Principles for Automatic Scale Selection

Tony Lindeberg

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

http://www.nada.kth.se/~tony
Email: tony@nada.kth.se

Abstract: An inherent property of objects in the world is that they only exist as meaningful entities over certain ranges of scale. If one aims at describing the structure of unknown real-world signals, then a multi-scale representation of data is of crucial importance. Whereas conventional scale-space theory provides a well-founded framework for dealing with image structures at different scales, this theory does not directly address the problem of how to *select* appropriate scales for further analysis. This chapter outlines a systematic methodology of how mechanisms for automatic scale selection can be formulated in the problem domains of feature detection and image matching (flow estimation), respectively.

For feature detectors expressed in terms of Gaussian derivatives, hypotheses about interesting scale levels can be generated from scales at which normalized measures of feature strength assume local maxima with respect to scale. It is shown how the notion of γ-normalized derivatives arises by necessity given the requirement that the scale selection mechanism should commute with rescalings of the image pattern. Specifically, it is worked out in detail how feature detection algorithms with automatic scale selection can be formulated for the problems of edge detection, blob detection, junction detection, ridge detection and frequency estimation. A general property of this scheme is that the selected scale levels reflect the size of the image structures.

When estimating image deformations, such as in image matching and optic flow computations, scale levels with associated deformation estimates can be selected from the scales at which normalized measures of uncertainty assume local minima with respect to scales. It is shown how an integrated scale selection and flow estimation algorithm has the qualitative properties of leading to the selection of coarser scales for larger size image structures and increasing noise level, whereas it leads to the selection of finer scales in the neighbourhood of flow field discontinuities.

Keywords: scale, scale-space, scale selection, normalized derivative, feature detection, blob detection, corner detection, frequency estimation, Gaussian derivative, stereo matching, optic flow, multi-scale representation, computer vision

1 Principles for Automatic Scale Selection

Tony Lindeberg

Computational Vision and Active Perception Laboratory (CVAP)
KTH (Royal Institute of Technology), Stockholm, Sweden

1.1 Introduction

An inherent property of objects in the world is that they only exist as meaningful entities over certain ranges of scale. If one aims at describing the structure of unknown real-world signals, then a multi-scale representation of data is of crucial importance. Whereas conventional scale-space theory provides a well-founded framework for dealing with image structures at different scales, this theory does not directly address the problem of how to *select* appropriate scales for further analysis.

This chapter outlines a systematic methodology for formulating mechanisms for automatic scale selection in the domains of feature detection and image matching.

1.2 Multi-scale differential image geometry

A natural and powerful framework for representing image data at the earliest stages of visual processing is by computing differential geometric image descriptors at multiple scales [Koenderink, 1990; Koenderink et al., 1992]. This section summarizes essential components of this *scale-space theory* [Lindeberg, 1994d], which also constitutes the vocabulary for expressing the scale selection mechanisms.

1.2.1 Scale-space representation

Given any continuous signal $g\colon R^D \to R$, its linear *scale-space representation* $L\colon R^D \times R_+ \to R$ is defined as the solution to the *diffusion equation*

$$\partial_t L = \frac{1}{2}\,\boldsymbol{\nabla}^2 L = \frac{1}{2}\sum_{d=1}^{D}\partial_{x_d x_d}L \tag{1.1}$$

with initial condition $L(\boldsymbol{x}; 0) = g(\boldsymbol{x})$. Equivalently, this family can be defined by convolution with *Gaussian kernels* $h(\boldsymbol{x}; t)$ of various width t

$$L(\boldsymbol{x}; t) = h(\boldsymbol{x}; t) * g(\boldsymbol{x}), \tag{1.2}$$

where $h\colon R^D \times R_+ \to R$ is given by

$$h(\boldsymbol{x}, t) = \frac{1}{(2\pi t)^{D/2}}\exp\left(-\frac{x_1^2 + \ldots + x_D^2}{2t}\right) \tag{1.3}$$

and $\boldsymbol{x} = [x_1, \ldots, x_D]^T$. There are several results [Koenderink, 1984], [Koenderink and van Doorn, 1987], [Koenderink and van Doorn, 1990], [Koenderink and van Doorn, 1992], [Babaud et al., 1986], [Yuille and Poggio, 1986], [Hummel and Moniot, 1989], [Lindeberg, 1990], [Lindeberg, 1994b],

[Lindeberg, 1994d], [Lindeberg, 1997a], [Florack et al., 1992], [Florack, 1993], [Florack et al., 1993], [Florack, 1997], [Pauwels et al., 1995], [Sporring et al., 1996] stating that within the class of linear transformations the Gaussian kernel is the unique kernel for generating a scale-space. The conditions that specify the uniqueness are essentially linearity and shift invariance combined with different ways of formalizing the notion that new structures should not be created in the transformation from a finer to a coarser scale (see also [Lindeberg, 1994c], [Lindeberg, 1996e], [Lindeberg and ter Haar Romeny, 1994a], [Lindeberg and ter Haar Romeny, 1994b], [Sporring et al., 1996] for reviews).

1.2.2 Gaussian derivative operators

From the scale-space representation, we can at any level of scale define *scale-space derivatives* by

$$L_{x^\alpha}(\boldsymbol{x}; t) = (\partial_{x^\alpha} L)(\boldsymbol{x}; t) = \partial_{x^\alpha}(h(\boldsymbol{x}; t) * g(\boldsymbol{x})) \qquad (1.4)$$

where $\alpha = [\alpha_1, \ldots, \alpha_D]^T$ and $\partial_{x^\alpha} L = L_{x_1^{\alpha_1} \ldots x_D^{\alpha_D}}$ constitute multi-index notation for the derivative operator ∂_{x^α}. Since differentiation commutes with convolution, the scale-space derivatives can be written

$$L_{x^\alpha}(\boldsymbol{x}; t) = (\partial_{x^\alpha} h(\cdot;\ t)) * g(\boldsymbol{x}) \qquad (1.5)$$

and correspond to convolving the original image g with Gaussian derivate kernels $\partial_{x^\alpha} h$. Figure 1.1 shows a few examples of such *Gaussian derivate operators*.

The Gaussian derivatives provide a compact way to characterize the local image structure around a certain image point at any scale. With access to Gaussian derivative responses of all orders at one image point $\boldsymbol{x}_0$, we can for any $\boldsymbol{x}$ in a neighbourhood of $\boldsymbol{x}_0$ reconstruct the original scale-space representation by a local Taylor expansion. With θ_α denoting the Taylor coefficient of order α, this reconstruction can be written

$$L(\boldsymbol{x}; t) = \sum_\alpha \theta_\alpha L_{x^\alpha}(\boldsymbol{x}_0;\ t)(\boldsymbol{x} - \boldsymbol{x}_0)^\alpha. \qquad (1.6)$$

Truncating this representation to derivatives up to order N, results in a so-called *N-jet representation* [Koenderink and van Doorn, 1987].

1.2.3 Directional derivatives and linear combinations of Gaussian derivatives

The Gaussian derivatives according to (1.4) correspond to partial derivatives along the Cartesian coordinate directions. Using the well-known expression for the nth-order directional derivative $\partial_{\bar\beta}^n$ of a function L in any
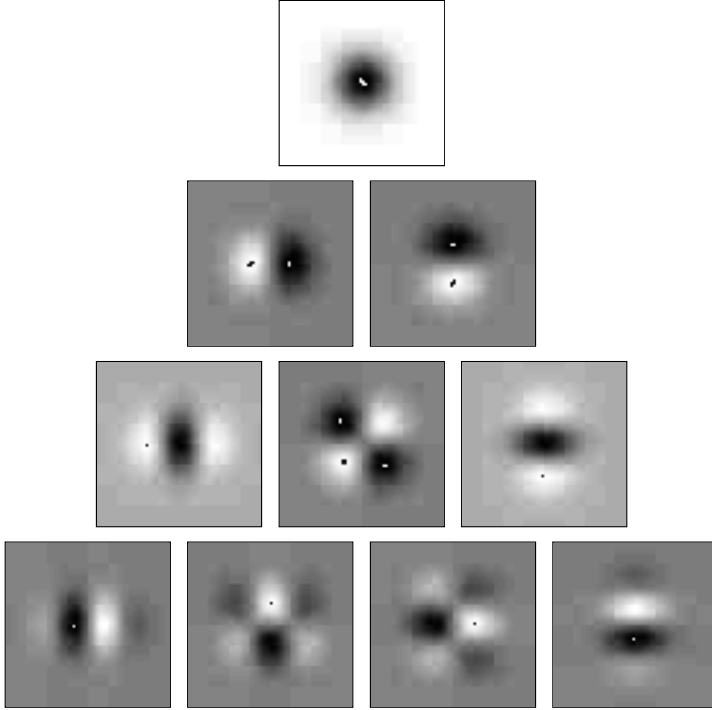
Figure 1.1: *Gaussian derivative kernels up to order three in the two-dimensional case.*

direction β,

$$\partial_{\bar{\beta}}^n L = (\cos\beta\,\partial_x + \sin\beta\,\partial_y)^n L. \qquad (1.7)$$

we can express a *directional derivative* in any direction $(\cos\beta, \sin\beta)$ as a linear combination of the Cartesian Gaussian derivatives. (This property is sometimes referred to as "steerability" [Freeman and Adelson, 1990],[Perona, 1992].) For orders up to three, the explicit expressions are

$$\partial_{\bar{\beta}} L = \cos\beta L_x + \sin\beta L_y, \qquad (1.8)$$

$$\partial_{\bar{\beta}}^2 L = \cos^2\beta L_{xx} + 2\cos\beta\sin\beta L_{xy} + \sin^2\beta L_{yy} \qquad (1.9)$$

$$\partial_{\bar{\beta}}^3 L = \cos^3\beta L_{xxx} + 3\cos^2\beta\sin\beta L_{xxy} + 3\cos\beta\sin^2\beta L_{xyy} + \sin^3\beta L_{yyy}. \qquad (1.10)$$

Figure 1.2 shows an example of computing first- and second-order directional derivatives in this way, based on the Gaussian derivative operators in Figure 1.1.
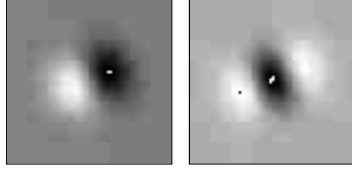
Figure 1.2: a *First- and* **b** *second-order directional derivative approximation kernels in the 22.5 degree direction computed as a linear combination of the Gaussian derivative operators.*

More generally, and with reference to (1.6), it is worth noting that the Gaussian derivatives at any scale (including the zero order derivative) serve as a complete linear basis, implying that any linear filter can be expressed as a (possibly infinite) linear combination of Gaussian derivatives.

1.2.4 Differential invariants

A problem with image descriptors as defined from (1.4) and (1.7) is that they depend upon the orientation of the coordinate system. A simple way to define image descriptors that are invariant to rotations in the image plane, is by considering directional derivatives in a preferred coordinate system aligned to the local image structure.

Figure 1.3: *In the (u, v) coordinate system, the v-direction is at every point parallel to the gradient direction of L, and the u-direction is parallel to the tangent to the level curve.*

One such choice of preferred directions is to introduce a local orthonormal coordinate system (u, v) at any point P_0, with the v-axis parallel to the gradient direction at P_0, and the u-axis perpendicular, i.e. $e_v = (\cos\varphi, \sin\varphi)^T$ and $e_u = (\sin\varphi, -\cos\varphi)^T$, where

$$e_v|_{P_0} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} = \frac{1}{\sqrt{L_x^2 + L_y^2}} \begin{pmatrix} L_x \\ L_y \end{pmatrix}\Bigg|_{P_0}. \qquad (1.11)$$

In terms of Cartesian coordinates, the corresponding local directional derivative operators can then be written

$$\partial_u = \sin\varphi\,\partial_x - \cos\varphi\,\partial_y. \quad \partial_v = \cos\varphi\,\partial_x + \sin\varphi\,\partial_y, \tag{1.12}$$

and for the lowest orders of differentiation we have

$$L_u = 0$$
$$L_v = \sqrt{L_x^2 + L_y^2} \tag{1.13}$$

$$L_v^2 L_{uu} = L_{xx}L_y^2 - 2L_{xy}L_xL_y + L_{yy}L_x^2$$
$$L_v^2 L_{uv} = L_xL_y(L_{xx} - L_{yy}) - (L_x^2 - L_y^2)L_{xy}$$
$$L_v^2 L_{vv} = L_x^2 L_{xx} + 2L_xL_yL_{xy} + L_y^2 L_{yy} \tag{1.14}$$

$$L_v^3 L_{uuu} = L_y(L_y^2 L_{xxx} + 3L_x^2 L_{xyy})$$
$$\qquad\qquad - L_x(L_x^2 L_{yyy} + 3L_y^2 L_{xxy})$$
$$L_v^3 L_{uuv} = L_x(L_y^2 L_{xxx} + (L_x^2 - 2L_y^2)L_{xyy})$$
$$\qquad\qquad + L_y(L_x^2 L_{yyy} + (L_y^2 - 2L_x^2)L_{xxy})$$
$$L_v^3 L_{uvv} = L_y(L_x^2 L_{xxx} + (L_y^2 - 2L_x^2)L_{xyy})$$
$$\qquad\qquad + L_x((2L_y^2 - L_x^2)L_{xxy} - L_y^2 L_{yyy})$$
$$L_v^3 L_{vvv} = L_x(L_x^2 L_{xxx} + 3L_y^2 L_{xyy})$$
$$\qquad\qquad + L_y(3L_x^2 L_{xxy} + L_y^2 L_{yyy}). \tag{1.15}$$

By definition, these differential definition are invariant under rotations of the image plane, and this (u, v)-coordinate system is characterized by the first-order directional derivatives L_u being zero.

