

On Automatic Selection of Temporal Scales in Time-Causal Scale-Space *

Tony Lindeberg

Computational Vision and Active Perception Laboratory (CVAP)
Department of Numerical Analysis and Computing Science
KTH, S-100 44 Stockholm, Sweden

Abstract. This paper outlines a general framework for automatic selection in temporal scale-space representations, and shows how the suggested theory applies to motion detection and motion estimation.

1 Introduction

A fundamental constraint on the design of a vision system originates from the fact that image structures are perceived as meaningful entities only over certain ranges of scale. In general situations, it is hardly ever possible to know in advance at what scales interesting structures can be expected to appear. For this reason, an image representation that explicitly incorporates the notion of scale is a crucially important tool when dealing with sensory data, such as images.

A multi-scale representation by itself, however, contains no explicit information about what image structures should be regarded as significant or what scales are appropriate for treating those. Early work addressing these problems for blob-like image structures was presented in (Lindeberg 1993a), leading to the notion of a scale-space primal sketch. Then, in (Lindeberg 1993b, 1996b) an extension to other aspects of image structures was presented by selecting scales for differential feature detectors (such as blobs, corners, edges and ridges) from maxima over scales of normalized differential entities.

The subject of this article is to address the problem of scale selection in the temporal domain, in order to deal with image data over time. Whereas, it is now rather generally accepted that some kind of “smoothing over time” is necessary when processing time-varying images, most current work on motion analysis is still carried out at a single temporal scale (see, e.g., (Barron *et al.* 1994; Beauchemin and Barron 1995)).

A main argument which will be advocated in this article, is that in analogy to earlier advances on spatial domains, the performance and robustness of algorithms operating over time can be improved substantially, if the spatio-temporal image data are considered at *several temporal scales simultaneously*, and if we incorporate *explicit mechanisms for automatic selection of temporal scales*.

* Technical report ISRN KTH NA/P-97/09-SE. Department of Numerical Analysis and Computing Science, Royal Institute of Technology, S-100 44 Stockholm, Sweden, Sep 1997. Also presented in Proc. AFPAC'97: Algebraic Frames for the Perception-Action Cycle (G. Sommer and J. J. Koenderink, eds.), vol. 1315 of Lecture Notes in Computer Science, (Kiel, Germany), pp. 94-113, Springer Verlag, Berlin, Sept. 1997.

To form the basis of a theory for temporal scale selection, we will start by showing how *time-causal normalized scale-space derivatives* can be defined for different types of time-causal scale-space concepts. Then, an adaptation of a previously proposed heuristic principle will be presented, stating that in the absence of further information, important clues for spatio-temporal scale selection can be obtained from the scales at which (possibly non-linear) combinations of *normalized spatio-temporal derivatives assume maxima over scales*. Specifically, it will be shown how this approach applies to motion detection and velocity estimation.

2 Spatial and temporal scale-space: Overview

Traditionally, most work on scale-space representation has been concerned with the spatial domain, in which the values of the input signal are available in all coordinate directions. Given any D -dimensional signal $f: \mathbb{R}^D \rightarrow \mathbb{R}$, its (spatial) scale-space representation $L: \mathbb{R}^D \times \mathbb{R}_+ \rightarrow \mathbb{R}$ is defined by convolution

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

with the (rotationally symmetric) Gaussian kernel

$$g(x; s) = \frac{1}{(2\pi s)^{N/2}} e^{-x^T x / 2s} \quad (2)$$

and scale-space derivatives are defined from this representation by $L_{x^\alpha}(\cdot; s) = \partial_{x^\alpha} L(\cdot; s)$ where $s \in \mathbb{R}_+$ is the scale parameter and $\alpha = (\alpha_1, \dots, \alpha_D)$ represents the order of differentiation. As has been shown by several authors (Witkin 1983; Koenderink 1984; Yuille and Poggio 1986; Koenderink and van Doorn 1992; Florack 1993; Lindeberg 1994; Pauwels *et al.* 1995), the choice of the Gaussian kernel and its derivatives is basically a unique choice, given natural assumptions on a visual front-end (scale-space axioms).

This scale-space concept, however, cannot be directly applied to temporal data, since in a real-time situation it is essential that image operators do not extend into the future. One suggestion for how to deal with this problem was given by (Koenderink 1988), who proposed to transform the time axis so as to map the present moment to the unreachable infinity. In the transformed domain, he then applied the traditional scale-space concept given by (1) and (2). Based on a classification of scale-space kernels in the continuous and discrete domains, which guarantee non-creation of local extrema and respect the time direction as causal (Lindeberg 1990; Lindeberg and Fagerström 1996; Lindeberg 1997), three other types of temporal scale-space approaches can be distinguished:

Continuous time and discrete scale parameter: For continuous time, it turns out that all time-causal scale-space kernels can be decomposed into convolution with primitive *truncated exponential kernels*

$$h_{prim}(t; \mu) = \frac{1}{\mu} e^{-t/\mu} \quad (t \geq 0) \quad (3)$$

having (possibly different) time constants μ . For each such primitive filter, the mean is μ and the variance μ^2 . Hence, if we couple k such filters in cascade, the equivalent convolution kernel will have a Laplace transform of the form

$$H_{composed}(s; \mu) = \int_{t=-\infty}^{\infty} (*_{i=1}^k h_{prim}(t; \mu_i)) e^{-st} dt = \prod_{i=1}^k \frac{1}{1 + \mu_i s}, \quad (4)$$

with mean (time delay) $\sum_{i=1}^k \mu_i$ and variance (effective integration time) $\sum_{i=1}^k \mu_i^2$.

Discrete time with discrete scale parameter. The discrete correspondence to the truncated exponential filters are *first-order geometric moving average filters* corresponding to the recurrence relation

$$f_{out}(t) - f_{out}(t-1) = \frac{1}{1 + \mu} (f_{in}(t) - f_{out}(t-1)). \quad (5)$$

Such a primitive filter has mean μ and variance $\mu^2 + \mu$. Coupling k such filters in cascade, gives a filter with generating function of the form

$$H_{composed}(z) = \sum_{n=-\infty}^{\infty} h_{composed}(n) z^n = \prod_{i=1}^k \frac{1}{1 - \mu_i (z-1)}, \quad (6)$$

with mean $\sum_{i=1}^k \mu_i$ and variance $\sum_{i=1}^k (\mu_i^2 + \mu_i)$. In the case of discrete time, also *time-shifted binomial kernels* satisfy temporal causality, and in this respect, discrete time allows for more degrees of freedom.

Discrete time with continuous scale parameter. The case of discrete time is special also in the sense that in this case, and only in this case, there is a non-trivial semi-group structure of scale-space kernels compatible with temporal causality. It corresponds to convolution with *Poisson kernels*

$$p(n; \lambda) = e^{-\lambda} \frac{\lambda^n}{n!} \quad (7)$$

which have mean λ , variance λ and generating function $P(z; \lambda) = e^{\lambda(z-1)}$. Intuitively, this filter can be interpreted as the limit case of repeated convolution of geometric moving average filters (6) having time constants $\mu = \lambda/m$

$$\lim_{m \rightarrow \infty} \left(H_{geom}(z; \frac{\lambda}{m}) \right)^m = \lim_{m \rightarrow \infty} \frac{1}{(1 - \frac{\lambda}{m}(z-1))^m} = P(z; \lambda). \quad (8)$$

For small values of the λ , these kernels are highly non-symmetric, whereas for large λ they approach Gaussian kernels (having the same mean and variance).

This temporal scale-space concept can be regarded as the *canonical time-causal scale-space model*, since it is the *only* time-causal scale-space concept having a semi-group structure with a *continuous time-scale parameter* and guaranteeing non-creation of local extrema with increasing scales.

