Scale-Space Behaviour and Invariance Properties of Differential Singularities [*]

Tony Lindeberg

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

Abstract. This article describes how a certain way of expressing low-level feature detectors, in terms of singularities of differential expressions defined at multiple scales in scale-space, simplifies the analysis of the effect of smoothing. It is shown how such features can be related across scales, and generally valid expressions for drift velocities are derived with examples concerning edges, junctions, Laplacean zero-crossings, and blobs. A number of invariance properties are pointed out, and a particular representation defined from such singularities, the scale-space primal sketch, is treated in more detail.

Keywords: scale-space, drift velocity, feature detection, primal sketch, singularity, invariance.

1 Introduction

A common way of implementing low-level feature detectors in computer vision and image processing is by applying non-linear operations to smoothed input data. Examples of this are edge detection, junction detection, and blob detection. The pre-smoothing step can be motivated either heuristically by the need for noise suppression in real-world signals, or by the fact that image structures only exist as meaningful entities over certain ranges of scale. The latter argument is one of the main motivations for the development of the multi-scale representation known as *scale-space representation*, in which a given signal is subjected to smoothing by Gaussian kernels of successively increasing width.

This aim of this article is to show why a certain way of formulating such low-level feature detectors, in terms of singularities of differential expressions defined from the scale-space representation, is attractive from theoretical viewpoint. An overview of the scheme proposed is shown in Fig. 1. Any given signal is subjected

---

Singularities (features)

$\uparrow$

| *Nearest-neighbour* comparisons |

|

Differential geometric descriptors

$\uparrow$

| *Point* operations |

|

Smoothed derivatives: $L_{x^i y^j}$

$\uparrow$

| *Local* derivative operators: $\partial_{x^i y^j}$ |

|

Smoothed signals: L

$\uparrow$

| *Large support* smoothing operator: $G*$ |

|

Continuous input signal: f
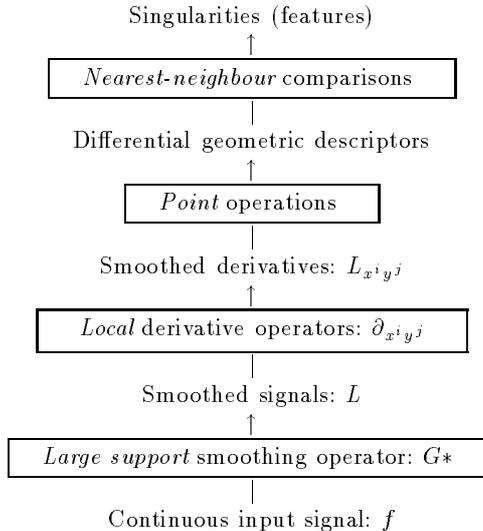
Fig. 1. Schematic overview of the different types of computations required for detecting features in terms of differential singularities at multiple scales.

to the following operations: (i) smoothing to a number of scales, (ii) derivative computations at each scale, (iii) combination of the derivatives at each scale into (non-linear) differential geometric entities, and (iv) detection of zero-crossings in these. It will be shown how the effect of the smoothing operation in this scheme can be analyzed by (i) showing how features defined in this way can be related across scales; a subject which can be referred to as the "deep structure of scale-space", and (ii) by deriving drift velocities for a large class of feature detectors. A number of invariance properties with respect to natural transformations of the spatial coordinates and the grey-level domain will also be listed.

Before starting it should be pointed out that this scheme is not presented as solely new. Some of the results presented in the paper are (at least partly) known, or have been touched upon before; see, e.g., Koenderink and van Doorn [8], who proposed the multi-scale N-jet representation, Florack *et al.* [5], who showed how a minimal set of differential invariants can be derived at any scale, or Lindeberg [11] who analysed the behaviour of scales of local extrema and related entities. The purpose with this presentation is rather to emphasize the role of the *singularities* in the scheme in Fig. 1, and to illustrate how they are attractive for the theoretical analysis of different feature detectors. For simplicity, the treatment is developed for two-dimensional signals. The approach, however, is valid in arbitrary dimensions.