Another natural choice of a preferred coordinate system is to align a local (p, q)-coordinate system to the eigenvectors of the Hessian matrix. To express directional derivatives in such coordinates, characterized by the mixed second-order derivative L_{pq} being zero, we can rotate the coordinate system by an angle ψ defined by

$$\cos\psi|_{(x_0,y_0)} = \sqrt{\frac{1}{2}\left(1 + \frac{L_{xx} - L_{yy}}{\sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}\right)}\Bigg|_{(x_0,y_0)} \tag{1.16}$$

$$\sin\psi|_{(x_0,y_0)} = (\mathrm{sgn}\,L_{xy})\sqrt{\frac{1}{2}\left(1 - \frac{L_{xx} - L_{yy}}{\sqrt{(L_{xx} - L_{yy})^2 + 4L_{xy}^2}}\right)}\Bigg|_{(x_0,y_0)}$$
$$\tag{1.17}$$

and define unit vectors in the p- and q-directions by $e_p = (\sin\psi, -\cos\psi)$ and $e_q = (\cos\psi, \sin\psi)$ with associated directional derivative operators

$$\partial_p = \sin\psi\,\partial_x - \cos\psi\,\partial_y, \qquad \partial_q = \cos\psi\,\partial_x + \sin\psi\,\partial_y. \tag{1.18}$$

Then, it is straightforward to verify that this definition implies that

$$\begin{aligned} L_{pq} = \partial_p\partial_q L &= (\cos\psi\,\partial_x + \sin\psi\,\partial_y)\,(\sin\psi\,\partial_x - \cos\psi\,\partial_y)L \\ &= \cos\psi\,\sin\psi\,(L_{xx} - L_{yy}) - (\cos^2\psi - \sin^2\psi)\,L_{xy} = 0. \end{aligned} \tag{1.19}$$

A more general class of (non-linear) differential invariants will be considered in Sect. 1.3.2.

1.2.5 Windowed spectral moment descriptors

The differential invariants defined so far depend upon the *local* differential geometry at any given *image point*. One way of defining *regional* image descriptors, which reflect the intensity distribution over *image patches*, is by considering windowed spectral moment descriptors (see also [Bigün et al., 1991; Rao and Schunk, 1991; Lindeberg, 1994d]). Using Plancherel's relation

$$\int_{\omega \in R^2} \hat{f}_1(\omega)\,\hat{f}_2^*(\omega)\,\mathrm{d}\omega = (2\pi)^2 \int_{x \in R^2} f_1(x)\,f_2^*(x)\,\mathrm{d}x, \tag{1.20}$$

where $\hat{h}_i(\omega)$ denotes the Fourier transform of $h_i(x)$ (with $\omega = 2\pi k$ as variable in the frequency domain) and by letting $f_1 = L_{x^\alpha}$ and $f_2 = L_{x^\beta}$, we have

$$\int_{\omega \in R^2} (i\omega)^{|\alpha|+|\beta|}|\hat{L}|^2(\omega)\,\mathrm{d}\omega = (2\pi)^2 \int_{x \in R^2} L_{x^\alpha}(x)\,L_{x^\beta}(x)\,\mathrm{d}x. \tag{1.21}$$

Let us next introduce a Gaussian window function $h(\cdot;\ s)$, depending on an *integration scale parameter* s, in addition to the *local scale parameter* t of the ordinary scale-space representation. Then, we can define the following *windowed spectral moments*

$$\begin{aligned} \mu_{20}(x;\ t,s) &= \int_{\xi \in R^2} L_x^2(\xi;\ t)\,h(x-\xi;\ s)\,\mathrm{d}\xi, \\ \mu_{11}(x;\ t,s) &= \int_{\xi \in R^2} L_x(\xi;\ t)\,L_y(\xi;\ t)\,h(x-\xi;\ s)\,\mathrm{d}\xi, \\ \mu_{02}(x;\ t,s) &= \int_{\xi \in R^2} L_y^2(\xi;\ t)\,h(x-\xi;\ s)\,\mathrm{d}\xi, \end{aligned} \tag{1.22}$$

and higher order spectral moment descriptors can be defined in an analogous fashion.
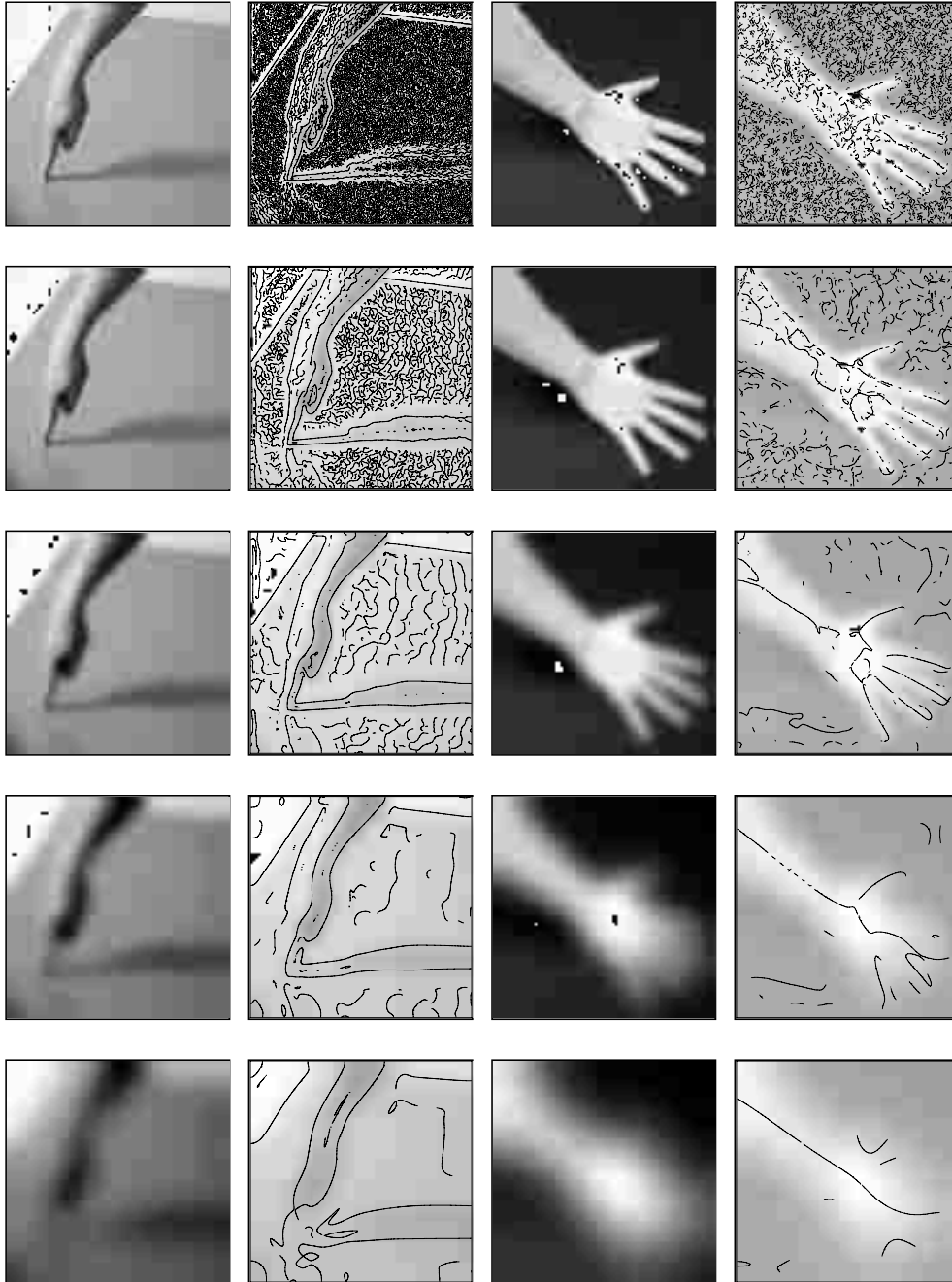
Scale-space representation   *Edges*   *Scale-space representation*   *Ridges*

Figure 1.4: *Edges and ridges computed at different scales in scale-space (scale levels t = 1.0, 4.0, 16.0, 64.0 and 256.0 from top to bottom) using a differential geometric edge detector and ridge detector, respectively. (Image size: 256*256 pixels.)*

1.2.6 The scale-space framework for a visual front-end

The image descriptors defined in Section 1.2.2–Section 1.2.5 provide a useful basis for expressing a large number of early visual operations, including image representation, feature detection, stereo matchning, optic flow and shape estimation. There is also a close connection to biological vision. Neurophysiological studies by Young [1985, 1987] have shown that there are receptive fields in the mammalian retina and visual cortex, which can be well modelled by Gaussian derivatives up to order four. In these respects, the scale-space representation with its associated Gaussian derivative operators can be seen as a canonical idealized model of a visual front-end.

When computing these descriptors at multiple scales, however, as is necessary to capture the complex multi-scale nature of our world, one can expect these descriptors to accurately reflect interesting image structures at certain scales, while the responses may be less useful at other scales. To simplify the interpretation tasks of later processing stages, a key problem of a visual front-end (in addition to making image structures more explicit by computing differential image descriptors at multiple scales) is to provide hypothesis about how to select locally appropriate scales for describing the data set.

1.2.7 The need for automatic scale selection

To illustrate the need for an explicit mechanism for automatic scale selection, let us first consider the problem of detecting edges. The left column in figure 1.4 shows the result of applying a standard edge detector (described in Section 1.4.1) to an image, which have been smoothed by convolution with Gaussian kernels of different widths.

As can be seen, different types of edge structures give rise to edge curves at different scales. For example, the shadow of the arm only appears as a connected edge curve at coarse scales. If such coarse scales are used at the finger tip, however, the shape distortions due to scale-space smoothing will be substantial. Hence, to extract this edge with a reasonable trade-off between detection and localization properties, the only reasonable choice is to *allow the scale levels to vary along the edge.*

The right column in figure 1.4 shows corresponding results for a ridge detector (described in Section 1.4.2). Different types of ridge structures give rise to qualitatively different types of ridge curves depending on the scale level. The fingers respond at $t \approx 16$, whereas the arm as a whole is extracted as a long ridge curve at $t \approx 256$.

For these reasons, and since the choice of scale levels crucially affects the performance of any feature detector, and different scale levels will,

in general, be required in different parts of the image, it is essential to *complement feature detectors by explicit mechanisms which automatically adapt the scale levels to the local image structure.*

1.3   A general scale selection principle

A powerful approach to perform local and adaptive scale selection is by detecting local extrema over scales of normalized differential entities. This chapters presents a general theory by first introducing the notion of normalized derivatives, and then showing how local extrema over scales of normalized differential entities reflect the characteristic size of corresponding image structures.

1.3.1   Normalized derivatives and intuitive idea for scale selection

A well-known property of the scale-space representation is that the amplitude of spatial derivatives

$$L_{x^\alpha}(\cdot;\ t) = \partial_{x^\alpha} L(\cdot;\ t) = \partial_{x_1^{\alpha_1}} \ldots \partial_{x_D^{\alpha_D}} L(\cdot;\ t)$$

in general *decrease with scale, i.e.,* if a signal is subject to scale-space smoothing, then the numerical values of spatial derivatives computed from the smoothed data can be expected to decrease. This is a direct consequence of the *non-enhancement property of local extrema,* which means that the value at a local maximum cannot increase, and the value at a local minimum cannot decrease [Babaud et al., 1986; Lindeberg, 1994d]. In other words, the amplitude of the variations in a signal will always decrease with scale.