Special properties of time-causal scale-spaces: A fundamental difference between the temporal scale-space concepts and the spatial multi-scale representations is that the convolution kernels are non-symmetric. Each temporal channel is associated with an inherent time delay, reflecting the fact that there is no way to access real-world data at the very present moment. Any measurement requires a finite amount of energy, and hence integration over a certain time interval. This, implies computations over non-zero time-scales, and non-zero time delays.

3 Automatic scale selection: A general principle

The presentation so far provides a theoretical framework for *representing* image data at different spatial and temporal scales. When to use it in practice, basic problems concern how to determine what structures should be regarded as significant and what scales are appropriate for handling those.

A general principle for scale selection for feature detectors defined in a spatial scale-space representation has been proposed in (Lindeberg 1993b, 1994, 1996b). It is formulated in terms of the evolution properties over scales of image descriptors expressed in terms of γ -normalized derivatives defined by

$$\partial_\xi = s^{\gamma/2} \partial_x, \quad (9)$$

where $s \in \mathbb{R}_+$ denotes the scale parameter, $\gamma > 0$ is a free parameter and ξ represents the γ -normalized coordinate of the (here, 1-D) variable x .

For an r th-order Gaussian derivative operator $g_{\xi^r}(\cdot; s)$ normalized in this way, it can be shown that the evolution over scales of its L_p -norm is given by

$$\|g_{\xi^r}(\cdot; s)\|_p = \sqrt{s}^{|r|(\gamma-1)+D(1/p-1)} \|g_{\xi^r}(\cdot; 1)\|_p. \quad (10)$$

Hence, this normalization corresponds to the L_p -norm of the equivalent normalized Gaussian derivative kernels $\partial_\xi^r g(x; s)$ being constant over scales, iff

$$p = \frac{1}{1 + \frac{|r|}{D}(1 - \gamma)}. \quad (11)$$

The basic idea of the scale selection method is that in the absence of further evidence, *the scale levels at which some (possibly non-linear) combination of such normalized derivatives assume maxima over scales can be treated as reflecting characteristic lengths of corresponding structures in the data.* As support for this approach, the following evidence can be presented (Lindeberg 1996b):

- A general theoretical analysis showing that for large classes of differential invariants, local maxima over scales of such normalized differential entities will be preserved under rescalings of the input pattern.
- Theoretical analysis of model signals for which closed-form analysis is tractable.
- Simulation results for real-world and synthetic images.

The first works on this scale selection methodology (Lindeberg 1993b, 1994) were concerned with the case $\gamma = 1$, and it was shown that for this value of γ , the scale selection methodology commutes with size variations of the input pattern. More generally, from the transformation property of these γ -normalized derivatives under a rescaling of the input $f(x) = f'(sx)$ by a factor s

$$\partial_{\xi}^r L(x; s) = s^{r(1-\gamma)} \partial_{\xi'}^r L'(x'; s'), \quad (12)$$

it is rather straightforward to show that homogeneous polynomial differential expressions scale according to a power law under size variations of the input, implying that local maxima over scales (as well as in space) will be preserved.

Conversely, given the idea that scale selection should be performed in an analogous way as image features are computed on a spatial domain — from local spatial maxima of operator responses — one may then ask how responses from operators of different size should be normalized. Indeed, it can be shown (Lindeberg 1996b) that the γ -normalized derivative concept arises *by necessity*, given the following natural assumptions:

- local maxima over scales should be preserved under rescalings of any (non-trivial) image pattern,
- the only additional source of information that could be used for normalizing the operation is the scale parameter,
- at any scale, the spatial maxima should be preserved for feature detectors expressed as homogeneous differential expressions.

Hence, the γ -normalized derivative concept spans the class of reasonable normalizations for a scale selection procedure based on local maxima over scales.

4 Dense frequency estimation based on quasi quadrature

The scale selection methodology presented in previous section has mainly been applied to the detection of *sparse image features*, such as blobs, corners, edges and ridges. In many situations, we are also interested in the computation of *dense* image descriptors, such as texture descriptors and optic flow.

An obvious problem that arises if a scale selection mechanism is to be based on a linear combination of partial derivatives, such as the Laplacian operator, is that there could be large spatial variations in the operator response. In signal processing, a common methodology for reducing this so-called phase dependency is by using *quadrature filter pairs*, defined (from a Hilbert transform) in such a way that the Euclidean sum of the filter responses will be phase independent for any sine wave. The Hilbert transform of a Gaussian derivative kernel is, however, not within the Gaussian derivative family, and we are here interested in operators of small support which can be expressed within the scale-space framework.

Given the normalized derivative concept, there is a straightforward way of combining Gaussian derivatives into an entity that gives an approximately constant operator response at the scale given by the scale selection mechanism. At

any scale t in the scale-space representation L of a one-dimensional signal f , define the following *quasi quadrature* entity in terms of normalized derivatives based on $\gamma = 1$ ¹ by

$$QL = L_\xi^2 + C L_{\xi\xi}^2 = s L_x^2 + C s^2 L_{xx}^2, \quad (13)$$

where C is a free parameter (to be determined). This approach bears close relationship to the idea by (Koenderink and van Doorn 1987) to regard derivatives of odd and even order as local sine and cosine functions. Specifically, for any sine wave $f(x) = \sin \omega_0 x$, we have

$$(QL)(x; s) = s \omega_0^2 e^{-\omega_0^2 s} (1 + (C s \omega_0^2 - 1) \sin^2 \omega_0 x). \quad (14)$$

As can be seen, the spatial variations in QL will be large when $s \omega_0^2$ is either much smaller or much larger than one, whereas the relative spatial oscillations decrease to zero when s approaches $1/(C \omega_0^2)$.

To obtain an intuitive understanding of how the choice of C affects local maxima of QL over scales, let us differentiate (14):

$$s_{QL}(x) = \frac{1}{\omega_0^2} \left(1 + \frac{2C \sin^2(\omega_0 x)}{\cos^2(\omega_0 x) + \sqrt{\cos^4(\omega_0 x) + 4C^2 \sin^4(\omega_0 x)}} \right). \quad (15)$$

¹ Since the differential expression QL is inhomogeneous, we must require $\gamma = 1$ for the scale selection procedure to commute with size variations in the input pattern.

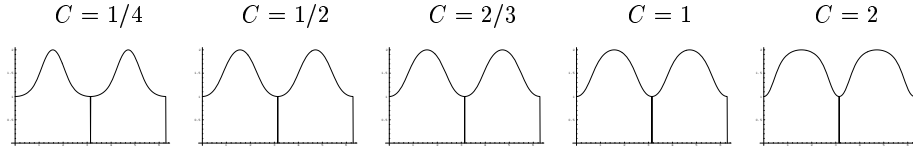


Fig. 1. Spatial variation of the *selected scale levels* when maximizing the quasi quadrature entity (13) over scales for different values of the free parameter C using a one-dimensional sine wave of unit frequency as input pattern. Observe that $C = 2/3$ (equation (16)) gives rise to the most symmetric variations in the selected scale values.

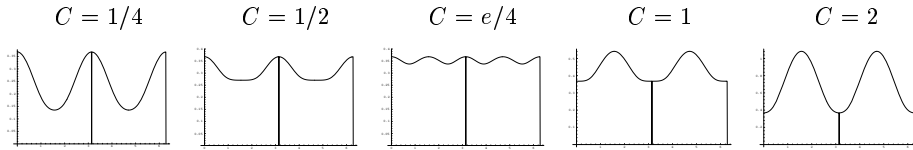


Fig. 2. Spatial variation of the *maximum value over scales* of the quasi quadrature entity (13) computed for different values of the free parameter C for a one-dimensional sine wave of unit frequency. As can be seen, the smallest spatial variations in the amplitude of the maximum response are obtained for $C = e/4$ (equation (17)).