The scale-space concept dealt with is the traditional diffusion-based scale-space for continuous signals developed by Witkin [16], Koenderink [7], and

Babaud *et al.* [1], which is given by the solution to

$$\partial_t L = \frac{1}{2}\nabla^2 L = \frac{1}{2}(L_{xx} + L_{yy}) \tag{1}$$

with initial condition $L(\cdot;\ 0) = f(\cdot)$. At any scale in this representation and at any point $P_0 = (x_0, y_0) \in \mathbb{R}^2$, denote by $\partial_{\bar{v}}$ the directional derivative in the gradient direction of L, and by $\partial_{\bar{u}}$ the derivative in the perpendicular direction. In terms of derivatives along the Cartesian coordinate directions it holds that

$$\partial_{\bar{u}} = \sin\phi\,\partial_x - \cos\phi\,\partial_y \ \ ; \ \ \ \partial_{\bar{v}} = \cos\phi\,\partial_x + \sin\phi\,\partial_y \ \ , \tag{2}$$

where $(\cos\phi, \sin\phi)$ is the normalized gradient direction of L at P_0.

2 Feature Detection from Singularities in Scale-Space

2.1 Examples of Differential Formulations of Feature Detectors

A natural way to define *edges* from a continuous grey-level image $L : \mathbb{R}^2 \to \mathbb{R}$ is as the union of the points for which the gradient magnitude assumes a maximum in the gradient direction. This method is usually referred to as "non-maximum suppression", (see e.g. Canny [4]). Assuming that the second and third order directional derivatives of L in the v-direction are not simultaneously zero, a necessary and sufficient condition for P_0 to be a gradient maximum in the gradient direction may be stated as:

$$\begin{cases} L_{\bar{v}\bar{v}} = 0 \ , \\ L_{\bar{v}\bar{v}\bar{v}} < 0 \ . \end{cases} \tag{3}$$

Since only the sign information is important, this condition can be restated as

$$\begin{cases} \tilde{L}_{\bar{v}\bar{v}} = \ L_{\bar{v}}^2 L_{\bar{v}\bar{v}} = \ L_x^2 L_{xx} + 2L_x L_y L_{xy} + L_y^2 L_{yy} = 0 \ , \\ \tilde{L}_{\bar{v}\bar{v}\bar{v}} = L_{\bar{v}}^3 L_{\bar{v}\bar{v}\bar{v}} = L_x^3 L_{xxx} + 3L_x^2 L_y L_{xxy} + 3L_x L_y^2 L_{xyy} + L_y^3 L_{yyy} < 0 \ . \end{cases} \tag{4}$$

An entity commonly used for *junction detection* is the curvature of level curves in intensity data (see e.g. Kitchen [6] or Koenderink and Richards [9]). In terms of directional derivatives it can be expressed as

$$\kappa = \frac{L_{\bar{u}\bar{u}}}{L_{\bar{v}}} \ \ . \tag{5}$$

In order to to give a stronger response near edges, the level curve curvature is usually multiplied by the gradient magnitude (see, e.g., Brunnström *et al.* [3])

$$|\tilde{\kappa}| = |L_{\bar{v}}^2 L_{\bar{u}\bar{u}}| = |L_y^2 L_{xx} - 2L_x L_y L_{xy} + L_x^2 L_{yy}| \ \ . \tag{6}$$

Assuming that the first- and second-order differentials of $\tilde{\kappa}$ are not simultaneously degenerate, a necessary and sufficient condition for a point P_0 to be a maximum in this rescaled level curve curvature is that:

$$\begin{cases} \partial_{\bar{u}}(\tilde{\kappa}) = 0 \ , \ \ \ \partial_{\bar{v}}(\tilde{\kappa}) = 0 \ , \\ \mathcal{H}(\tilde{\kappa}) = \tilde{\kappa}_{\mathcal{H}} = \tilde{\kappa}_{\bar{u}\bar{u}}\tilde{\kappa}_{\bar{v}\bar{v}} - \tilde{\kappa}_{\bar{u}\bar{v}}^2 > 0 \ , \\ \mathrm{sign}(\tilde{\kappa})\tilde{\kappa}_{\bar{u}\bar{u}} < 0 \ . \end{cases} \tag{7}$$

Zero-crossings of the Laplacean

$$\nabla^2 L = L_{\bar{u}\bar{u}} + L_{\bar{v}\bar{v}} = L_{xx} + L_{yy} = 0 \tag{8}$$

have been used for *stereo matching* (see, e.g., Marr [15]) and *blob detection* (see, e.g., Blostein and Ahuja [2]). Blob detection methods can also be formulated in terms of local extrema (see, e.g., Lindeberg and Eklundh [10]).