As a simple example of this, consider a sinusoidal input signal of some given angular frequency $\omega_0 = 2\pi k_0$; for simplicity in one dimension,

$$g(x) = \sin \omega_0 x. \tag{1.23}$$

It is straightforward to show that the solution of the diffusion equation is given by

$$L(x;\ t) = e^{-\omega_0^2 t/2} \sin \omega_0 x. \tag{1.24}$$

Thus, the amplitude of the scale-space representation, L_{max}, as well as the amplitude of the mth order smoothed derivative, $L_{x^m,max}$, decrease exponentially with scale

$$L_{max}(t) = e^{-\omega_0^2 t/2}, \quad L_{x^m,max}(t) = \omega_0^m\, e^{-\omega_0^2 t/2}.$$

Let us next introduce a *γ-normalized derivative operator* defined by

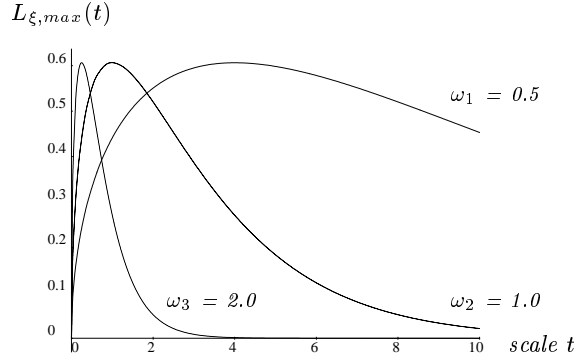$$\partial_{\xi,\gamma-norm} = t^{\gamma/2}\, \partial_x, \tag{1.25}$$

Figure 1.5: *The amplitude of first order normalized derivatives as function of scale for sinusoidal input signals of different frequency ($\omega_1 = 0.5$, $\omega_2 = 1.0$ and $\omega_3 = 2.0$).*

which corresponds to the change of variables

$$\xi = \frac{x}{t^{\gamma/2}}. \tag{1.26}$$

For the sinusoidal signal, the amplitude of an mth order normalized derivative as function of scale is given by

$$L_{\xi^m,max}(t) = t^{m\gamma/2}\,\omega_0^m\,e^{-\omega_0^2 t/2}, \tag{1.27}$$

i.e., it first increases and then decreases. Moreover, it assumes a unique maximum at $t_{max,L_{\xi^m}} = \frac{\gamma\,m}{\omega_0^2}$. If we define a scale parameter σ of dimension length by $\sigma = \sqrt{t}$ and introduce the wavelength λ_0 of the signal by $\lambda_0 = 2\pi/\omega_0$, we can see that the scale at which the amplitude of the γ-normalized derivative assumes its maximum over scales is *proportional* to the wavelength, λ_0, of the signal:

$$\sigma_{max,L_{\xi^m}} = \frac{\sqrt{\gamma\,m}}{2\pi}\lambda_0. \tag{1.28}$$

The maximum value over scales is

$$L_{\xi^m,max}(t_{max,L_{\xi^m}}) = \frac{(\gamma m)^{\gamma m/2}}{e^{\gamma m/2}}\,\omega_0^{(1-\gamma)m}. \tag{1.29}$$

In the case when $\gamma = 1$, this maximum value is *independent* of the frequency of the signal (see figure 1.5), and the situation is highly symmetric, *i.e.*, given any scale t_0, the maximally amplified frequency is given by $\omega_{max} = \sqrt{m/t_0}$, and for any ω_0 the scale with maximum amplification is $t_{max} = m/\omega_0^2$. In other words, for normalized derivatives with $\gamma = 1$ it holds that sinusoidal signals are treated in a similar (scale invariant) way independent of their frequency (see figure 1.5).

1.3.2   A general principle for automatic scale selection

The example shows that the scale at which a normalized derivative assumes its maximum over scales is for a sinusoidal signal proportional to the wavelength of the signal. In this respect, maxima over scales of normalized derivatives reflect the scales over which spatial variations take place in the signal. This property is, however, not restricted to sine wave patterns or to image measurements in terms of linear derivative operators of a certain order. Contrary, it applies to a large class of image descriptors which can be formulated as multi-scale differential invariants expressed in terms of Gaussian derivatives. In [Lindeberg, 1994d], the following scale selection principle was proposed:

> *In the absence of other evidence, assume that a scale level, at which some (possibly non-linear) combination of normalized derivatives assumes a local maximum over scales, can be treated as reflecting a characteristic length of a corresponding structure in the data.*

1.3.3   Properties of the scale selection principle

A basic justification for the abovementioned statement can be obtained from the fact that for a large class of (possibly non-linear) combinations of normalized derivatives it holds that maxima over scales have a nice behaviour under rescalings of the intensity pattern. If the input image is rescaled by a constant scaling factor s, then the scale at which the maximum is assumed will be multiplied by the same factor (if measured in units of $\sigma = \sqrt{t}$). This is a fundamental requirement on a scale selection mechanism, since it guarantees that image operations commute with size variations.

Scaling properties.   For two signals g and g' related by

$$g(x) = g'(sx), \tag{1.30}$$

the corresponding normalized derivatives defined from the scale-space representations L and L' on the two domains are related according to

$$\partial_{\xi^m} L(x;\ t) = s^{m(1-\gamma)}\, \partial_{\xi'^m} L'(x';\ t'), \tag{1.31}$$

and when $\gamma = 1$, the normalized derivatives are equal at corresponding points $(x;\ t)$ and $(x';\ t') = (sx;\ s^2 t)$.

When $\gamma \neq 1$, a weaker scale invariance properties holds. Let us consider a homogeneous polynomial differential invariant $\mathcal{D}L$ of the form

$$\mathcal{D}L = \sum_{i=1}^{I} c_i \prod_{j=1}^{J} L_{x^{\alpha_{ij}}}, \tag{1.32}$$

where the sum of the orders of differentiation in a certain term

$$\sum_{j=1}^{J} |\alpha_{ij}| = M \tag{1.33}$$

does not depend on the index i of that term. Then, normalized differential expressions in the two domains are related by

$$\mathcal{D}_{\gamma-norm}L = s^{M(1-\gamma)} \mathcal{D}'_{\gamma-norm}L', \tag{1.34}$$

i.e., magnitude measures scale according to a power law. Local maxima over scales are, however, still preserved

$$\partial_t \left(\mathcal{D}_{\gamma-norm}L\right) = 0 \qquad \Leftrightarrow \qquad \partial_{t'} \left(\mathcal{D}'_{\gamma-norm}L'\right) = 0, \tag{1.35}$$

which gives sufficient scale invariance to support the scale selection methodology. More generally, it can be shown that the notion of γ-normalized derivatives arises *by necessity*, given natural requirements of a scale selection mechanism [Lindeberg, 1996a].

1.3.4   Interpretation of normalized derivatives

L_p-**norms.**   For a D-dimensional signal, it can be shown that the variation over scales of the L_p-norm of an mth order normalized Gaussian derivative kernel is given by

$$\|h_{\xi^m}(\cdot; \; t)\|_p = \sqrt{t}^{m(\gamma-1)+D(1/p-1)} \, \|h_{\xi^m}(\cdot; \; t)\|_p. \tag{1.36}$$

In other words, the L_p-norm of the mth order Gaussian derivative kernel is constant over scales if and only if

$$p = \frac{1}{1 + \frac{m}{D}(1-\gamma)}. \tag{1.37}$$

Hence, the γ-normalized derivative concept be interpreted as an L_p-normalization of the Gaussian derivative kernels over scales for a specific value of p, which depends upon γ, the dimension as well as the order m of differentiation.

The perfectly scale invariant case $\gamma = 1$ gives $p = 1$ for all orders m and corresponds to L_1-normalization of the Gaussian derivative kernels.

Power spectra.   For a signal $g: R^2 \to R$ having a power spectrum of the form

$$S_g(\omega_1, \omega_2) = (\hat{f}\hat{f}^*)(\omega_1, \omega_2) = |\omega|^{-2\alpha} = (\omega_1^2 + \omega_2^2)^{-\alpha}, \tag{1.38}$$

it can be shown that the variation over scales of the following energy measure

$$P_{L(\cdot;\,t)} = \int_{x \in R^2} |\boldsymbol{\nabla}L(x;\,t)|^2 \, dx. \tag{1.39}$$

is given by

$$P_{\mathrm{norm}}(\cdot;\,t) = t^\gamma \, P_{L(\cdot;\,t)} \sim t^{\alpha+\gamma-2}. \tag{1.40}$$

This expression is independent of scale if and only if $\alpha = 2 - \gamma$. In other words, in the two-dimensional case the normalized derivative model is *neutral* with respect to power spectra of the form (and natural images often have power spectra of this form [Field, 1987])

$$S_g(\omega) = |\omega|^{-2(2-\gamma)}. \tag{1.41}$$

1.4 Feature detection with automatic scale selection

This section shows how the general principle for automatic scale selection described in Section 1.3.2 can be integrated with various types of feature detectors.

1.4.1 Edge detection

At any scale in scale-space, let us define an edge point as a point at which the second directional derivative L_{vv} in the v-direction is zero, and the third directional derivative L_{vvv} is negative:

$$\begin{cases} L_{vv} = 0, \\ L_{vvv} < 0, \end{cases} \tag{1.42}$$

The edges in figure 1.4 have been computed according to this definition [Canny, 1986], [Korn, 1988], [Lindeberg, 1993b], [Lindeberg, 1994d].

In view of the scale selection principle, a natural extension of this notion of *non-maximum suppression* is by defining a *scale-space edge* a curve on this *edge surface*, such that some measure of edge strength $\mathcal{E}_{\gamma-norm}L$ assumes locally maxima with respect to scale on this curve.

$$\begin{cases} \partial_t(\mathcal{E}_{\gamma-norm}L(x,y;\,t)) = 0, \\ \partial_{tt}(\mathcal{E}_{\gamma-norm}L(x,y;\,t)) < 0, \\ L_{vv}(x,y;\,t) = 0, \\ L_{vvv}(x,y;\,t) < 0. \end{cases} \tag{1.43}$$

Based on the γ-parameterized normalized derivative concept in, we shall here consider the following two edge strength measures:

$$\mathcal{G}_{\gamma-norm}L = t^\gamma \, L_v^2, \tag{1.44}$$

$$\mathcal{T}_{\gamma-norm}L = -t^{3\,\gamma} \, L_v^3 \, L_{vvv}. \tag{1.45}$$

Qualitative properties. For a *diffuse step edge*, defined as the primitive function of a one-dimensional Gaussian,

$$f_{t_0}(x, y) = \int_{x'=-\infty}^{x} h(x';\ t_0)\, dx',$$

each of these edge strength measures assumes a unique maximum over scales at $t_{\mathcal{G}_{\gamma-norm}} = t_{\mathcal{T}_{\gamma-norm}} = \frac{\gamma}{1-\gamma} t_0$. Requiring this maximum to occur at t_0 gives $\gamma = \frac{1}{2}$.

For a local model of an *edge bifurcation*, expressed as

$$L(x;\ t) = \tfrac{1}{4}x^4 + \tfrac{3}{2}x^2(t - t_b) + \tfrac{3}{4}(t - t_b)^2,$$

with edges at $x_1(t) = (t_b - t)^{1/2}$ when $t \leq t_b$, we have $(\mathcal{G}_{\gamma-norm}L)(x_1(t);\ t) = 4\, t^\gamma (t_b - t)^3$, and the selected scales are $t_{\mathcal{G}_{\gamma-norm}} = \frac{\gamma}{3+\gamma}\, t_b$ and $t_{\mathcal{T}_{\gamma-norm}} = \frac{3\gamma}{5+3\gamma}\, t_b$.

On other words, the scale selection method has the qualitative property of *reflecting the degree of diffuseness of the edge.* Moreover, since the edge strength decreases rapidly at a bifurcation, the selected scales will tend away from bifurcation scales.

Results of edge detection. Let us now apply the integrated edge detection scheme to different real-world images. In brief, edges are extracted as follows [Lindeberg, 1996b]: The differential descriptors in the edge definition (1.43) are rewritten in terms of partial derivatives in Cartesian coordinates and are computed at a number of scales in scale-space. Then, a polygon approximation is constructed of the intersections of the two zero-crossing surfaces of L_{vv} and $\partial_t(\mathcal{E}_{\gamma-norm})$ that satisfy the sign conditions $L_{vvv} < 0$ and $\partial_t(\mathcal{E}_{\gamma-norm}) < 0$. Finally, a significance measure is computed for each edge by integrating the normalized edge strength measure along the curve

$$H(\Gamma) = \int_{(x;\ t) \in \Gamma} \sqrt{(\mathcal{G}_{\gamma-norm}L)(x;\ t)}\, ds, \qquad (1.46)$$

$$T(\Gamma) = \int_{(x;\ t) \in \Gamma} \sqrt[4]{(\mathcal{T}_{\gamma-norm}L)(x;\ t)}\, ds. \qquad (1.47)$$

Fig. 1.6 shows the result of applying this scheme to two real-world images. As can be seen, the sharp edges due to object boundaries are extracted as well as the diffuse edges due to illumination effects (the occlusion shadows on the arm and the cylinder, the cast shadow on the table, as well as the reflection on the table). (Recall from fig. 1.4 that for this image it is impossible to capture the entire shadow edge at one scale without introducing severe shape distortions at the finger tip.)