Notably, the extreme values $s_{QL}|_{\omega_0 x=0} = \frac{1}{\omega_0^2}$ and $s_{QL}|_{\omega_0 x=\frac{\pi}{2}} = \frac{2}{\omega_0^2}$ are independent of C , and graphs showing the spatial variation for a few values of C are displayed in figure 1. Given the form of these curves, a natural symmetry requirement can be stated as

$$s_{QL}|_{\omega_0 x=\frac{\pi}{4}} = \frac{1}{2} \left(s_{QL}|_{\omega_0 x=0} + s_{QL}|_{\omega_0 x=\frac{\pi}{2}} \right) \quad \Rightarrow \quad C = \frac{2}{3} \approx 0.6667. \quad (16)$$

In this respect, $C = \frac{2}{3}$ gives the most symmetric variation of selected scales w.r.t. the information contents in the first-order and second-order derivatives.

Another interesting factor to analyse is the variation in magnitude at the selected scales. Insertion of the scale values according to (15) into the quasi quadrature measure (13) gives spatial variations of as displayed in figure 2. To determine C , a simple minimum-ripple condition is to require that

$$\underset{s = s_{\mathcal{Q},0}}{QL}|_{\omega_0 x=0} = \underset{s = s_{\mathcal{Q},\frac{\pi}{2}}}{QL}|_{\omega_0 x=\frac{\pi}{2}} \quad \Rightarrow \quad C = \frac{e}{4} \approx 0.6796. \quad (17)$$

In other words, also a determination of C based on small spatial variations in the magnitude measure computed at the selected scales gives rise to an approximately similar value of C as the abovementioned symmetry requirement.

Moreover, note that $C = e/4$ corresponds to normalizing the first- and second-order Gaussian derivative kernels to having the same L_1 -norm.

5 Normalized time-causal scale-space derivatives

The theory for automatic scale selection in section 3 applies to a scale-space on a continuous domain having a continuous scale parameter. Notably, however, the temporal scale-space concepts given by (4), (6) and (7) imply that either time or time-scale have to be discretized. Hence, a fundamental problem concerns how to normalize the derivative (or derivative approximation) responses over scales. Given the basic properties in section 3, natural constructions to consider are:

- multiply the derivative/difference operator by the variance of the smoothing kernel raised to the order of differentiation (in analogy with (9)),
- normalize the equivalent derivative/difference kernels to constant L_p -norm (or discrete l_p -norm) over scales (in analogy with (11)).

In this section, we shall describe properties of these approaches for the Poisson-type scale-space given by (7). A corresponding analysis for the two other temporal scale-space concepts is given in (Lindeberg 1996d).

To simplify the calculations, let us restrict ourselves to $\gamma = 1$, corresponding to normalization to constant L_1/l_1 -norm for any order of differentiation.

Normalized differences in the Poisson-type scale-space. For the first-order backward difference of the Poisson kernel

$$(\delta_t p)(n; \lambda) = p(t; \lambda) - p(t-1; \lambda) = -\left(\frac{t}{\lambda} - 1\right) p(t; \lambda), \quad (18)$$

we can use the fact that $p(n; \lambda)$ assumes a local maximum at $n = [\lambda]$ to compute

$$\|(\delta_t p)(\cdot; \lambda)\|_1 = \sum_{n=-\infty}^{\infty} |p(n; \lambda) - p(n-1; \lambda)| = 2p([\lambda]; \lambda) = \frac{2e^{-\lambda}\lambda^{[\lambda]}}{[\lambda]}.$$

For small λ ($\lambda < 1$), we have $\|(\delta_t p)(\cdot; \lambda)\|_1 = 2e^{-\lambda} = 2(1 - \lambda + \mathcal{O}(\lambda^2))$, whereas for large λ Stirlings formula $n! = n^n e^{-n} \sqrt{2\pi n} (1 + \frac{1}{12n} + \mathcal{O}(\frac{1}{n^2}))$ gives

$$\|(\delta_t p)(\cdot; \lambda)\|_1 = \frac{2}{\sqrt{2\pi[\lambda]}} \left(1 - \frac{(\lambda - [\lambda])^2}{2[\lambda]} - \frac{1}{12[\lambda]} + \mathcal{O}\left(\frac{1}{[\lambda]^2}\right)\right). \quad (19)$$

Concerning the second-order differences, we can use the inflexion points at $n_{1,2} = \lambda + \frac{1}{2} \pm (\lambda + \frac{1}{4})^{1/2}$ to reduce the result to first-order differences (18)

$$\|(\delta_{tt} p)(\cdot; \lambda)\|_1 = 2 \sum_{n=n_1}^{n_2} |(\delta_{tt} p)(n; \lambda)| = 2((\delta_t p)([n_1]; \lambda) - (\delta_t p)([n_2]; \lambda)).$$

Unfortunately, it is hard to simplify this expression. For small λ , however, $\|(\delta_{tt} p)(\cdot; \lambda)\|_1 = -\delta_{tt} p(1; \lambda) = (1 - \lambda)e^{-\lambda} = 1 - 2\lambda + \mathcal{O}(\lambda^2)$, and for large λ

$$\|(\delta_{tt} p)(\cdot; \lambda)\|_1 \approx \frac{4}{\sqrt{2\pi e}} \frac{1}{\lambda}. \quad (20)$$

As could be anticipated, variance-based normalization and normalization to constant l_1 -norm approach each other² with increasing temporal scales (and decreasing effects of grid sampling). For small λ , on the other hand, where the sampling effects may be strong, the results differ significantly.

6 Selection of temporal scales: Intuitive ideas

To understand the consequences of selecting local maxima over scales of normalized temporal derivatives (derivative approximations), let us first consider the response properties for a (phase shifted) sine wave, which has been used as an illustrative example in the early developments of the scale selection methodology:

$$f(n) = \cos(\nu n + \varphi) \quad (n \in \mathbf{Z}). \quad (21)$$

² Recall that for the L_1 -norms of the first- and second-order normalized Gaussian derivatives, we have $\|g_\xi(\cdot; s)\|_1 = (2/\pi)^{1/2}$ and $\|g_{\xi^2}(\cdot; s)\|_1 = (8/(\pi e))^{1/2}$.

The Fourier transform of the Poisson kernel is $\mathcal{F}(p)(\omega; \lambda) = \sum_{n=-\infty}^{\infty} p(n; \lambda) e^{-in\omega} = P(e^{-i\omega}; \lambda) = e^{-\lambda(1-\cos\omega)} e^{i\lambda\sin\omega}$ and for the r :th order backward difference

$$\mathcal{F}(\delta_-^r)(\omega) = (1 - e^{-i\omega})^r = (1 - \cos\omega + i\sin\omega)^r = (2\sin(\frac{\omega}{2}))^r e^{ir(\pi-\omega)/2}.$$

Thus, the closed-form expression for the r :th order scale-space derivative of f is

$$(\delta_-^r L)(n; \lambda) = e^{-\lambda(1-\cos\nu)} (2\sin(\frac{\nu}{2}))^r \cos(n\nu + \varphi + \lambda\sin\nu + r\frac{\pi-\nu}{2})$$

and the amplitude varies according to $\hat{L}_{tr}(\lambda) = (2\sin(\frac{\nu}{2}))^r e^{-\lambda(1-\cos\nu)}$.