2.2 Invariance Properties of Differential Singularities

One of the main reasons why the formulation in terms of singularities is important is because these singularities do not depend on the actual numerical values of the differential geometric entities, but only on their *relative* relations. In this way, they will be less affected by scale-space smoothing, which is well-known to decrease the amplitude of the variations in a signal and its derivatives.

In fact, the differential entities used above are *invariant* to a number of primitive transformations of both the original and the smoothed grey-level signal; *translations*, *rotations*, and (uniform) *rescalings* in space as well as *affine intensity transformations*. (This set is similar but not equal to the set of transformations used by Florack *et al.* [5] to derive necessity results about differential invariants from intensity data; the main difference is that [5] considers invariance with respect to arbitrary *monotone intensity transformations*, while the differential singularities used here are invariant to uniform rescalings of the coordinate axes, i.e., *size changes*.)

To give a precise formulation of this, let $L_{\bar{u}^m \bar{v}^n} = L_{\bar{u}^\alpha}$ denote a mixed directional derivative of order $|\alpha| = m + n$, where $\alpha = (m, n)$, and let $\mathcal{D}$ be a (possibly non-linear) homogeneous differential expression of the form

$$\mathcal{D}L = \sum_{i=1}^{I} c_i \prod_{j=1}^{J} L_{\bar{u}^{\alpha_{ij}}} \ , \tag{9}$$

where $|\alpha_{ij}| > 0$ for all $i = [1..I]$ and $j = [1..J]$, and $\sum_{j=1}^{J} |\alpha_{ij}| = N$ for all $i \in [1..I]$. Moreover, let $\mathcal{S}_\mathcal{D}L$ denote the *singularity set* of this operator, i.e., $\mathcal{S}_\mathcal{D}L = \{(x;\ t) : \mathcal{D}L(x;\ t) = 0\}$, and let $\mathcal{G}$ be the Gaussian smoothing operator, i.e., $L = \mathcal{G}f$. Under these transformations of the spatial domain (represented by $x \in \mathbb{R}^2$) and the intensity domain (represented by either the unsmoothed f or the smoothed L) the singularity sets transform as follows:

Transformation	Definition	Invariance
translation	$(\mathcal{T}L)(x;\ t) = L(x + \Delta x;\ t)$	$\mathcal{S}_\mathcal{D}\,\mathcal{G}\,\mathcal{T}\,f = \mathcal{S}_\mathcal{D}\,\mathcal{T}\,\mathcal{G}\,f = \mathcal{T}\,\mathcal{S}_\mathcal{D}\,\mathcal{G}\,f$
rotation	$(\mathcal{R}L)(x;\ t) = L(Rx;\ t)$	$\mathcal{S}_\mathcal{D}\,\mathcal{G}\,\mathcal{R}\,f = \mathcal{S}_\mathcal{D}\,\mathcal{R}\,\mathcal{G}\,f = \mathcal{R}\,\mathcal{S}_\mathcal{D}\,\mathcal{G}\,f$
uniform scaling	$(\mathcal{U}L)(x;\ t) = L(sx;\ t)$	$\mathcal{S}_\mathcal{D}\,\mathcal{G}\,\mathcal{U}\,f = \mathcal{S}_\mathcal{D}\,\mathcal{U}\,\mathcal{G}\,f = \mathcal{U}\,\mathcal{S}_\mathcal{D}\,\mathcal{G}\,f$
affine intensity	$(\mathcal{A}L)(x;\ t) = aL(x;\ t) + b$	$\mathcal{S}_\mathcal{D}\,\mathcal{G}\,\mathcal{A}\,f = \mathcal{S}_\mathcal{D}\,\mathcal{A}\,\mathcal{G}\,f = \mathcal{S}_\mathcal{D}\,\mathcal{G}\,f$

Above, R is a rotation matrix, Δx is a vector ($\in \mathbb{R}^2$), while a, b and s are scalar constants. The definitions of the transformed singularity sets are as follows; $\mathcal{T}\mathcal{S}_\mathcal{D}L = \{(x;\ t) : \mathcal{D}L(x + \Delta x;\ t) = 0\}$, $\mathcal{R}\mathcal{S}_\mathcal{D}L = \{(x;\ t) : \mathcal{D}L(Rx;\ t) = 0\}$,

