelli, E.P., Freeman, W.T., Adelson, E.H., Heeger, D.J.: Shiftable multi-scale transforms. IEEE Trans. Information Theory **38**(2), 587–607 (1992)

65. Sporring, J., Nielsen, M., Florack, L., Johansen, P. (eds.): Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory. Series in Mathematical Imaging and Vision. Springer, Copenhagen (1996)

66. ter Haar Romeny, B. (ed.): Geometry-Driven Diffusion in Computer Vision. Series in Mathematical Imaging and Vision. Springer, Dordrecht (1994)

67. ter Haar Romeny, B.: Front-End Vision and Multi-Scale Image Analysis. Springer, Dordrecht (2003)

68. ter Haar Romeny, B., Florack, L., Nielsen, M.: Scale-time kernels and models. In: Scale-Space and Morphology: Proc. Scale-Space'01, Vancouver, Canada, July 2001. Lecture Notes in Computer Science. Springer, Berlin (2001)

69. Tuytelaars, T., van Gool, L.: Matching widely separated views based on affine invariant regions. Int. J. Comput. Vis. **59**(1), 61–85 (2004)

70. Valois, R.L.D., Cottaris, N.P., Mahon, L.E., Elfer, S.D., Wilson, J.A.: Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. Vis. Res. **40**(2), 3685–3702 (2000)

71. Weickert, J.: Anisotropic Diffusion in Image Processing. Teubner-Verlag, Stuttgart (1998)

72. Weickert, J., Ishikawa, S., Imiya, A.: Linear scale-space has first been proposed in Japan. J. Math. Imaging Vis. **10**(3), 237–252 (1999)

73. Witkin, A.P.: Scale-space filtering. In: Proc. 8th Int. Joint Conf. Art. Intell., Karlsruhe, Germany, Aug. 1983, pp. 1019–1022 (1983)

74. Young, R.A.: The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles. Technical Report GMR-4920, Computer Science Department, General Motors Research Lab., Warren, Michigan (1985)

75. Young, R.A.: The Gaussian derivative model for spatial vision: I. Retinal mechanisms. Spat. Vis. **2**, 273–293 (1987)

76. Young, R.A., Lesperance, R.M.: The Gaussian derivative model for spatio-temporal vision: II. Cortical data. Spat. Vis. **14**(3, 4), 321–389 (2001)

77. Young, R.A., Lesperance, R.M., Meyer, W.W.: The Gaussian derivative model for spatio-temporal vision: I. Cortical model. Spat. Vis. **14**(3, 4), 261–319 (2001)

78. Yuille, A.L., Poggio, T.A.: Scaling theorems for zero-crossings. IEEE Trans. Pattern Anal. Mach. Intell. **8**, 15–25 (1986)

79. Zhang, H., Zhang, W.: Maximum principles and bounds in a class of fourth-order uniformly elliptic equations. J. Phys. A: Math. Gen. **35**, 9245–9250 (2002)

Tony Lindeberg is a Professor in Computer Science at KTH (Royal Institute of Technology) in Stockholm, Sweden and a Research Fellow at the Royal Swedish Academy of Sciences. He was born in Stockholm, Sweden in 1964, received his MSc degree in 1987, his PhD degree in 1991, became docent in 1996, and was appointed professor in 2000.

His research interests in computer vision relate to scale-space representation, image features, object recognition, spatio-temporal recognition, focus-of-attention and shape. He has developed theories and methodologies for continuous and discrete scale-space representation, detection of salient image structures, automatic scale selection, scale-invariant image features, affine invariant features, affine and Galilean normalization, temporal and spatio-temporal scale-space concepts as well as spatial and spatio-temporal image descriptors for image-based recognition. He has also worked on applications in medical image analysis and gesture recognition. He is author of the book Scale-Space Theory in Computer Vision.