Variance-based normalization. If we normalize the discrete derivative approximation operator based on the variance of the equivalent convolution kernel, this corresponds to the *normalized backward difference operator*

$$\delta_\tau^r = \delta_{-,norm}^r = \lambda^{r/2} \delta_-^r, \quad (22)$$

where τ represents the temporal coordinate normalized with respect to temporal scale. Hence, the response will first increase and then decrease with scale,

$$\hat{L}_{\tau r}(\lambda) = (2\sin(\frac{\nu}{2}))^r \lambda^{r/2} e^{-\lambda(1-\cos\nu)}, \quad (23)$$

and there is a unique maximum over scales at

$$\lambda_{\hat{L}_{\tau r},max} = \frac{r}{4\sin^2(\frac{\nu}{2})}. \quad (24)$$

Thus, in agreement with the results from the spatial domain (Lindeberg 1993b, 1994), *a local maximum over scales provides a qualitative measure of the approximate range of scales over which temporal variations occur.* If we insert the scale value (24) into (23), we see that the maximum normalized response

$$\hat{L}_{\tau r}(\lambda_{\hat{L}_{\tau r},max}) = \left(\frac{r}{e}\right)^{r/2} \quad (25)$$

is independent of ν . Thus, *all frequencies are treated in a uniform manner.* Moreover, there are strong similarities between the results from this construction and their counterparts based on the Gaussian scale-space concept. For a sine wave with frequency ν , there is a unique maximum over scales at $s = r/\nu^2$, and the maximum normalized response over scales is of the same form as (25).

This similarity is, in fact, not surprising. If we rewrite the Fourier transform of the Poisson kernel as $\mathcal{F}(p)(\omega; \lambda) = \exp(-\lambda(2\sin(\frac{\omega}{2}))^2/2) \exp(i\lambda\sin\omega)$, we see that the magnitude of $\mathcal{F}(p)(\omega; \lambda)$ is equal to the magnitude of the Fourier transform of the Gaussian kernel $\mathcal{F}(g)(\omega; s) = \exp(-s\omega^2/2)$ with ω replaced by $2\sin(\omega/2)$. Thus, disregarding the phase information, these kernels can be (formally) mapped to each other by a simple frequency warping.

Normalization to constant l_1 -norm. To study the effect of normalization to constant l_1 -norm, divide (22) by the l_1 -norm of the first-order difference of the Poisson kernel (19). This gives a normalized response of the form

$$\hat{L}_{\tau^r, discrete}(\lambda) = \sin\left(\frac{\nu}{2}\right) [\lambda]! \frac{e^{\lambda \cos \nu}}{\lambda^{|\lambda|}}. \quad (26)$$

Clearly, this function decreases with λ when $\cos \nu < 1$, showing that a single frequency with $|\nu| \geq \pi/2$ cannot give rise to a local maximum with $\lambda > 0$. If we insert this frequency into the continuous expression (24), we obtain a rule of thumb saying that if we would like to compute a derivative of order r , the scale level should preferably not be lower than $\lambda_{min}(r) = r/2$.

When $[\lambda] = 0$, *i.e.*, when $\lambda \in [0, 1[$, the normalized response increases with λ if $|\nu| < \pi/2$ (and decreases otherwise). If $[\lambda] = 1$, *i.e.*, when $\lambda \in [1, 2[$, there is a local maximum at $\lambda = 1/\cos \nu$ in this interval if $\cos \nu \in [\frac{1}{2}, 1[$. Similarly, if $[\lambda] = 2$, *i.e.*, when $\lambda \in [2, 3[$, there is a local maximum at $\lambda = 2/(\cos \nu)$ in this interval if $\cos \nu \in [\frac{2}{3}, 1[$. This pattern shows how in the case of l_1 -normalization, a single frequency may give rise to multiple responses over scales.

Cascade-coupled first-order integrators. If we couple k truncated exponential filters having equal time constants $\mu_i = \mu$ in cascade, and define a variance-based *normalized derivative operator* by

$$\partial_{\tau}^r = (k \mu^2)^{r/2} \partial_t^r, \quad (27)$$

then the maximum over scales will be assumed in layer

$$k_{\hat{L}_{\tau^r, max}} \approx \frac{r}{\log(1 + \mu^2 \nu^2)} = \frac{r}{\mu^2 \nu^2} \left(1 + \mathcal{O}\left(\frac{1}{\mu^2 \nu^2}\right)\right) \quad (28)$$

corresponding to variance $\lambda_{\hat{L}_{\tau^r, max}} = k_{\hat{L}_{\tau^r, max}} \mu^2$ of the equivalent convolution kernel. The maximum normalized response is

$$\hat{L}_{\tau^r}(\mu, k_{\hat{L}_{\tau^r, max}}) = \left(\frac{r}{e}\right)^{r/2} \left(\frac{\mu^2 \nu^2}{\log(1 + \mu^2 \nu^2)}\right)^{r/2}. \quad (29)$$

Cascade-coupled first-order recursive filters. If we in an analogous way couple k geometric moving average filters having equal time constants $\mu_i = \mu$ in cascade, and define a variance-based normalized difference operator by

$$\delta_{-, norm}^r = (k(\mu^2 + \mu))^{r/2} \delta_-^r, \quad (30)$$

the maximum over scales will be assumed in layer

$$k_{\hat{L}_{\tau^r, max}} \approx \frac{r}{\log(1 + 4(\mu^2 + \mu) \sin^2(\frac{\nu}{2}))} \quad (31)$$

corresponding to variance $\lambda_{\hat{L}_{\tau^r, max}} = k_{\hat{L}_{\tau^r, max}} (\mu^2 + \mu)$ of the equivalent convolution kernel. The maximum normalized amplitude over scales is

$$\hat{L}_{\tau^r}(\mu, k_{\hat{L}_{\tau^r, max}}) = \left(\frac{r}{e}\right)^{r/2} \left(\frac{4(\mu^2 + \mu) \sin^2(\frac{\nu}{2})}{\log(1 + 4(\mu^2 + \mu) \sin^2(\frac{\nu}{2}))}\right)^{r/2}. \quad (32)$$

Thus, local maxima over scales of these normalized derivative operators reflect similar properties as for the Poisson-type scale-space.

Scale invariance properties. Concerning the behaviour of the scale-selection method under size variations of the input pattern, we cannot, of course, aim at perfect scale invariance for the temporal scale-space concepts defined on discrete grids. For the temporal scale-space for a continuous domain (4), on the other hand, it can be shown that perfect scale invariance can be accomplished if we allow the time constants in the discrete set of filters to be variable.

7 Response properties for basic model patterns

To carry out closed-form analysis for signals having richer frequency contents, consider the behaviour in the Poisson-type scale-space of the following signals

$$f_{blob}(n) = p(n; \lambda_0), \quad (33)$$

$$f_{edge}(n) = \Psi(n; \lambda_0) = \sum_{i=-\infty}^n p(i; \lambda_0). \quad (34)$$

f_{blob} can be interpreted as idealized models of a time pulse, while f_{edge} models the edge of a new object that enters or leaves the visual field. The temporal extent of the pulse and the diffuseness of the edge are determined by λ_0 .