and $\mathcal{U}\mathcal{S}_{\mathcal{D}}L = \{(x;\ t) : \mathcal{D}L(sx;\ \underline{s^2 t}) = 0\}$. The commutative properties of $\mathcal{G}$ with $\mathcal{T}$, $\mathcal{R}$, and $\mathcal{A}$ are trivial consequences of the translational invariance, rotational invariance, and linearity of Gaussian smoothing. Under uniform rescalings $f'(x,y) = f(sx, sy)$, however, it holds that $L' = \mathcal{G}f'$ is related to L by $L'(x, y;\ t_0) = L(sx, sy;\ s^2 t_0)$, which means that $\mathcal{U}$ applied to a singularity set also affects the scale levels. The commutative properties of $\mathcal{S}_{\mathcal{D}}$ with $\mathcal{T}$, $\mathcal{U}$, and $\mathcal{A}$ follow from corresponding invariance properties of linear derivative operators combined with the homogeneity of (9), while the commutativity with respect to $\mathcal{R}$ follows from the rotational invariance of the directional derivatives $L_{\bar{u}^{\alpha_i j}}$.

3 Relating Differential Singularities at Different Scales

Consider a feature, which at any level of scale can be defined by

$$h(x, y;\ t) = 0 \tag{10}$$

for some function $h : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}^N$, where N is either 1 or 2. Using the implicit function theorem it is easy to analyze the dependence of (x, y) on t in the solution to (10). The results to be derived give estimates of the drift velocity of different features due to scale-space smoothing, and provides a theoretical basis for relating corresponding features at adjacent scales. It does hence enable well-defined linking and/or identification of features across scales.

3.1 Zero-Dimensional Entities (Points)

Assume first that N is equal to 2, that is, that $h(x, y;\ t) = (h_1(x, y;\ t), h_2(x, y;\ t))$ for some functions $h_1, h_2 : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}$. The derivative of the mapping h at a point $P_0 = (x_0, y_0;\ t_0)$ is

$$h'|_{P_0} = \left. \begin{pmatrix} \partial_x h_1 & \partial_y h_1 & \partial_t h_1 \\ \partial_x h_2 & \partial_y h_2 & \partial_t h_2 \end{pmatrix} \right|_{P_0} = \left. \left(\frac{\partial(h_1, h_2)}{\partial(x, y)} \quad \frac{\partial(h_1, h_2)}{\partial(t)} \right) \right|_{P_0}. \tag{11}$$

If $\partial(h_1, h_2)/\partial(x, y)$ is a non-singular matrix at P_0, then the solution (x, y) to $h(x, y;\ t_0) = 0$ will be an isolated point. Moreover, the implicit function theorem guarantees that there exists some local neighbourhood around P_0 where (x, y) can be expressed as a function of t. The derivative of that mapping $t \mapsto (x, y)$ is:

$$\left. \begin{pmatrix} \partial_t x \\ \partial_t y \end{pmatrix} \right|_{P_0} = - \left. \begin{pmatrix} \partial_x h_1 & \partial_y h_1 \\ \partial_x h_2 & \partial_y h_2 \end{pmatrix} \right|_{P_0}^{-1} \left. \begin{pmatrix} \partial_t h_1 \\ \partial_t h_2 \end{pmatrix} \right|_{P_0}. \tag{12}$$

If h is a function of the spatial derivatives of L only, which is the case, for example, for the feature extractors treated in Sect. 2.1, then the fact that spatial derivatives of L satisfy the diffusion equation $\partial_t = (\partial_{xx} + \partial_{yy})/2$, can be used for replacing derivatives with respect to t by derivatives with respect to x and y. Hence, closed form expression can be obtained containing only partial derivatives of L with respect to x and y. For example, the *junction candidates* given by (7)

satisfy $(\tilde{\kappa}_{\bar{u}}, \tilde{\kappa}_{\bar{v}}) = (0,0)$. In terms of directional derivatives, (12) can then be
written

$$\left. \begin{pmatrix} \partial_t u \\ \partial_t v \end{pmatrix} \right|_{P_0} = - \left. \begin{pmatrix} \tilde{\kappa}_{\bar{u}\bar{u}} & \tilde{\kappa}_{\bar{u}\bar{v}} \\ \tilde{\kappa}_{\bar{u}\bar{v}} & \tilde{\kappa}_{\bar{v}\bar{v}} \end{pmatrix}^{-1} \right|_{P_0} \left. \begin{pmatrix} \partial_t \tilde{\kappa}_{\bar{u}} \\ \partial_t \tilde{\kappa}_{\bar{v}} \end{pmatrix} \right|_{P_0} . \tag{13}$$