Fig. 1.7 illustrates the ranking on significance obtained from the integrated edge strength along the curve. Whereas there are inherent limitations in using such an entity as the *only* measure of saliency, note that this measure captures essential information.

Fig. 1.8 gives a three-dimensional illustration of how the selected scale levels vary along the edges. The scale-space edges have been drawn as three-dimensional curves in scale-space, overlayed on a low-contrast copy of the original grey-level image in such a way that the height over the image plane represents the selected scale. Observe that coarse scales are selected for the diffuse edge structures due to illumination effects and that finer scales are selected for the sharp edge structures due to object boundaries.

Fig. 1.9 shows the result of applying edge detection with scale selection based on local maxima over scales of $\mathcal{T}_{\gamma-norm}L$ to an image containing a large amount of fine-scale information. From a first view, these results may appear very similar to the result of traditional edge detection at a fixed (very fine) scale. A more detailed study, however, reveals that a number
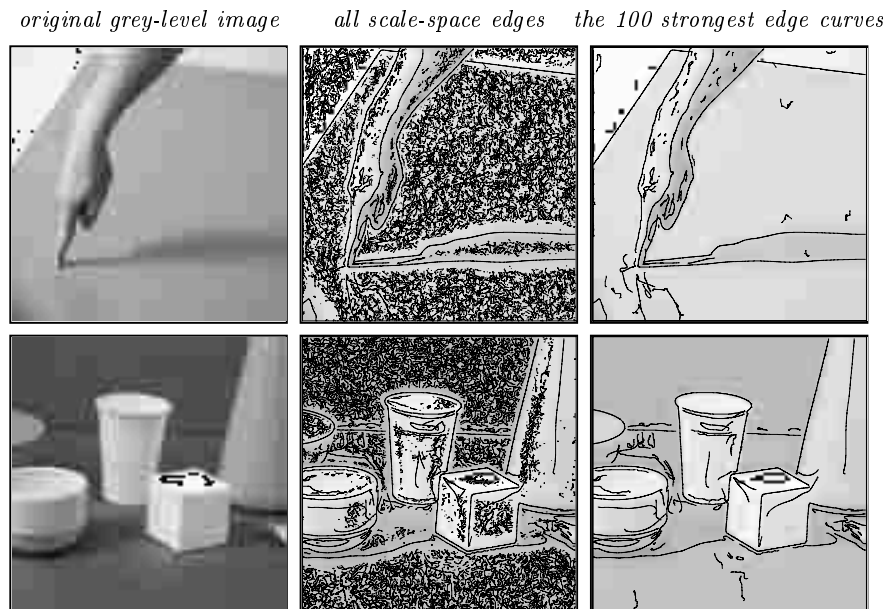
original grey-level image   *all scale-space edges*   *the 100 strongest edge curves*

Figure 1.6: *The result of edge detection with automatic scale selection based on local maxima over scales of the first order edge strength measure $\mathcal{G}_{\gamma-norm}L$ with $\gamma = \frac{1}{2}$. The middle column shows all the scale-space edges, whereas the right column shows the 100 edge curves having the highest significance values. Image size: 256×256 pixels.*

50 most significant edges   20 most significant edges   10 most significant edges
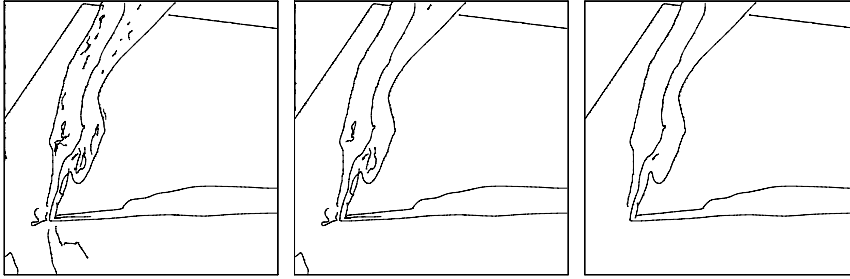

Figure 1.7: *Illustration of the ranking on saliency obtained from the integrated γ-normalized gradient magnitude along the scale-space edges. Here, the 50, 20, and 10 most significant edges, respectively, have been selected from the arm image.*
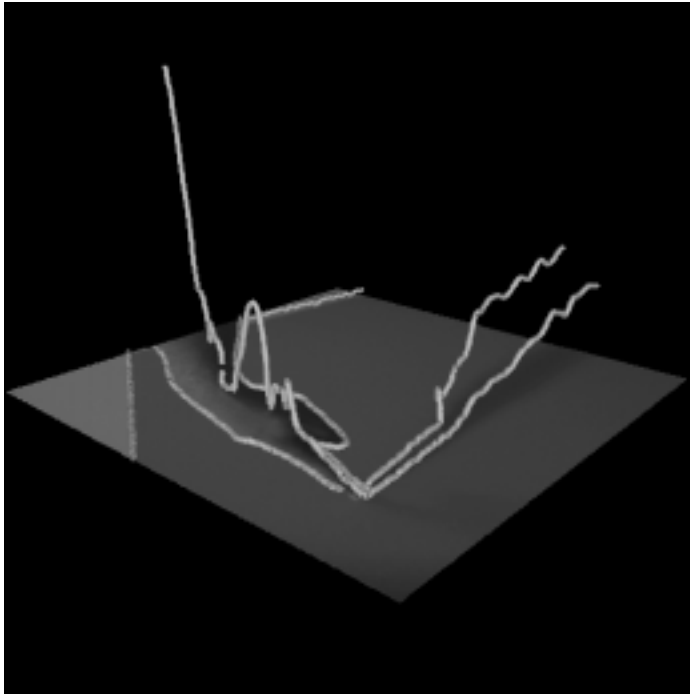

Figure 1.8: *Three-dimensional view of the 10 most significant scale-space edges extracted from the arm image. From the vertical dimension representing the selected scale measured in dimension length (in units of $\sqrt{t}$), it can be seen how coarse scales are selected for the diffuse edge structures (due to illumination effects) and that finer scales are selected for the sharp edge structures (the object boundaries).*
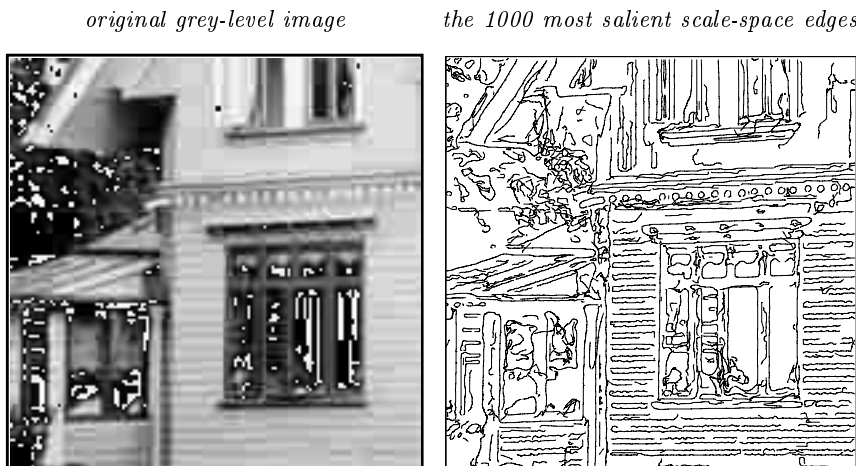
original grey-level image   *the 1000 most salient scale-space edges*

Figure 1.9: *The 1000 strongest scale-space edges extracted using scale selection based on local maxima over scales of* $\mathcal{T}_{\gamma-norm}L$ *(with* $\gamma = \frac{1}{2}$*).   (Image size:* 256×256 *pixels.)*

of shadow edges are extracted, which would be impossible to detect at the same scale as the dominant fine-scale information. In this context, it should be noted that the fine-scale edge detection in this case is not the result of any manual setting of tuning parameters. It is a direct *consequence* of the scale-space edge concept, and is the result of applying the same mechanism as extracts coarse scale levels for diffuse image structures.

Summary.   To conclude, for both these measures of edge strength, this scale selection scheme has the desirable property of adapting the scale levels to the local image structure such that *the selected scales reflect the degree of diffuseness of the edge.*

1.4.2   Ridge detection

By a slight reformulation, ridge detection algorithms can be expressed in a similar way. If we follow a differential geometric approach, and define a bright (dark) ridge point as a point for which the brightness assumes a maximum (minimum) in the main eigendirection of the Hessian matrix [Haralick, 1983; Eberly et al., 1994; Koenderink and van Doorn, 1994; Lindeberg, 1994c], then in the (p, q)-system this definition can be stated as

$$\left\{ \begin{array}{rcl} L_p & = & 0, \\ L_{pp} & < & 0, \\ |L_{pp}| & \geq & |L_{qq}|, \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{rcl} L_q & = & 0, \\ L_{qq} & < & 0, \\ |L_{qq}| & \geq & |L_{pp}|. \end{array} \right.$$

In the (u, v)-system, this condition can for non-degenerate L equivalently be written

$$\left\{ \begin{array}{rcl} L_{uv} & = & 0, \\ L_{uu}^2 - L_{vv}^2 & > & 0, \end{array} \right. \tag{1.48}$$

where the sign of L_{uu} determines the polarity; $L_{uu} < 0$ corresponds to bright ridges, and $L_{uu} > 0$ to dark ridges. Figure 1.4 shows the results of applying this ridge detector at different scales.

In analogy with Section 1.4.1, let us next sweep out a *ridge surface* in scale-space by applying this ridge definition at all scales. Then, given a measure $\mathcal{R}_{\gamma-norm}L$ of normalized ridge strength, define a *scale-space ridge* as a curve on this surface along which the ridge strength measure assumes local maxima with respect to scale.

$$\left\{ \begin{array}{l} \partial_t(\mathcal{R}_{\text{norm}}L(x, y; \; t)) = 0, \\ \partial_{tt}(\mathcal{R}_{\text{norm}}L(x, y; \; t)) < 0, \end{array} \right.$$
$$\left\{ \begin{array}{l} L_p(x, y; \; t) = 0, \\ L_{pp}(x, y; \; t) < 0. \end{array} \right. \tag{1.49}$$

Here, we consider the following ridge strength measures:

$$\mathcal{N}_{\gamma-norm}L = t^{4\,\gamma}\,(L_{pp}^2 - L_{qq}^2)^2 = t^{4\,\gamma}(L_{xx} + L_{yy})^2((L_{xx} - L_{yy})^2 + 4\,L_{xy}^2) \tag{1.50}$$

Qualitative properties. For a Gaussian ridge defined by $g(x, y) = h(x; \; t_0)$, it can be shown that the selected scale will then be $t_{\mathcal{R}_{\gamma-norm}} = \frac{2\,\gamma}{3-2\,\gamma}\,t_0$. Requiring this scale to be $t_{\mathcal{R}_{\gamma-norm}} = t_0$, gives $\gamma = \frac{3}{4}$.

Results of ridge detection. Fig. 1.10 shows the result of applying such a ridge detector to two images and selecting the 100 and 10 strongest bright ridges, respectively, by integrating a measure of normalized ridge strength along each curve.

For the arm image, observe how a coarse-scale descriptor is extracted for the arm as a whole, whereas that the individual fingers give rise to ridge curves at finer scales.