Descriptors based on the Poisson-type scale-space. From the semi-group property $p(\cdot; \lambda_1) * p(\cdot; \lambda_2) = p(\cdot; \lambda_1 + \lambda_2)$, it follows that the Poisson-type scale-space representations of these signals are given by

$$L_{blob}(n; \lambda) = p(n; \lambda + \lambda_0) \quad (35)$$

$$L_{edge}(n; \lambda) = \Psi(n; \lambda + \lambda_0), \quad (36)$$

and the first- and second-order normalized derivative approximations are

$$(\delta_\tau L_{blob})(n; \lambda) = \sqrt{\lambda} (\delta_t p)(n; \lambda + \lambda_0), \quad (37)$$

$$(\delta_{\tau\tau} L_{blob})(n; \lambda) = \lambda (\delta_{tt} p)(n; \lambda + \lambda_0), \quad (38)$$

$$(\delta_\tau L_{edge})(n; \lambda) = \sqrt{\lambda} p(n; \lambda + \lambda_0), \quad (39)$$

$$(\delta_{\tau\tau} L_{edge})(n; \lambda) = \lambda (\delta_t p)(n; \lambda + \lambda_0), \quad (40)$$

with $\delta_t p$ according to (18) and the second-order differences $\delta_{tt} p$ defined by

$$(\delta_{tt} p)(n; \lambda) = (\delta_i^2 p)(n; \lambda) = \left(\frac{n(n-1)}{\lambda^2} - \frac{2n}{\lambda} + 1 \right) p(t; \lambda). \quad (41)$$

Figure 3–4 show $L_{blob}(n; \lambda)$ and $L_{edge}(n; \lambda)$ with their first- and second-order normalized differences as grey-level functions of n and λ . Figure 5 shows corresponding results for a quadrature measure defined by

$$QL = L_\tau^2 + C L_{\tau\tau}^2 = \lambda (\delta_t L)^2 + C \lambda^2 (\delta_{tt} L)^2, \quad (42)$$

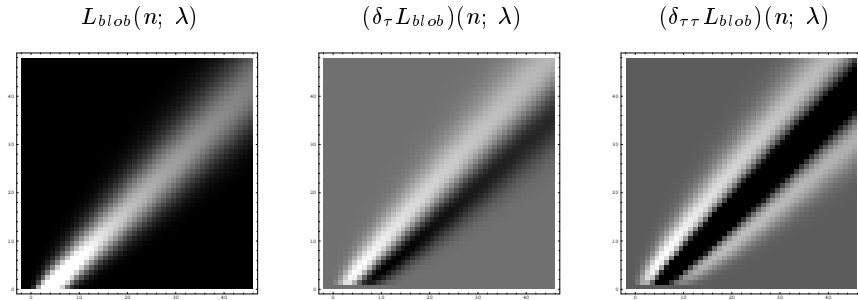


Fig. 3. The zero-, first- and second-order normalized responses in the Poisson-type scale-space representation of an *idealized blob* (time pulse) with a Poisson-shaped profile with temporal extent given by $\lambda_0 = 4$. Observe the characteristic increase in the time delay with increasing values of the scale parameter. Moreover, note that the scale at which the maximum over scales is assumed *increases with time and then decreases*. (Horizontal axis: time, vertical axis: temporal scale.)

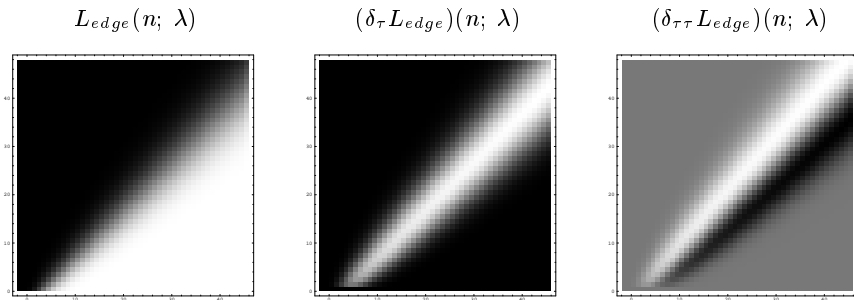


Fig. 4. The zero-, first- and second-order normalized responses in the Poisson-type scale-space representation of an *idealized edge* (time ramp) with intensity profile corresponding to an integrated Poisson kernel with diffuseness $\lambda_0 = 4$. Observe the characteristic increase in the time delay with increasing values of the scale parameter. Moreover, note that the scale at which the maximum over scales is assumed *increases monotonically with scale*. (Horizontal axis: time, vertical axis: time-scale.)

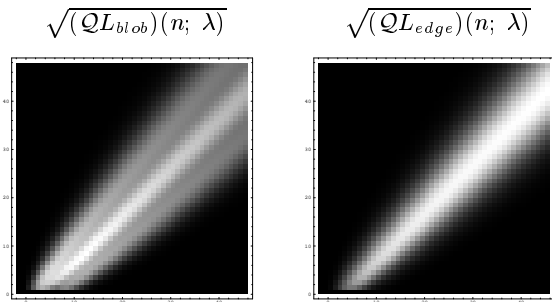


Fig. 5. The result of computing (the square root of) the *quasi quadrature measure* $QL = L_\tau^2 + C L_{\tau\tau}^2$ in the Poisson-type scale-space representations of the *idealized blob* signal in figure 3 and for the *idealized edge* signal in figure 4. Observe how this entity is less sensitive to the local phase information, while the qualitative properties of the scale selection are preserved. (Horizontal axis: time, vertical axis: time-scale.)

where L_ξ and $L_{\xi\xi}$ represent normalized discrete derivative approximations. Since for the non-symmetric backward difference operator, derivative approximation operators of different orders are associated with different time delays, the symmetric time-shifted difference operator $\delta_0 f(t) = (f(t) - f(t-2))/2$ was used for computing the first-order derivative approximation. By this discretization, the first- and second-order difference operators will have the same time delay.

Basic effects of the scale selection method. For both signals, we see that there is a characteristic increase in the time delay when the scale parameter is increased. For small t (compared to λ_0), the maximum over scales is assumed at fine scales, and this scale (as well as the maximum magnitude over scales) increase with time. These results illustrate one of the basic properties of temporal scale-space—for a new (isolated) object that enters the visual field, fine scales will be selected initially, since the object is only visible at the finest temporal scales. Then, with increasing time, this scale can be expected to increase, if we maximize a differential entity such as Q over scales.

8 Spatio-temporal patterns and motion estimation

Let us now apply simultaneous selection of spatial scales and temporal scales to the computation of spatio-temporal derivatives and analyse its implications with respect to the analysis of motion data.

8.1 Motion of a one-dimensional sine wave

Consider first a one-dimensional sine wave moving with velocity c

$$f(x, t) = \cos(\nu(x - ct)), \quad (43)$$

and define the separable spatio-temporal scale-space representation of this signal as the tensor product of the spatial and temporal scale-space concepts. Moreover, to obtain compact closed-form expressions for the results (with the discretization aspects suppressed), let us first model the temporal scale-space representation as the convolution with Gaussian kernels having the same mean and variance

$$g(x; \lambda) = \frac{1}{\sqrt{2\pi\lambda}} e^{-(x-\lambda)^2/(2\lambda)}. \quad (44)$$

Whereas this family does not satisfy temporal causality, it constitutes a reasonable approximation of the behaviour at coarse temporal scales, and we obtain a (two-parameter) spatio-temporal scale-space representation L of f of the form

$$L(x, t; s, \lambda) = e^{-\nu^2 s/2} e^{-\nu^2 c^2 \lambda/2} \cos(\nu(x - ct + \lambda)). \quad (45)$$

If we independently select spatial and temporal scale levels from the maxima over scales of the amplitude of the first-order normalized spatial and temporal

derivatives respectively, it follows that the selected spatial scale level will be $s = \frac{1}{\nu^2}$ and the temporal scale level $\lambda = 1/(c^2\nu^2)$. In other words, *the ratio between the selected spatial and temporal scales reflects the velocity*

$$c = \sqrt{\frac{s}{\lambda}}. \quad (46)$$

This behaviour bears close relationships to frequency-based motion approaches (Adelson and Bergen 1985; Heeger 1988; Fleet 1992). (A fundamental difference, however, is that there is no need for specifying frequencies manually, and the velocity is obtained fully automatically.) Moreover, from straightforward differentiation (assuming that it $L_x \neq 0$), it follows that the velocity estimate is

$$\hat{c} = -\frac{L_t}{L_x} = c = \sqrt{\frac{s}{\lambda}}, \quad (47)$$

i.e., for this ideal and noise free signal the resulting velocity estimate is independent of the scale parameters s and λ .