By differentiating the expressions for $\tilde{\kappa}_{\bar{u}}$ and $\tilde{\kappa}_{\bar{v}}$ with respect to t, by using the
fact that the spatial derivatives satisfy the diffusion equation, and by expressing
the result in directional derivatives in the u- and v-directions the following ex-
pressions can be obtained (the calculations have been done using Mathematica):

$$
\begin{aligned}
\tilde{\kappa}_{\bar{u}\bar{u}} &= L_{\bar{v}}^2 L_{\bar{u}\bar{u}\bar{u}\bar{u}} + 2L_{\bar{u}\bar{u}}(L_{\bar{u}\bar{u}}L_{\bar{v}\bar{v}} - L_{\bar{u}\bar{v}}^2) + 2L_{\bar{v}}(L_{\bar{u}\bar{v}}L_{\bar{u}\bar{u}\bar{u}} - L_{\bar{u}\bar{u}}L_{\bar{u}\bar{u}\bar{v}}), \\
\tilde{\kappa}_{\bar{u}\bar{v}} &= L_{\bar{v}}^2 L_{\bar{u}\bar{u}\bar{u}\bar{v}} + 2L_{\bar{u}\bar{v}}(L_{\bar{u}\bar{u}}L_{\bar{v}\bar{v}} - L_{\bar{u}\bar{v}}^2) + 2L_{\bar{v}}(L_{\bar{v}\bar{v}}L_{\bar{u}\bar{u}\bar{u}} - L_{\bar{u}\bar{v}}L_{\bar{u}\bar{u}\bar{v}}), \\
\tilde{\kappa}_{\bar{v}\bar{v}} &= L_{\bar{v}}^2 L_{\bar{u}\bar{u}\bar{v}\bar{v}} + 2L_{\bar{v}\bar{v}}(L_{\bar{u}\bar{u}}L_{\bar{v}\bar{v}} - L_{\bar{u}\bar{v}}^2) + 2L_{\bar{v}}(L_{\bar{u}\bar{u}}L_{\bar{v}\bar{v}\bar{v}} + 2L_{\bar{v}\bar{v}}L_{\bar{u}\bar{u}\bar{v}} - 3L_{\bar{u}\bar{v}}L_{\bar{u}\bar{v}\bar{v}}), \\
\partial_t \tilde{\kappa}_{\bar{u}} &= L_{\bar{v}}^2 (L_{\bar{u}\bar{u}\bar{u}\bar{u}\bar{u}} + L_{\bar{u}\bar{u}\bar{u}\bar{v}\bar{v}})/2 \\
&\quad + (L_{\bar{u}\bar{u}}L_{\bar{v}\bar{v}} - L_{\bar{u}\bar{v}}^2)(L_{\bar{u}\bar{u}\bar{u}} + L_{\bar{u}\bar{v}\bar{v}}) + L_{\bar{v}}(L_{\bar{u}\bar{u}\bar{u}}L_{\bar{v}\bar{v}\bar{v}} - L_{\bar{u}\bar{u}\bar{v}}L_{\bar{u}\bar{v}\bar{v}}), \\
\partial_t \tilde{\kappa}_{\bar{v}} &= L_{\bar{v}}^2 (L_{\bar{u}\bar{u}\bar{u}\bar{u}\bar{v}} + L_{\bar{u}\bar{u}\bar{v}\bar{v}\bar{v}})/2 + (L_{\bar{u}\bar{u}}L_{\bar{v}\bar{v}} - L_{\bar{u}\bar{v}}^2)(L_{\bar{u}\bar{u}\bar{v}} + L_{\bar{v}\bar{v}\bar{v}}) \\
&\quad + L_{\bar{v}}(L_{\bar{v}\bar{v}}(L_{\bar{u}\bar{u}\bar{u}\bar{u}} + L_{\bar{u}\bar{u}\bar{v}\bar{v}}) + L_{\bar{u}\bar{u}}(L_{\bar{v}\bar{v}\bar{v}\bar{v}} + L_{\bar{u}\bar{u}\bar{v}\bar{v}}) - 2L_{\bar{u}\bar{v}}(L_{\bar{u}\bar{u}\bar{u}\bar{v}} + L_{\bar{u}\bar{v}\bar{v}\bar{v}})) \\
&\quad + L_{\bar{v}}(L_{\bar{u}\bar{u}\bar{v}}(L_{\bar{u}\bar{u}\bar{u}} + L_{\bar{v}\bar{v}\bar{v}}) - L_{\bar{u}\bar{v}\bar{v}}(L_{\bar{u}\bar{u}\bar{u}} + L_{\bar{u}\bar{v}\bar{v}})).
\end{aligned}
$$