1.4.3 Blob detection

The Laplacian operator $\boldsymbol{\nabla}^2 L = L_{xx} + L_{yy}$ is a commonly used entity for blob detection, since it gives a strong response at the center of blob-like image structures [Marr, 1982; Blostein and Ahuja, 1989; Voorhees and
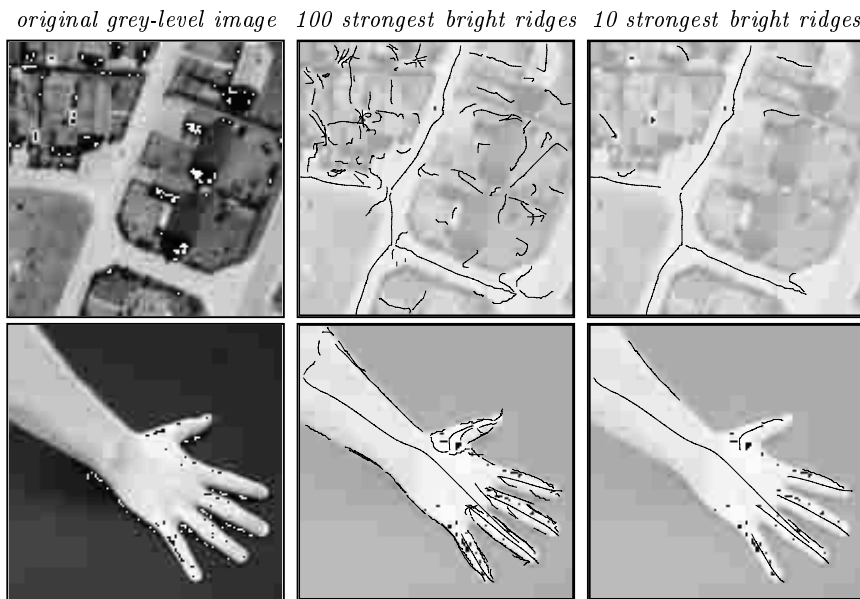
original grey-level image *100 strongest bright ridges* *10 strongest bright ridges*

Figure 1.10: *The 100 and 10 strongest bright ridges respectively extacted using scale selection based on local maxima over scales of $\mathcal{A}_{\gamma-norm}$ (with $\gamma = \frac{3}{4}$). Image size: 128×128 pixels in the top row, and 140×140 pixels in the bottom row.*

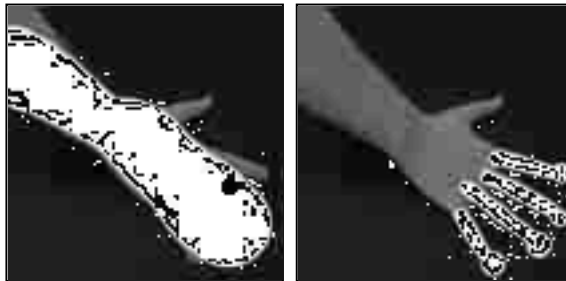*backprojection of ridge 1* *backprojection of ridges 2–5*

Figure 1.11: *Alternative illustration of the five strongest scale-space ridges extracted from the image of the arm in figure 1.10. Each ridge is backprojected onto a dark copy of the original image as the union of a set of circles centered on the ridge curve with the radius proportional to the selected scale at that point.*

Poggio, 1987]. To formulate a blob detector with automatic scale selection, we can consider the points in scale-space at which the the square of the normalized Laplacian

$$\nabla^2_{\mathrm{norm}} L = t(L_{xx} + L_{yy}) \tag{1.51}$$

assumes maxima with respect to space and scale. Such points are referred to as *scale-space extrema* of $(\nabla^2_{\mathrm{norm}} L)^2$.

Qualitative properties.   For a Gaussian blob defined by

$$g(x,y) = h(x,y;\ t_0) = \frac{1}{2\pi t_0} e^{-(x^2+y^2)/2t_0} \tag{1.52}$$

it can be shown that the selected scale at the center of the blob is given by

$$\partial_t(\nabla^2_{\mathrm{norm}} L)(0,0;\ t) = 0 \quad \Longleftrightarrow \quad t_{\nabla^2 L} = t_0. \tag{1.53}$$

Hence, the selected scale directly reflects the width t_0 of the Gaussian blob.

Results of blob detection.   Figure 1.12–and Figure1.13 shows the result of applying this blob detector to an image of a sunflower field. In figure 1.12, each blob feature detected as a scale-space maximum is illustrated by a circle, with its radius proportional to the selected scale. Figure 1.13 shows a three-illustration of the same data set, by marking the scale-space extrema by spheres in scale-space. Observe how the size variations in the image are captured by this structurally very simple operation.
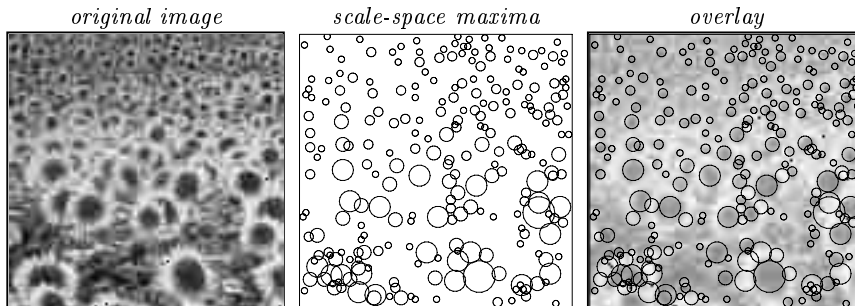

Figure 1.12: Blob detection by detection of scale-space maxima of the normalized Laplacian operator: (a) Original image. (b) Circles representing the 250 scale-space maxima of $(\nabla_{\mathrm{norm}} L)^2$ having the strongest normalized response. (c) Circles overlayed on image.
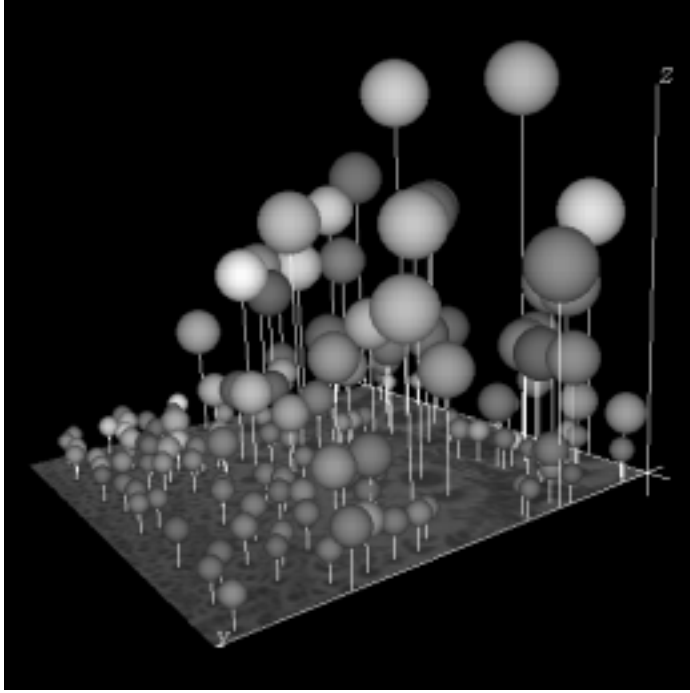
Figure 1.13: *Three-dimensional view of the 150 strongest scale-space maxima of the square of the normalized Laplacian of the Gaussian computed from the sunflower image.*

1.4.4   Corner detection

A commonly used technique for detecting junction candidates in grey-level images is to detect extrema in the curvature of level curves multiplied by the gradient magnitude raised to some power [Kitchen and Rosenfeld, 1982; Koenderink and Richards, 1988]. A special choice is to multiply the level curve curvature by the gradient magnitude raised to the power of three. This leads to the differential invariant $\tilde{\kappa} = L_v^2 L_{uu}$, with the corresponding normalized expression

$$\tilde{\kappa}_{\mathrm{norm}} = t^{2\gamma} L_v^2 L_{uu} \tag{1.54}$$

Qualitative properties.   For a *diffuse L-junction* $g(x_1, x_2) = \Phi(x_1;\ t_0)\,\Phi(x_2;\ t_0)$ modelled as the product of two diffuse step edges $\Phi(x_i;\ t_0) = \int_{x'=-\infty}^{x_i} h(x';\ t_0)\,dx'$, it can be shown that variation of $\tilde{\kappa}_{\mathrm{norm}}$ at the origin is given by

$$|\tilde{\kappa}_{\mathrm{norm}}(0,0;\ t)| = \frac{t^{2\gamma}}{8\pi^2 (t_0 + t)^2}. \tag{1.55}$$
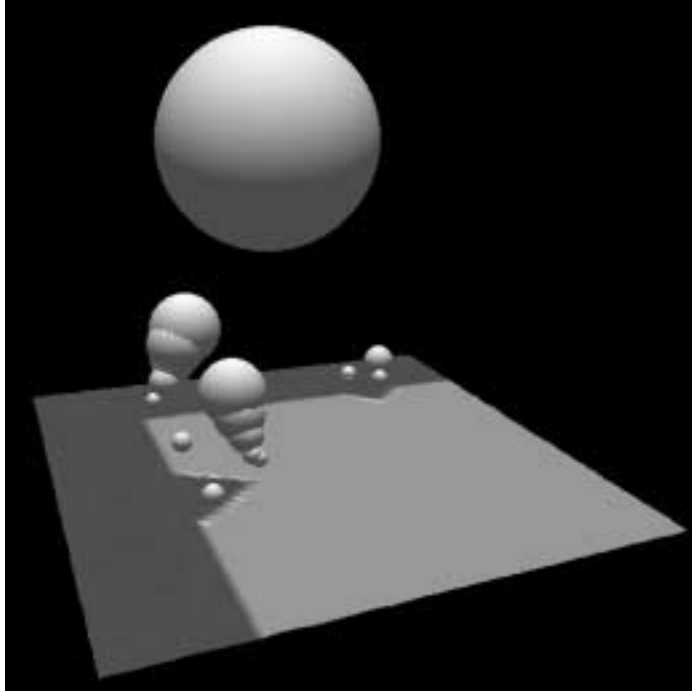
Figure 1.14: *Three-dimensional view of scale-space maxima of $\tilde{\kappa}^2_{\mathrm{norm}}$ computed for a large scale corner with superimposed corner structures at finer scales.*

When $\gamma = 1$, this entity *increases monotonically* with scale, whereas for $\gamma \in \;]0, 1[$, $\tilde{\kappa}_{\mathrm{norm}}(0, 0; \; t)$ assumes a unique maximum over scales at $t_{\tilde{\kappa}} = \frac{\gamma}{1-\gamma} \, t_0$. On the other hand, for a *non-uniform Gaussian blob* $L(x_1, x_2; \; t) = h(x_1; \; t_1 + t) \, h(x_2; \; t_2 + t)$, the normalized response always *decreases* with scale at sufficiently coarse scales.

This analysis indicates that when $\gamma = 1$, $\tilde{\kappa}^2_{\mathrm{norm}}$ can be expected to increase with scales when a single corner model of infinite extent constitutes a reasonable approximation, whereas $\tilde{\kappa}^2_{\mathrm{norm}}$ can be expected to decrease with scales when so much smoothing is applied that the overall shape of the object is substantially distorted.

Hence, selecting scale levels (and spatial points) where $\tilde{\kappa}^2_{\mathrm{norm}}$ assumes maxima over scales can be expected to give rise to scale levels in the intermediate scale range (where a finite extent junction model constitutes a reasonable approximation) and the selected scale levels reflect thus reflect over how large region a corner model is valid. In practice, a slightly smaller value of $\gamma = 7/8$ is used.

Results of corner detection.   Figure 1.14 shows the result of detecting scale-space extrema from an image with corner structures at multiple scales. Observe that a coarse scale response is obtained for the large scale corner structure as a whole, whereas the superimposed corner structures of smaller size give rise to scale-space maxima at finer scales. (More results on real images will be shown in Section 1.5.)

1.4.5   Local frequency estimation

To extend the abovementioned application of the scale selection methodology from the detection of sparse image features to the computation of dense image descriptors, a natural starting point is to consider the theory of quadrature filter pairs defined (from a Hilbert transform) in such a way as to be phase independent for any sine wave.

To approximate such operators within the Gaussian derivative framework, we can define a corresponding quasi quadrature measure in the one-dimensional case by [Koenderink and van Doorn, 1987; Lindeberg, 1997b]

$$\mathcal{P}L = L_\xi^2 + C L_{\xi\xi}^2 = t L_x^2 + C t^2 L_{xx}^2 \tag{1.56}$$

where a good choice of the parameter C is $C = e/4$. Note that, in order to achieve scale invariance, it is necessary to use normalized derivatives with $\gamma = 1$ in the (inhomogeneous) linear combination of Gaussian derivatives of different order. The γ-normalized derivative concept, however, leaves a degree of freedom, which can be parameterized by

$$\mathcal{P}L = t^{-\Gamma} (t L_x^2 + C t^2 L_{xx}^2). \tag{1.57}$$

To extend this entity to two-dimensional signals, we can consider

$$\mathcal{P}L = t^{-\Gamma} \left(t(L_x^2 + L_y^2) + C\, t^2 \left(L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2 \right) \right), \tag{1.58}$$

$$\mathcal{R}L = t^{-\Gamma} \left(t\left(L_x^2 + L_y^2 \right) + C\, t^2 \left((L_{xx} - L_{yy})^2 + 4\, L_{xy}^2 \right) \right). \tag{1.59}$$

Both these differential expressions are invariant under rotations and reduce to the form (1.57) for a one-dimensional signal. The second order differential expression in (1.58),

$$\mathcal{S}_{\gamma-norm}L = t^{2\,\gamma} \left(L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2 \right) \tag{1.60}$$

however, is a natural measure of the total amount of second-order information in the signal, whereas the second order differential expression in (1.59)

$$\mathcal{A}_{\gamma-norm}L = t^{2\,\gamma} \left(L_{pp} - L_{qq} \right)^2 = t^{2\,\gamma} \left((L_{xx} - L_{yy})^2 + 4\, L_{xy}^2 \right) \tag{1.61}$$

is more specific to elongated structures (e.g. ridges). The specific choice of $\Gamma = 1/2$ means that (1.59) and can be interpreted as a linear combination of the edge strength measure (1.44) with $\gamma = 1/2$ and the ridge strength measure (1.61) with $\gamma = 3/4$.