To formulate an algorithm that allows for simultaneous determination of spatial and temporal scales adapted to the scale levels at which dominant variations occur in the signal, a straightforward approach is to extend the definition of the quadrature measure (13) to the spatio-temporal domain

$$\max_{s,\lambda} (\mathcal{Q}_{prod(x,t)} L)(x, t; s, \lambda) = \max_{s,\lambda} (L_\xi^2 + C L_{\xi\xi}^2) (L_\tau^2 + C L_{\tau\tau}^2). \quad (48)$$

and to maximize the entity over spatial as well as and temporal scales.

Figure 6 shows the result of computing this entity at different spatio-temporal scales for an arbitrary point (x, t) in space-time and for a few combinations of ν and c . (Here, the scale parameters have been parameterized by the effective spatial scale $s_{\text{eff}} = \log_2 s$ and the effective temporal scale $\lambda_{\text{eff}} = \log_2 \lambda/2$.) Observe how we in this way obtain information about the dominant spatial and temporal frequencies around (x, t) . Specifically, the ratio between these scale levels serves as a direct estimate of the dominant motion of the first- and second-order image structures at this spatio-temporal scale.

Figure 7 shows corresponding results for two superimposed sine waves, of different spatial frequencies ν_i , which move with different velocities c_i :

$$L(x, t; s, \lambda) = e^{-\nu_1^2 s/2} e^{-\nu_1^2 c_1^2 \lambda/2} \cos(\nu_1(x - c_1 t + \lambda)) \quad (49)$$

$$+ e^{-\nu_2^2 s/2} e^{-\nu_2^2 c_2^2 \lambda/2} \cos(\nu_2(x - c_2 t + \lambda)). \quad (50)$$

Note how multiple responses over scales are obtained, indicating the ability of this approach to capture multiple transparent motions.

8.2 Velocity estimation

A more traditional approach for computing velocity estimates is in terms of spatio-temporal derivatives or other filter outputs. Here, it will be illustrated how the proposed scale selection methodology can be applied to such problems.

Consider the motion constraint equation (Horn and Schunck 1981) in the case of a one-dimensional spatial domain, and differentiate this relation once:

$$\begin{cases} L_x + L_t = 0, \\ cL_{xx} + L_{xt} = 0. \end{cases} \quad (51)$$

In terms of normalized spatio-temporal derivatives, we can write:

$$\begin{cases} L_\xi + \sqrt{\frac{s}{\lambda}} L_\tau = 0, \\ cL_{\xi\xi} + \sqrt{\frac{s}{\lambda}} L_{\xi\tau} = 0. \end{cases} \quad (52)$$

Ideally, these equations should lead to the same solution. On real-world data, however, such consistency can hardly be expected. Therefore, let us solve them in a least-squares sense using weights as imposed by the normalized derivatives:

$$\min_c (cL_\xi + \sqrt{\frac{s}{\lambda}} L_\tau)^2 + C (cL_{\xi\xi} + \sqrt{\frac{s}{\lambda}} L_{\xi\tau})^2 = \min_c \alpha c^2 + 2\beta c + \gamma. \quad (53)$$

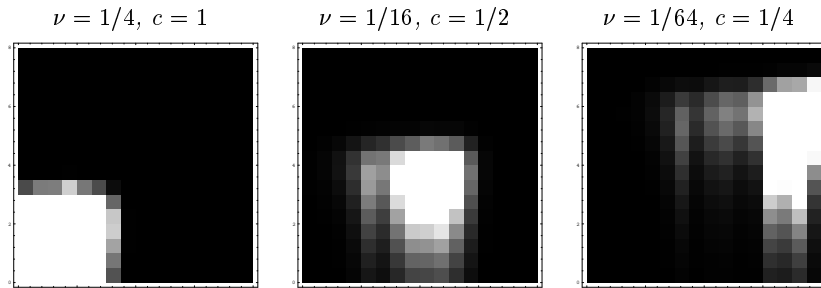


Fig. 6. $\lambda_{\text{eff}}-s_{\text{eff}}$ -diagram showing the normalized spatio-temporal quasi quadrature measure $\mathcal{P}_{\text{prod}(x,t)}L$ as function of effective temporal scale $\lambda_{\text{eff}} = \log_2 \lambda/2$ (horizontal axis) and effective spatial scale $s_{\text{eff}} = \log_2 s/2$ (vertical axis). Observe how the dominant peak at $(\lambda, s) \approx 1/\nu^2(1/c^2, 1)$, serves as an indicator of the dominant spatial frequency ν as well as the velocity c .

$$\begin{aligned} \nu_1 &= 1/64, c_1 = 1. \\ \nu_2 &= 1/4, c_2 = 1/4. \end{aligned}$$

$$\begin{aligned} \nu_1 &= 1/64, c_1 = 1/4. \\ \nu_2 &= 1/16, c_2 = 1. \end{aligned}$$

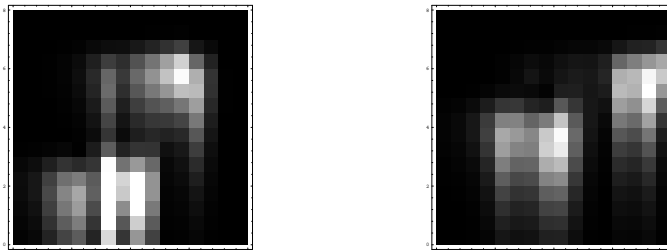


Fig. 7. Corresponding $\lambda_{\text{eff}}-s_{\text{eff}}$ -diagram of $\mathcal{P}_{\text{prod}(x,t)}L$ for two superimposed sine waves of frequencies ν_1 and ν_2 which move with velocities c_1 and c_2 , respectively. Notably, such transparent motion gives rise to multiple responses in spatio-temporal scale-space provided that the difference in frequency and velocity is sufficiently large.

Differentiation with respect to c shows that the velocity estimate is given by

$$c = -\frac{\beta}{\alpha} = -\sqrt{\frac{\sigma}{\lambda}} \left(\frac{L_{\xi}L_{\tau} + L_{\xi\xi}L_{\xi\tau}}{L_{\xi}^2 + L_{\xi\xi}^2} \right), \quad (54)$$

where the differential expression within parentheses gives the velocity estimates in units of current spatio-temporal scale. Insertion of this value into (53) gives a residual of the form $\rho = (\alpha\gamma - \beta^2)/\alpha$, and a *normalized residual*

$$\rho_{norm} = \frac{\rho}{\alpha c^2} = \frac{\alpha\gamma - \beta^2}{\alpha^2 c^2} = \frac{\alpha\gamma}{\beta^2} - 1. \quad (55)$$

In (Lindeberg 1996c), a scale selection methodology for stereo matching and flow estimation was presented based on the *region-based* flow estimation scheme by (Lukas and Kanade 1981; Bergen *et al.* 1992). The basic idea was to extend the least-squares methodology by (Lukas and Kanade 1981; Bergen *et al.* 1992) for computing the velocity estimates to *the minimization of a normalized residual over scales*. Given the least-squares formulation in (53), we propose to express a corresponding scale selection methodology for a *point-based* flow estimation scheme, by minimizing the normalized residual (55) over scales.

Figure 8 show three examples of computing flow estimates with automatic spatio-temporal scale selection in this way, from on a discrete separable temporal scale-space representation based on the time-causal temporal scale-space model in (Lindeberg and Fagerström 1996; Lindeberg 1997). To allow for the handling of two-dimensional image data, and to reduce the influence of the aperture problem, the least-squares formulation in (53) was extended in the following ways:

- from a one-dimensional to a two-dimensional spatial domain,
- from spatial derivatives up to order two to derivatives up to order four,
- the summation extended to nearest-neighbours (over a 3×3 -neighbourhood) to enforce spatial consistency of local motions at the pixel level.