Qualitative properties. These differential expressions inherit similar scale selection properties for sine waves as those described in Section 1.3.1; see [Lindeberg, 1996a],[Lindeberg, 1997b] for an analysis.
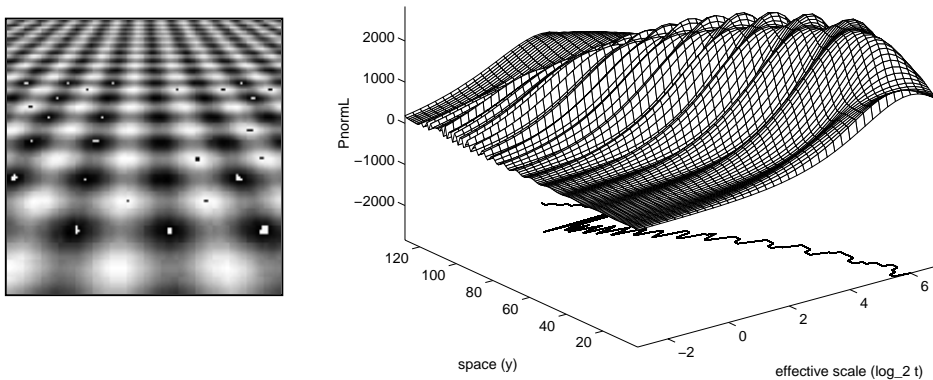

Figure 1.15: *Dense scale selection by maximizing the quasi quadrature measure (1.58) over scales: (left) Original grey-level image. (right) The variations over scales of the quasi quadrature measure $\mathcal{P}L$ computed along a vertical cross-section through the center of the image. The result is visualized as a surface plot showing the variations over scale of the quasi quadrature measure as well as the position of the first local maximum over scales.*

Results of frequency estimation. Figure 1.15 shows an example result of estimating local frequencies in this way, by detecting local maxima over scale of $\mathcal{P}L$ along a vertical cross-section in an image of a periodic pattern. Observe how the selected scale levels capture the variations in frequency caused by the perspective effects.

1.5 Feature localization with automatic scale selection

The scale selection techniques presented so far are useful in the stage of detecting image features, and the role of the scale selection mechanism is to estimate the approximate size of the image structures the feature detector responds to. When computing features at coarse scales in scale-space,

however, the shape distortions can be significant, and in many cases it is desirable to complement feature detection modules by an explicit *feature localization* stage.

The subject of this section is show how mechanism for automatic scale selection can be formulated in this context, by minimizing measures of inconsistency over scales.

1.5.1   Corner localization

Given an approximate estimate x_0 of the location and the size s of a corner (computed according to Section 1.4.4), an improved estimate of the corner position can be computed as follows [Förstner and Gülch, 1987; Lindeberg, 1994d]: Consider at every point $x' \in R^2$ in a neighbourhood of x_0, the line $l_{x'}$ perpendicular to the gradient vector $(\boldsymbol{\nabla} L)(x') = (L_{x_1}, L_{x_2})^T (x')$ at that point:

$$D_{x'}(x) = ((\boldsymbol{\nabla} L)(x'))^T (x - x') = 0. \tag{1.62}$$

Then, minimize the perpendicular distance to all lines $l_{x'}$ in a neighbourhood of x_0, weighted by the gradient magnitude, *i.e.* determine the point $x \in R^2$ that minimizes

$$\min_{x \in R^2} \int_{x' \in R^2} (D_{x'}(x))^2 \, w_{x_0}(x'; \, s) \, dx' \tag{1.63}$$

for a Gaussian window function $w_{x_0}(\cdot; \, s) \colon R^2 \to R$ with integration scale s set from the detection scale $t_{tilde\kappa}$ of the corner and centered at the candidate junction x_0. After expansion, this minimization problem be expressed as a standard least squares problem

$$\min_{x \in R^2} x^T A \, x - 2 \, x^T b + c \quad \Longleftrightarrow \quad A \, x = b, \tag{1.64}$$

where $x = (x_1, x_2)^T$, and A, b, and c are determined by the local statistics of the gradient directions $\boldsymbol{\nabla} L(\cdot; \, t)$ at scale t in a neighbourhood of x_0 (compare with (1.22))

$$A(x; \, t, s) = \int_{x' \in R^2} (\boldsymbol{\nabla} L)(x') \, (\boldsymbol{\nabla} L)^T(x') \, w_{x_0}(x'; \, s) \, dx',$$

$$b(x; \, t, s) = \int_{x' \in R^2} (\boldsymbol{\nabla} L)(x') \, (\boldsymbol{\nabla} L)^T(x') \, x' \, w_{x_0}(x'; \, s) \, dx', \tag{1.65}$$

$$c(x; \, t, s) = \int_{x' \in R^2} x'^T \, (\boldsymbol{\nabla} L)(x') \, (\boldsymbol{\nabla} L)^T(x') \, x' w_{x_0}(x'; \, s) \, dx'.$$

1.5.2 Scale selection principle for feature localization

To express a scale selection mechanism for this corner localizer, let us extend the minimization problem (1.64) from a single scale to optimization over multiple scales [Lindeberg, 1994a]

$$\min_{t \in R_+} \min_{x \in R^2} \frac{x^T A\, x - 2\, x^T b + c}{\text{norm}(t)} = \min_{t \in R_+} \min_{x \in R^2} \frac{c - b^T A^{-1} b}{\text{trace}\, A} \qquad (1.66)$$

and introduce a normalization factor norm(t) to relate minimizations at different scales. The particular choice of norm(t) = trace A implies that the *normalized residual*

$$\tilde{r} = \min_{x \in R^2} \frac{\int_{x' \in R^2} \left| ((\boldsymbol{\nabla} L)(x'))^T (x - x') \right|^2 w_{x_0}(x';\ s)\, dx'}{\int_{x' \in R^2} \left| (\boldsymbol{\nabla} L)(x') \right|^2 w_{x_0}(x';\ s)\, dx'} \qquad (1.67)$$

has dimension [length]^2 can be interpreted as a weighted estimate of the localization error. Specifically, scale selection according to (1.66), by minimizing the normalized residual $\tilde{r}$ (1.67) over scales, corresponds to selecting the scale that minimizes the estimated inaccuracy in the localization estimate.
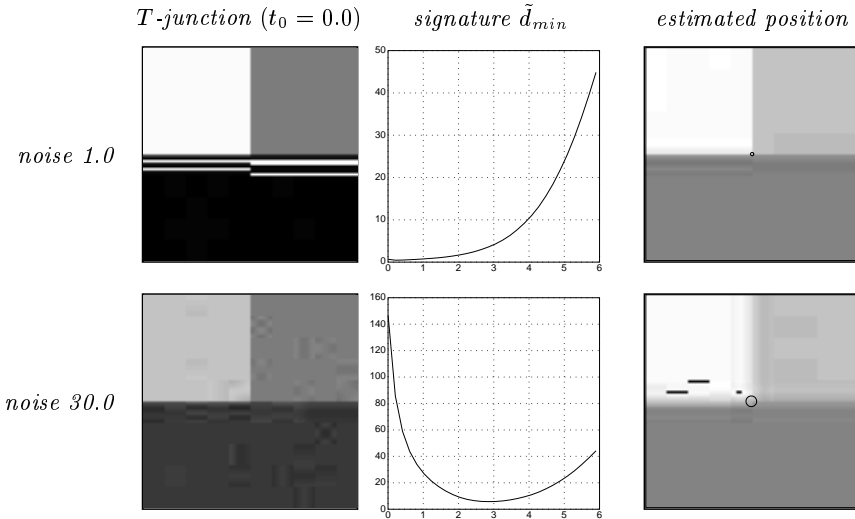

T-junction ($t_0 = 0.0$)   *signature $\tilde{d}_{min}$*   *estimated position*

noise 1.0

noise 30.0

Figure 1.16: *Corner localization by minimizing the normalized residuals over scales for two corner structures. A basic property of the scale selection mechanism is that an increase in the noise level implies that the minimum over scales is assumed at a coarser scale.*

Qualitative effects. Figure 1.16 shows the result of performing corner localization in this way, by minimizing the normalized residual $\tilde{r}$ over scales for two corner structures. Observe how an increase in the noise level implies that the minimum over scales is assumed at a coarser scale.
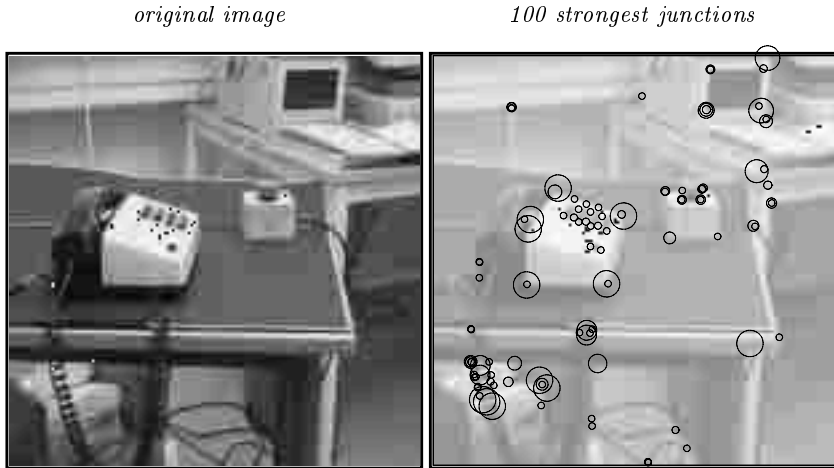
<div align="center">original image   100 strongest junctions</div>

Figure 1.17: Results of composed two-stage junction detection followed by junction localization. (left) Original grey-level image. (right) The 100 strongest junction responses ranked according to the scale-space maxima of $\tilde{\kappa}_{\mathrm{norm}}^2$ and illustrated by circles with their radii proportional to the detection scales of the junction responses.

Results of corner localization. Figure 1.17 shows the result of integrating this corner localization module with the corner detector in Section 1.4.4. The resulting two-stage corner automatically adapts its detection scales and localization scales to size variations and noise variations in the image structures. In addition, the region of interest associated with each corner is useful for purposes such as matching.

Edge localization. In [Lindeberg, 1994d], [Lindeberg, 1996a] it is outlined how minimization over scales of a similar normalized residual applies to the problem of edge localization.

1.6   Stereo matching with automatic scale selection

This section shows how a scale selection mechanism can be formulated for a differential stereo matchning scheme expressed within the Gaussian derivative framework. With appropriate modifications, similar ideas apply to the problem of flow estimation.

1.6.1   Background: Least squares differential stereo matching

Let us assume that the flow field between the scale-space representations L and R of two images can be approximated by a constant flow field v over the support region of a window function w. Following [Lukas and Kanade, 1981], [Bergen et al., 1992], [Förstner, 1993], [Lindeberg, 1995], [Lindeberg, 1996d] and several others, consider the discrete form of the *motion constraint equation* [Horn and Schunck, 1981]

$$(\boldsymbol{\nabla} L)(\xi)^T (\Delta \xi) + (L(\xi) - R(\xi)) = \mathcal{O}(|\Delta \xi|^2). \tag{1.68}$$

and integrate the square of this relation using w as window function. After expansion (and dropping the arguments) this gives the least squares problem

$$\min_{v \in R^2} v^T A v + 2b^T v + c, \tag{1.69}$$

where A, b, and c are defined by

$$A = \int_{\xi \in R^2} (\boldsymbol{\nabla} L)(\boldsymbol{\nabla} L)^T \, w \, d\xi$$

$$b = \int_{\xi \in R^2} (R - L) \, (\boldsymbol{\nabla} L) \, w \, d\xi \tag{1.70}$$

$$c = \int_{\xi \in R^2} (R - L)^2 \, w \, d\xi.$$

If we interpret $R - L$ as a discrete approximation to a temporal derivative, these image descriptors fall within the class of windowed spectral moments in (1.22). Assuming that A according to (??) is non-degenerate, the explicit solution of the flow estimate is

$$v = -A^{-1} b. \tag{1.71}$$

1.6.2   Scale selection principle for estimating image deformations

When implementing this scheme in practice, it is natural to express it within a coarse-to-fine multi-scale framework. When performing image matching from coarse to fine scales, however, it is not obvious what should be the finest scales. If we attempt to match image features beyond a certain resolution, we could expect the errors in the deformation estimates to increase rather than to decrease.

In analogy with Section 1.5.2, a natural way to formulate a scale selection mechanism for the differential stereo matching scheme is by extending the least squares estimation (1.69) to the two-parameter least squares problem

$$\min_{t \in R_+} \min_{v \in R^2} \frac{v^T A v + 2b^T v + c}{\text{norm}(t)}, \tag{1.72}$$

where the normalization factor norm(t) determines how the information at different scales should be compared. Again, we choose norm(t) = trace A, while one could also conceive other normalization approaches, such as the minimum eigenvalue of A.

The resulting *normalized residual* is of dimension [length]^2 and constitutes a first-order approximation of the following error measure

$$
\begin{aligned}
\mathcal{E}_{\boldsymbol{\nabla} L} &= \frac{\int_{\xi \in R^2} |(\boldsymbol{\nabla} L)^T (v - \Delta \xi)|^2 \, w(\xi) \, d\xi}{\int_{\xi \in R^2} |\boldsymbol{\nabla} L|^2 \, w(\xi) \, d\xi} \\
&= \frac{\int_{\xi \in R^2} (v - \Delta \xi)^T (\boldsymbol{\nabla} L)(\boldsymbol{\nabla} L)^T (v - \Delta \xi) \, w(\xi) \, d\xi}{\int_{\xi \in R^2} (\boldsymbol{\nabla} L)^T (\boldsymbol{\nabla} L) \, w(\xi) \, d\xi}
\end{aligned}
\tag{1.73}
$$

where v is the regional flow estimate and $\Delta \xi$ a pointwise flow estimate that satisfies (1.68). In other words, $\mathcal{E}_{\boldsymbol{\nabla} L}$ can be seen as a measure of the internal consistency of the estimated flow field, weighted by the gradient magnitude.

1.6.3 Properties of matching with automatic scale selection

Selection of coarser scales for larger size image structures. Figure 1.18 shows two synthetic image patterns which have been subject to a uniform expansion. The underlying patterns are identical except for the size of the texture elements which differs by a factor of four, and 10 % white Gaussian noise added to each image independently after the deformation. Observe for the small size pattern in the first row the minima over scales in $\tilde{r}$ are assumed at the finest scales, while when the size of the image structures is increased in the second row, the minimum over scales is assumed at coarser scales. This behaviour agrees with the intuitive notion that coarser scales should be selected for patterns containing larger size image structures.

Selection of coarser scales with increasing noise level.   In figure 1.19 the image pattern is the same, whereas the noise level is varied. Observe that with an increasing amount of interfering fine scale structures, the minimum in $\tilde{r}$ over scales is assumed at coarser scales. This behavior agrees with the intuitive notion that a larger amount of smoothing is required for noisy data than otherwise similar data with less noise.

Selection of finer scales near discontinuities in the deformation field.
Figure 1.20 shows the behavior of the scale selection method in the neighborhood of a discontinuity in the flow field. For a "wedding-cake type" random dot stereo pair to which 1 % white Gaussian noise has been added, the results are shown of accumulating the scale-space signature of the normalized residual in three windows with different distance to the discontinuity.
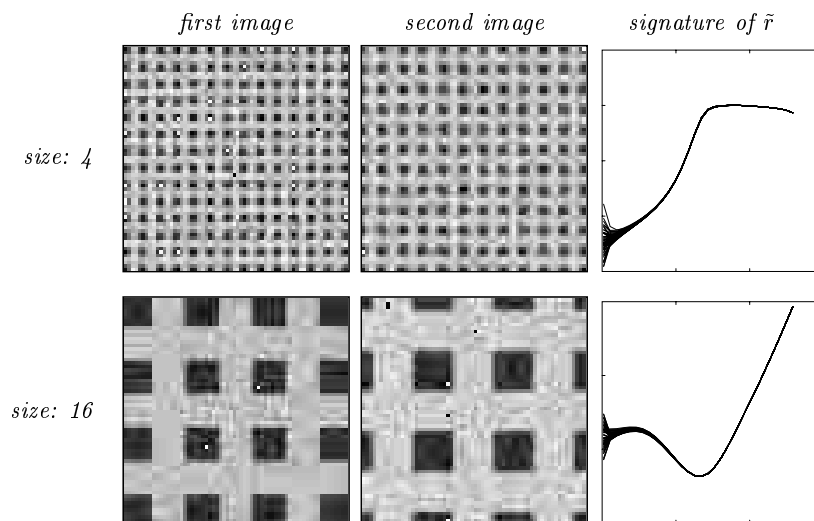
Figure 1.18: *Scale-space signatures of the normalized residual $\tilde{r}$ for synthetic expanding patterns with structures at different scales. Notice that with increasing size of the texture elements, the minimum over scales in the normalized residual is assumed at coarser scales.*
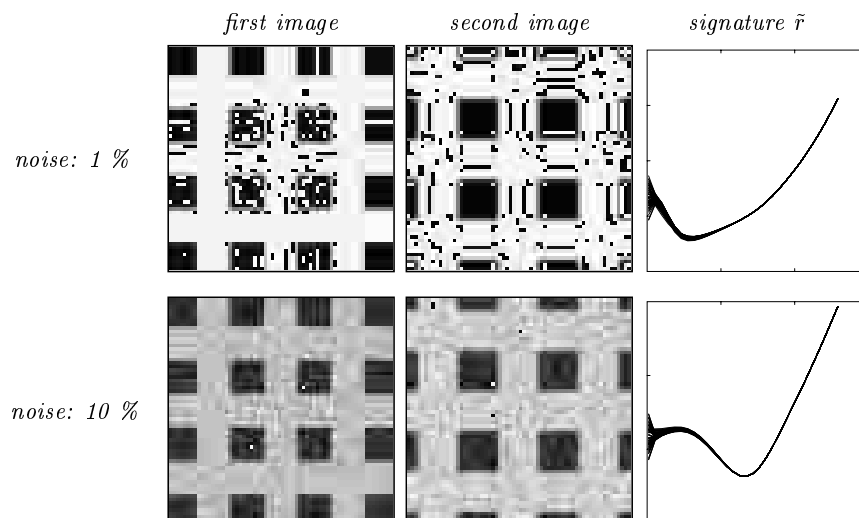

Figure 1.19: *Scale-space signatures of the normalized residual $\tilde{r}$ for a synthetic expanding pattern with different amounts of added white Gaussian noise. Observe that with increasing noise level, the minimum over scales in the normalized residual is assumed at coarser scales.*
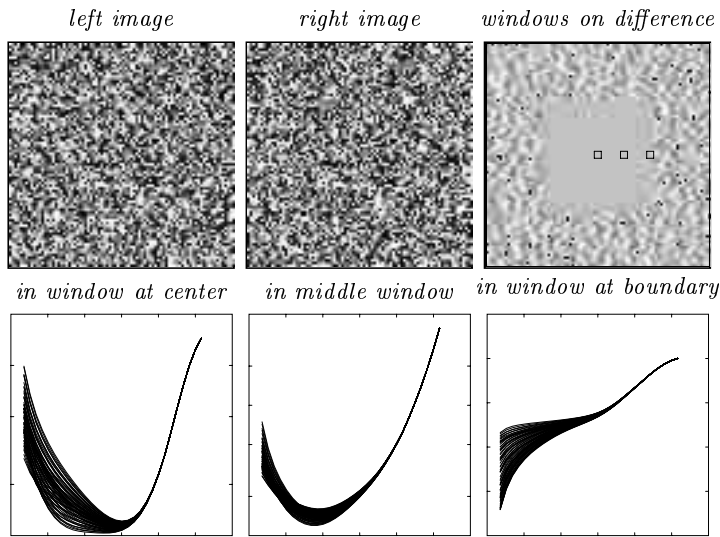
Figure 1.20: *The qualitative behaviour of the scale selection method at a discontinuity in the deformation field. The bottom row shows scale-space signatures of the normalized residual computed in three windows with different distance to the discontinuity (with their positions indicated in the upper right image by three squares overlayed on the pointwise difference between the left and the right image). Observe that with decreasing distance to the discontinuity, the minimum over scales is assumed at finer scales.*

Figure 1.21: *Selected scale levels along a central horizontal cross-section through the wedding-cake type random dot stereo pair in figure 1.20. Observe that distinct minima are obtained at the two discontinuities in the disparity field.*

These windows have been uniformly spaced from the image center to one of the discontinuities in the disparity field as shown in figure 1.20(c).

Observe that with decreasing distance to the discontinuity, the minimum over scales is assumed at finer scales. This qualitative behavior agrees with the intuitive notion that smaller windows for matching should be selected for image structures near a discontinuity in the disparity field than when matching otherwise similar image structures in a region where the disparity varies smoothly.

Notably, this rapid decrease of the selected scale levels could also provide a clue for detecting flow field discontinuities and signaling possible occlusions.

1.7 Summary and conclusions

The scale-space framework provides a canonical way to model early visual operations in terms of linear and non-linear combinations of Gaussian derivatives of different order, orientation and scale. This chapter has shown how scale-space descriptors can be complemented by mechanisms for automatic scale selection

For feature detectors expressed in terms of Gaussian derivatives, hypotheses about interesting scale levels can be generated from scales at which normalized measures of feature strength assume local maxima with respect to scale. The notion of γ-normalized derivatives arises by necessity given the requirement that the scale selection mechanism should commute with rescalings of the image pattern. Specific examples have been shown how feature detection algorithms with automatic scale selection can be formulated for the problems of edge detection, blob detection, junction detection, ridge detection and frequency estimation. A general property of this scheme is that the selected scale levels reflect the size of the image structures.

When estimating image deformations, such as in image matching and optic flow computations, scale levels with associated deformation estimates can be selected from the scales at which normalized measures of uncertainty assume local minima with respect to scales. It has been illustrated how an integrated scale selection and flow estimation algorithm has the qualitative properties of leading to the selection of coarser scales for larger size image structures and increasing noise level, whereas it leads to the selection of finer scales in the neighbourhood of flow field discontinuities.

Further reading. The chapter is based on an invited presentation at the First International Conference on Scale-Space Theory in Computer Vision held in Utrecht, Netherlands in July 1997. The material has been selected from [Lindeberg, 1993c], [Lindeberg, 1994a], [Lindeberg, 1994d], [Linde-

berg, 1995], [Lindeberg, 1996c], [Lindeberg, 1996a], [Lindeberg, 1996d]; see also [Lindeberg, 1991], [Lindeberg, 1993a] for complementary works.

Applications of these scale selection principles to various problems in computer vision have been presented in [Lindeberg and Gårding, 1993], [Lindeberg and Gårding, 1997], [Gårding and Lindeberg, 1996], [Lindeberg and Li, 1995], [Lindeberg and Li, 1997], [Bretzner and Lindeberg, 1998a], [Bretzner and Lindeberg, 1997], [Bretzner and Lindeberg, 1998b], [Almansa and Lindeberg, 1996], [Almansa and Lindeberg, 1998], [Wiltschi et al., 1997b], [Wiltschi et al., 1997a], [Lindeberg, 1997b].

For related works see [Mallat and Zhong, 1992], [Mallat and Hwang, 1992], [Korn, 1988], [Zhang and Bergholm, 1993], [Pizer et al., 1994], [Eberly et al., 1994], [Koller et al., 1995], [Jägersand, 1995], [Kanade and Okutomi, 1994], [Battiti et al., 1990], [Niessen and Maas, 1996], [Elder and Zucker, 1996], [Yacoob and Davis, 1997].

Bibliography

A. Almansa and T. Lindeberg. Enhancement of fingerprint images by shape-adapted scale-space operators. In J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors, *Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory*, Copenhagen, Denmark, May. 1996a. Kluwer Academic Publishers.

A. Almansa and T. Lindeberg. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale selection. *Technical Report ISRN KTH NA/P–98/03–SE.*, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, Apr. 1998.

J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Trans. Pattern Analysis and Machine Intell.*, 8(1):26–33, 1986.

R. Battiti, E. Amaldi, and C. Koch. Computing optical flow across multiple scales: An adaptive coarse-to-fine strategy. *Int. J. of Computer Vision*, 6(2):133–145, 1990.

J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In G. Sandini, editor, *Proc. 2nd European Conf. on Computer Vision*, volume 588 of *Lecture Notes in Computer Science*, pages 237–252, Santa Margherita Ligure, Italy, May. 1992. Springer-Verlag.

J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. Pattern Analysis and Machine Intell.*, 13(8):775–790, Aug. 1991.

D. Blostein and N. Ahuja. Shape from texture: integrating texture element extraction and surface estimation. *IEEE Trans. Pattern Analysis and Machine Intell.*, 11(12):1233–1251, Dec. 1989.

L. Bretzner and T. Lindeberg. On the handling of spatial and temporal scales in feature tracking. In ter Haar Romeny et al., editor, *Proc. 1st Int. Conf. on Scale-Space Theory in Computer Vision*, pages 128–139, Utrecht, The Netherlands, July 1997. Springer Verlag, New York.

L. Bretzner and T. Lindeberg. Feature tracking with automatic selection of spatial scales. *Computer Vision and Image Understanding*, 1998a. (To appear).