Finally, at each point, the global minimum of a corresponding normalized residual has been selected, and the flow estimate computed at that point. Whereas a more general scale selection methodology should also include a mechanism for explicit handling of multiple minima over scales as well as a mechanism for velocity adaptation (Lindeberg 1997), we can nevertheless notice that the following effects *arise as consequences* of the scale selection methodology:

- selection of larger scales with increasing size of image structures (compare the results in column 1 and column 2),
- selection of finer spatial scales near flow discontinuities (see column 2).

This algorithm was tested on the image sequences used by (Barron *et al.* 1994) and gave results which by visual inspection looked at least comparable to those reported in the evaluation. (One example is shown in third column.) A more detailed evaluation on calibrated reference data under variations of the size of image structures, the velocity and the noise level of superimposed Gaussian noise

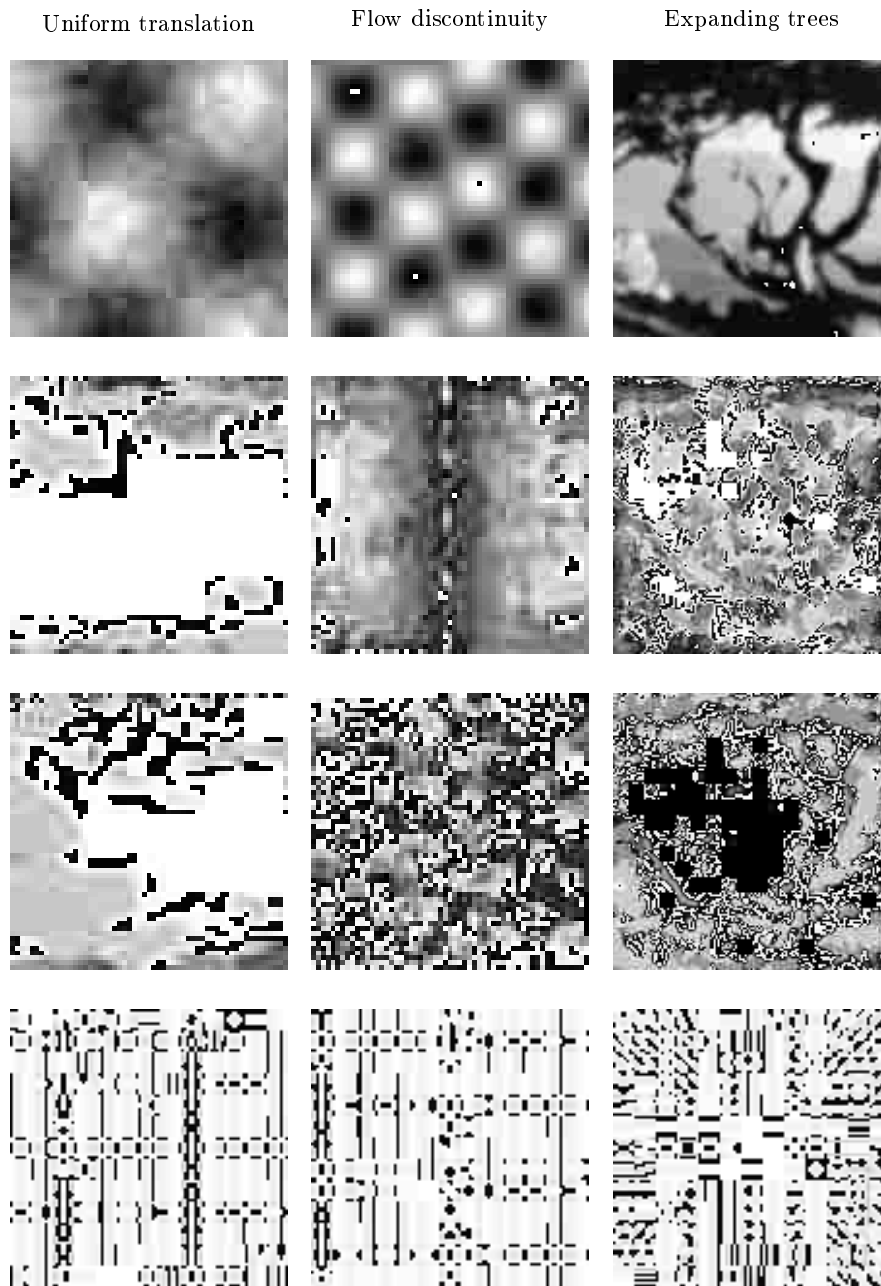


Fig. 8. Results of flow estimation with automatic scale selection for three image sequences: (left) noisy sine wave with uniform velocity, (middle) sine wave pattern where the left half moves downwards and the right half moves upwards, (right) expanding trees. The figures show from top to bottom: (to) a grey-level image from the sequence, (top middle) selected spatial scales, (bottom middle) selected temporal scales, (bottom) flow estimate. In these results, the following qualitative effects of the scale selection methodology can be seen: (i) large image structures lead to the selection of coarser scales than small image structures (compare left and right column), (ii) finer scales will be selected when we approach a flow discontinuity (see middle column).

showed that the ratio between the error in the velocity estimates computed at the spatio-temporal scales given by this scale selection criterion and corresponding velocity estimates computed at the best spatio-temporal scales (as defined from comparisons with reference data) was typically within a factor of two.

9 Relations to previous work

The motion literature is large and it is impossible to give a fair review here. (See (Barron *et al.* 1994; Beauchemin and Barron 1995) for more extensive overviews.) As pointed out earlier, the velocity detector resulting from this scale selection methodology bears close relationships to frequency based motion approaches such as (Adelson and Bergen 1985; Heeger 1988; Fleet 1992). There is also a close relationship to Gabor-based approaches for estimating motion energy, such as the work by (Grzywacz and Yuille 1990). Compared to this type of motion detection, least-squares based motion approaches such as (Lukas and Kanade 1981; Bergen *et al.* 1992) and tensor filtering approaches such as (Bigün *et al.* 1991) allow motion estimates to be computed using a smaller set of spatio-temporal filters. A motion energy scheme, such as this one, on the other hand, is able to handle situations with multiple transparent motions, and is not restricted to computing the average motion direction.

(Jones and Malik 1992) performed stereo matching based on the responses of Gaussian derivative filters at different scales and of different orientations. Motion constraint equations involving derivatives of higher order have been studied by (Werkhoven and Koenderink 1990; Arnspang 1991). (Florack and Nielsen 1994) have analysed the effect of differentiating motion constraints equations to higher order and shown how this information relates to higher-order flow fields. The least squares formulation in (53) involving Gaussian derivatives of multiple orders in the motion constraint equation can be seen as a combination of these ideas with spatial integration approach in (Lukas and Kanade 1981; Bergen *et al.* 1992) for restricting parameterized motion models of low order by overdetermined systems of equations.

Minimizing a measure of uncertainty over scales bears relationship to the statistical methodology in (Kanade and Okutomi 1994) for adapting the window size for stereo correlation. The closest relation to previous works, however, is that it is shown how the general scale selection methodology proposed in (Lindeberg 1993b, 1994) can be applied to time-causal scale-space concepts over temporal domains, and be integrated with the region-based scale selection principle for stereo matching and flow estimation in (Lindeberg 1996c) to express mechanisms for motion detection and flow estimation based small-support point operations from the N -jet (Koenderink and van Doorn 1987).

10 Summary and discussion

The subject of scale selection is an essential complement to traditional scale-space representation concerning many computer vision problems. The main sub-

ject of this paper has been to outline a foundation for expressing mechanisms for automatic selection of temporal scales in multi-scale representations based on time-causal image operations. At a more technical level, the following contributions have been presented.