L. Bretzner and T. Lindeberg. Use your hand as a 3-D mouse, or, relative orientation from extended sequences of sparse point and line correspondences using the affine trifocal tensor. In *Proc. 5th European Conference on Computer Vision* (Hans Burkhardt and Bernd Neumann, eds.), vol. 1406 of *Lecture Notes in Computer Science*, (Freiburg, Germany), pp. 141–157, Springer Verlag, Berlin, June 1998.

J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intell.*, 8(6):679–698, 1986.

D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach. Ridges for image analysis. *J. of Mathematical Imaging and Vision*, 4(4):353–373, 1994.

J. H. Elder and S. W. Zucker. Local scale control for edge detection and blur estimation. In *Proc. 4th European Conference on Computer Vision*, volume 1064 of *Lecture Notes in Computer Science*, Cambridge, UK, Apr. 1996. Springer Verlag, Berlin.

D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. of the Optical Society of America*, 4:2379–2394, 1987.

L. M. J. Florack. *The Syntactical Structure of Scalar Images*. PhD thesis, Dept. Med. Phys. Physics, Univ. Utrecht, NL-3508 Utrecht, Netherlands, 1993.

L. M. J. Florack. *Image Structure*. Series in Mathematical Imaging and Vision. Kluwer Academic Publishers, Dordrecht, Netherlands, 1997.

L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Scale and the differential structure of images. *Image and Vision Computing*, 10(6):376–388, Jul. 1992.

L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Cartesian differential invariants in scale-space. *J. of Mathematical Imaging and Vision*, 3(4):327–348, 1993.

W. A. Förstner. Image matching. In R. M. Haralick and L. G. Shapiro, editors, *Computer and Robot Vision*, volume II, pages 289–378. Addison-Wesley, 1993.

W. A. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centers of circular features. In *Proc. Intercommission Workshop of the Int. Soc. for Photogrammetry and Remote Sensing*, Interlaken, Switzerland, 1987.

W. T. Freeman and E. H. Adelson. Steerable filters for early vision, image analysis and wavelet decomposition. In *Proc. 3rd Int. Conf. on Computer Vision*, Osaka, Japan, Dec. 1990. IEEE Computer Society Press.

J. Gårding and T. Lindeberg. Direct computation of shape cues using scale-adapted spatial derivative operators. *Int. J. of Computer Vision*, 17(2): 163–191, 1996.

R. M. Haralick. Ridges and valleys in digital images. *Computer Vision, Graphics, and Image Processing*, 22:28–38, 1983.

B. K. P. Horn and B. G. Schunck. Determining optical flow. *J. of Artificial Intelligence*, 17:185–204, 1981.

R. A. Hummel and R. Moniot. Reconstructions from zero-crossings in scale-space. *IEEE Trans. Acoustics, Speech and Signal Processing*, 37 (12):2111–2130, 1989.

M. Jägersand. Saliency maps and attention selection in scale and spatial coordinates: An information theoretic approach. In *Proc. 5th International Conference on Computer Vision*, pages 195–202, Cambridge, MA, June 1995.

T. Kanade and M. Okutomi. A stereo matching algorithm with an adaptive window: Theory and experiment. *IEEE Trans. Pattern Analysis and Machine Intell.*, 16(9):920–932, 1994.

L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, 1(2):95–102, 1982.

J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50: 363–370, 1984.

J. J. Koenderink. *Solid Shape*. MIT Press, Cambridge, Massachusetts, 1990.

J. J. Koenderink, A. Kaeppers, and A. J. van Doorn. Local operations: The embodiment of geometry. In G. Orban and H.-H. Nagel, editors, *Artificial and Biological Vision Systems*, pages 1–23, 1992.

J. J. Koenderink and W. Richards. Two-dimensional curvature operators. *J. of the Optical Society of America*, 5:7:1136–1141, 1988.

J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

J. J. Koenderink and A. J. van Doorn. Receptive field families. *Biological Cybernetics*, 63:291–298, 1990.

J. J. Koenderink and A. J. van Doorn. Generic neighborhood operators. *IEEE Trans. Pattern Analysis and Machine Intell.*, 14(6):597–605, Jun. 1992.

J. J. Koenderink and A. J. van Doorn. Two-plus-one-dimensional differential geometry. *Pattern Recognition Letters*, 15(5):439–444, 1994.

T. M. Koller, G. Gerig, G. Szèkely, and D. Dettwiler. Multiscale detection of curvilinear structures in 2-D and 3-D image data. In *Proc. 5th International Conference on Computer Vision*, pages 864–869, Cambridge, MA, June 1995.

A. F. Korn. Toward a symbolic representation of intensity changes in images. *IEEE Trans. Pattern Analysis and Machine Intell.*, 10(5):610–625, 1988.

T. Lindeberg. Scale-space for discrete signals. *IEEE Trans. Pattern Analysis and Machine Intell.*, 12(3):234–254, Mar. 1990.

T. Lindeberg. *Discrete Scale-Space Theory and the Scale-Space Primal Sketch*. Ph. D. dissertation, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, May. 1991. ISRN KTH/NA/P- -91/08--SE. An extended and revised version published as book "Scale-Space Theory in Computer Vision" in The Kluwer International Series in Engineering and Computer Science.

T. Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *Int. J. of Computer Vision*, 11(3):283–318, Dec. 1993a.

T. Lindeberg. Discrete derivative approximations with scale-space properties: A basis for low-level feature extraction. *J. of Mathematical Imaging and Vision*, 3(4):349–376, Nov. 1993b.

T. Lindeberg. On scale selection for differential operators. In K. H. K. A. Høgdra, B. Braathen, editor, *Proc. 8th Scandinavian Conf. on Image Analysis*, pages 857–866, Tromsø, Norway, May. 1993c. Norwegian Society for Image Processing and Pattern Recognition.

T. Lindeberg. Junction detection with automatic selection of detection scales and localization scales. In *Proc. 1st International Conference on Image Processing*, volume I, pages 924–928, Austin, Texas, Nov. 1994a. IEEE Computer Society Press.

T. Lindeberg. On the axiomatic foundations of linear scale-space: Combining semi-group structure with causality vs. scale invariance. Technical Report ISRN KTH/NA/P--94/20--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, Aug. 1994b. Extended version to appear in J. Sporring and M. Nielsen and L. Florack and P. Johansen (eds.) Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory, Copenhagen, Denmark, Kluwer Academic Publishers, May 1996.

T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2):225–270, 1994c. Supplement *Advances in Applied Statistics: Statistics and Images: 2*.

T. Lindeberg. *Scale-Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers, Dordrecht, Netherlands, 1994d.

T. Lindeberg. Direct estimation of affine deformations of brightness patterns using visual front-end operators with automatic scale selection. In *Proc. 5th International Conference on Computer Vision*, pages 134–141, Cambridge, MA, June 1995.

T. Lindeberg. Feature detection with automatic scale selection. Technical Report ISRN KTH/NA/P--96/18--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, May. 1996a. *International Journal of Computer Vision*, vol. 30, no. 2, 1998. in press.

T. Lindeberg. Edge detection and ridge detection with automatic scale selection. Technical Report ISRN KTH/NA/P--96/06--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, Jan. 1996b. *International Journal of Computer Vision*, vol. 30, no. 2, 1998. in press.

T. Lindeberg. Edge detection and ridge detection with automatic scale selection. In *Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, 1996*, pages 465–470, San Francisco, California, June 1996c. IEEE Computer Society Press.

T. Lindeberg. A scale selection principle for estimating image deformations. Technical Report ISRN KTH/NA/P--96/16--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, Apr. 1996d. Image and Vision Computing, (in press).

T. Lindeberg. Scale-space theory: A framework for handling image structures at multiple scales. In *Proc. CERN School of Computing*, pages 27–38, Egmond aan Zee, The Netherlands, Sep. 1996e. Tech. Rep. CERN 96-08.

T. Lindeberg. Linear spatio-temporal scale-space. In ter Haar Romeny et al., editor, *Proc. 1st Int. Conf. on Scale-Space Theory in Computer Vision*, Utrecht, The Netherlands, July 1997a. Springer Verlag, New York.

T. Lindeberg. On automatic selection of temporal scales in time-casual scale-space. In G. Sommer and J. J. Koenderink, editors, *Proc. AFPAC'97: Algebraic Frames for the Perception-Action Cycle*, volume 1315 of *Lecture Notes in Computer Science*, pages 94–113, Kiel, Germany, Sept. 1997b. Springer Verlag, Berlin.

T. Lindeberg and J. Gårding. Shape from texture from a multi-scale perspective. In H.-H. N. et. al., editor, *Proc. 4th Int. Conf. on Computer Vision*, pages 683–691, Berlin, Germany, May. 1993. IEEE Computer Society Press.

T. Lindeberg and J. Gårding. Shape-adapted smoothing in estimation of 3-D depth cues from affine distortions of local 2-D structure. *Image and Vision Computing*, 15:415–434, 1997.

T. Lindeberg and M. Li. Segmentation and classification of edges using minimum description length approximation and complementary junction cues. In G. Borgefors, editor, *Proc. 9th Scandinavian Conference on Image Analysis*, pages 767–776, Uppsala, Sweden, June 1995. Swedish Society for Automated Image Processing.

T. Lindeberg and M. Li. Segmentation and classification of edges using minimum description length approximation and complementary junction cues. *Computer Vision and Image Understanding*, 67(1):88–98, 1997.

T. Lindeberg and B. ter Haar Romeny. Linear scale-space I: Basic theory. In B. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, Series in Mathematical Imaging and Vision, pages 1–41. Kluwer Academic Publishers, Dordrecht, Netherlands, 1994a.

T. Lindeberg and B. ter Haar Romeny. Linear scale-space II: Early visual operations. In B. ter Haar Romeny, editor, *Geometry-Driven Diffusion in Computer Vision*, Series in Mathematical Imaging and Vision, pages 43–77. Kluwer Academic Publishers, Dordrecht, Netherlands, 1994b.

B. D. Lukas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Image Understanding Workshop*, 1981.

S. G. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Trans. Information Theory*, 38(2):617–643, 1992.

S. G. Mallat and S. Zhong. Characterization of signals from multi-scale edges. *IEEE Trans. Pattern Analysis and Machine Intell.*, 14(7):710–723, 1992.

D. Marr. *Vision*. W.H. Freeman, New York, 1982.

W. Niessen and R. Maas. Optic flow and stereo. In J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors, *Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory*, Copenhagen, Denmark, May. 1996. Kluwer Academic Publishers.

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

P. Perona. Steerable-scalable kernels for edge detection and junction analysis. In *Proc. 2nd European Conf. on Computer Vision*, pages 3–18, Santa Margherita Ligure, Italy, May. 1992.

S. M. Pizer, C. A. Burbeck, J. M. Coggins, D. S. Fritsch, and B. S. Morse. Object shape before boundary shape: Scale-space medial axis. *J. of Mathematical Imaging and Vision*, 4:303–313, 1994.

A. R. Rao and B. G. Schunk. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, 53(2):157–185, Mar. 1991.

J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors. *Gaussian Scale-Space Theory: Proc. PhD School on Scale-Space Theory*. Series in Mathematical Imaging and Vision. Kluwer Academic Publishers, Copenhagen, Denmark, May 1996.

H. Voorhees and T. Poggio. Detecting textons and texture boundaries in natural images. In *Proc. 1st Int. Conf. on Computer Vision*, London, England, 1987.

K. Wiltschi, T. Lindeberg, and A. Pinz. Classification of carbide distributions using scale selection and directional distributions. Technical Report ISRN KTH/NA/P--97/10--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, 1997a. Shortened version in Proc. 4th International Conference on Image Processing, (Santa Barbara, California), vol. II, pp. 122–125, IEEE, Oct. 1997.

K. Wiltschi, A. Pinz, and T. Lindeberg. Classification of carbide distributions using scale selection and directional distributions. In *Proc. 4th International Conference on Image Processing*, volume II, pages 122–125, Santa Barbara, California, Oct. 1997b. IEEE.

Y. Yacoob and L. S. Davis. Estimating image motion using temporal multiscale models of flow and acceleration. In M. Shah and R. Jain, editors, *Motion-Based Recognition*. Kluwer Academic Publishers, 1997.

R. A. Young. 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.

R. A. Young. The Gaussian derivative model for spatial vision: I. Retinal mechanisms. *Spatial Vision*, 2:273–293, 1987.

A. L. Yuille and T. A. Poggio. Scaling theorems for zero-crossings. *IEEE Trans. Pattern Analysis and Machine Intell.*, 8:15–25, 1986.

W. Zhang and F. Bergholm. An extension of Marr's signature based edge classification and other methods for determination of diffuseness and height of edges, as well as line width. In H.-H. N. et. al., editor, *Proc. 4th Int. Conf. on Computer Vision*, pages 183–191, Berlin, Germany, May. 1993. IEEE Computer Society Press.

Index