- It has described how the extension of the general scale selection methodology in (Lindeberg 1993b) to γ -normalized derivatives (Lindeberg 1996a) corresponds to *normalization in L_p -norm*, and how the γ -normalized derivative concept arises by *necessity* given natural commutative properties of a scale selection methodology under size variations (section 3).
- It has been shown how the concept of quasi quadrature can be defined to allow for *dense scale selection*, and the relative weighting of first- and second-order derivative information has been analysed (section 4).
- It has been shown how *time-causal normalized derivatives* can be defined for the three different types of time-causal temporal scale-space concepts that guarantee non-creation of local extrema with increasing scale (section 5–6).
- The effect of performing local maximization of normalized derivatives over scales has been analysed for different types of model patterns, and it has been shown how the resulting scales reflect characteristic length of corresponding (spatio-)temporal image structures as well as an analysis of how the inherent time delay in a temporal multi-scale representation affects the relations between image structures at different temporal scales (section 6–7).
- It has been shown how this scale selection methodology can be used for capturing motion energy at different spatio-temporal scales, and how velocity estimates can be obtained from the spatio-temporal scales selected by the scale selection procedure (section 8.1).
- It is shown how a velocity estimation scheme can be formulated by applying the motion constraint equation to the N -jet, and how the minimization of a normalized residual over scales in a least-squares formulation provides a complementary methodology for automatic scale selection (section 8.2).

In summary, these results show how the scale selection principles previously defined on spatial domains carry over to temporal data, and how they can be given well-founded formulations based on the strictly time-causal operations which are necessary to handle real-time image data.

Acknowledgments

I would like to thank Daniel Fagerström for his kind help with the experiments underlying figure 8 as well as Kostas Daniliides for valuable discussions.

The support from the Swedish Research Council for Engineering Sciences, TFR, is gratefully acknowledged.

References

- [Adelson and Bergen, 1985] E.H. Adelson and J.R. Bergen. “Spatiotemporal energy models for the perception of motion”. *JOSA, A* 2:284–299, 1985.

- [Arnspang, 1991] J. Arnspang. *Motion Constraint Equations in Vision Calculus*. Doctoral dissertation. Doctoral dissertation, Dept. Med. Phys. Physics, Univ. Utrecht, NL-3508 Utrecht, Netherlands, 1991.
- [Barron *et al.*, 1994] J. J. Barron; D. J. Fleet, and S. S. Beauchemin. "Performance of Optical Flow Techniques". *IJCV*, 12(1), 1994.
- [Beauchemin and Barron, 1995] S.S. Beauchemin and J.L. Barron. "The computation of optical flow". *ACM Computing Surveys*, 27:433–467, 1995.
- [Bergen *et al.*, 1992] J. R. Bergen; P. Anandan; K. J. Hanna, and R. Hingorani. "Hierarchical Model-Based Motion Estimation". In *2nd ECCV*, 237–252, 1992.
- [Bigün *et al.*, 1991] J. Bigün; G. H. Granlund, and J. Wiklund. "Multidimensional orientation estimation with applications to texture analysis and optical flow". *IEEE-PAMI*, 13(8):775–790, 1991.
- [Fleet and Langley, 1995] D. J. Fleet and K. Langley. "Recursive Filters for Optical Flow". *IEEE-PAMI*, 17(1):61–67, 1995.
- [Fleet, 1992] D. Fleet. *Measurement of Image Velocity*. Kluwer, 1992.
- [Florack and Nielsen, 1994] L. Florack and M. Nielsen. "The Intrinsic Structure of the Optic Flow Field". Technical Report ERCIM-07/94-R033, 1994.
- [Florack, 1993] L. M. J. Florack. *The Syntactical Structure of Scalar Images*. PhD thesis. , Dept. Med. Physics, Utrecht Univ., NL-3508 Utrecht, Netherlands, 1993.
- [Grzywacz and Yuille, 1990] N. Grzywacz and A.L. Yuille. "A model for the estimate of local image velocity by cells in the visual cortex". *Proc. Royal Soc. London*, B 239:129–161, 1990.
- [Heeger, 1988] D. Heeger. "Optical flow using spatiotemporal filters". *IJCV*, 1:279–302, 1988.
- [Horn and Schunck, 1981] B. K. P. Horn and B. G. Schunck. "Determining Optical Flow". *AI*, 17:185–204, 1981.
- [Jones and Malik, 1992] D. G. Jones and J. Malik. "A computational framework for determining stereo correspondences from a set of linear spatial filters". In *2nd ECCV*, 395–410, 1992.
- [Kanade and Okutomi, 1994] T. Kanade and M. Okutomi. "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment". *IEEE-PAMI*, 16(9):920–932, 1994.
- [Koenderink and van Doorn, 1987] J. J. Koenderink and A. J. van Doorn. "Representation of Local Geometry in the Visual System". *Biol. Cyb.*, 55:367–375, 1987.
- [Koenderink and van Doorn, 1992] J. J. Koenderink and A. J. van Doorn. "Generic neighborhood operators". *IEEE-PAMI*, 14(6):597–605, 1992.
- [Koenderink, 1984] J. J. Koenderink. "The structure of images". *Biol. Cyb.*, 50:363–370, 1984.
- [Koenderink, 1988] J. J. Koenderink. "Scale-Time". *Biol. Cyb.*, 58:159–162, 1988.
- [Lindeberg and Fagerström, 1996] T. Lindeberg and D. Fagerström. "Scale-Space with causal time direction". In *4th ECCV*, volume 1064, 229–240, Cambridge, UK, 1996.
- [Lindeberg, 1990] T. Lindeberg. "Scale-Space for Discrete Signals". *IEEE-PAMI*, 12(3):234–254, 1990.
- [Lindeberg, 1993a] T. Lindeberg. "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention". *IJCV*, 11(3):283–318, 1993.
- [Lindeberg, 1993b] T. Lindeberg. "On Scale Selection for Differential Operators". In *8th SCIA*, 857–866, 1993.
- [Lindeberg, 1994] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer, Netherlands, 1994.

- [Lindeberg, 1996a] T. Lindeberg. “Edge detection and ridge detection with automatic scale selection”. In *Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, 1996*, 465–470, San Francisco, California, June 1996.
- [Lindeberg, 1996b] T. Lindeberg. “Feature Detection with Automatic Scale Selection”. Technical Report ISRN KTH/NA/P--96/18--SE, KTH, Stockholm, Sweden, 1996. To appear in IJCV.
- [Lindeberg, 1996c] T. Lindeberg. “A Scale Selection Principle for Estimating Image Deformations”. Technical Report ISRN KTH/NA/P--96/16--SE, KTH, Stockholm, Sweden, 1996. To appear in IVC. Shortened version in 5th ICCV, 1995.
- [Lindeberg, 1996d] T. Lindeberg. “On Automatic Selection of Temporal Scales in Time-Causal Scale-Space”. in preparation, 1996.
- [Lindeberg, 1997] T. Lindeberg. “Linear spatio-temporal scale-space”. In *Proc. 1st Int. Conf. on Scale-Space Theory in Computer Vision*, July 1997. (To appear).
- [Lukas and Kanade, 1981] B. D. Lukas and T. Kanade. “An iterative image registration technique with an application to stereo vision”. In *Image Understanding Workshop*, 1981.
- [Pauwels *et al.*, 1995] E. J. Pauwels; P. Fiddelaers; T. Moons, and L. J. van Gool. “An extended class of scale-invariant and recursive scale-space filters”. *IEEE-PAMI*, 17(7):691–701, 1995.
- [Werkhoven and Koenderink, 1990] P. Werkhoven and J. J. Koenderink. “Extraction of motion parallax structure in the visual system”. *Biol. Cyb.*, 1990.
- [Witkin, 1983] A. P. Witkin. “Scale-space filtering”. In *8th IJCAI*, 1019–1022, 1983.
- [Yuille and Poggio, 1986] A. L. Yuille and T. A. Poggio. “Scaling Theorems for Zero-Crossings”. *IEEE-PAMI*, 8:15–25, 1986.

Technical reports on related subjects can be fetched from <http://www.nada.kth.se/~tony>.