(These expressions simplify somewhat if we make use of $L_{\bar{u}\bar{u}\bar{u}}|_{P_0} = 0$, which
follows from $\tilde{\kappa}_{\bar{u}} = 0$.) Note that as long as the Hessian matrix of $\tilde{\kappa}$ is non-
degenerate, the sign of the $\tilde{\kappa}_{\mathcal{H}}$ and $\tilde{\kappa}_{\bar{u}\bar{u}}$ will be constant. This means that the
type of extremum will be the same. For *local extrema* of the grey-level landscape,
given by $(L_x, L_y) = (0,0)$, the expression for the drift velocity reduces to

$$r_t = -\frac{1}{2}(\mathcal{H}L)^{-1}\nabla^2(\nabla L) , \tag{14}$$

where $\mathcal{H}L$ denotes the Hessian matrix and ∇L the gradient vector. Observe that
regularity presents no problem, since L satisfies the diffusion equation, and for
$t > 0$ any solution to the diffusion equation is infinitely differentiable.

3.2 One-Dimensional Entities (Curves)

If N is equal to 1, then there will no longer be any unique correspondence between
points at adjacent scales. An ambiguity occurs, very similar to what is called the
aperture problem in motion analysis. Nevertheless, we can determine the drift
velocity in the normal direction of the curve. Given a function $h : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}$
consider the solution to $h(x,y; t) = 0$. Assume that $P_0 = (x_0, y_0; t_0)$ is a solution
to this equation and that the gradient of the mapping $(x,y) \mapsto h(x,y; t_0)$ is
non-zero. Then, in some neighbourhood around (x_0, y_0) the solution (x,y) to
$h(x,y; t_0) = 0$ defines a curve. Its normal at (x_0, y_0) is given by $(\cos\phi, \sin\phi) = (h_x, h_y)/(h_x^2 + h_y^2)^{1/2}$ at P_0. Consider the function $\tilde{h} : R \times \mathbb{R}_+ \to \mathbb{R}$ $\tilde{h}(s; t) = h(x_0 + s\cos\phi, y_0 + s\sin\phi; t)$, which has the derivative

$$\tilde{h}_s(0; t_0) = h_x(x_0, y_0; t_0)\cos\phi + h_y(x_0, y_0; t_0)\sin\phi = \left. \sqrt{h_x^2 + h_y^2} \right|_{P_0} . \tag{15}$$

Since this derivative is non-zero, we can apply the implicit function theorem.
It follows that there exists some neighbourhood around P_0 where $\tilde{h}(s; t) = 0$
defines s as a function of t. The derivative of this mapping is

$$\partial_t s \big|_{P_0} = -\left.\tilde{h}_s\right|_{P_0}^{-1} \left.\tilde{h}_t\right|_{P_0} = -\left.\frac{h_t}{\sqrt{h_x^2 + h_y^2}}\right|_{P_0} \quad . \tag{16}$$

As an example of this consider an *edge* given by non-maximum suppression

$$h = \alpha = L_x^2 L_{xx} + 2 L_x L_y L_{xy} + L_y^2 L_{yy} = 0 \quad . \tag{17}$$

By differentiating (17), by using the fact that the derivatives of L satisfy the diffusion equation, and by expressing the result in terms of the directional derivatives we get

$$
\begin{aligned}
\alpha &= L_{\bar{v}}^2 L_{\bar{v}\bar{v}} = 0, \\
\alpha_{\bar{u}} &= L_{\bar{v}}^2 L_{\bar{u}\bar{v}\bar{v}} + 2 L_{\bar{v}} L_{\bar{u}\bar{v}} L_{\bar{u}\bar{u}}, \\
\alpha_{\bar{v}} &= L_{\bar{v}}^2 L_{\bar{v}\bar{v}\bar{v}} + 2 L_{\bar{v}} L_{\bar{u}\bar{v}}^2, \\
\alpha_t &= L_{\bar{v}}^2 (L_{\bar{u}\bar{u}\bar{v}\bar{v}} + L_{\bar{v}\bar{v}\bar{v}\bar{v}})/2 + L_{\bar{v}} L_{\bar{u}\bar{v}}(L_{\bar{u}\bar{u}\bar{u}} + L_{\bar{u}\bar{v}\bar{v}}) \quad .
\end{aligned}
\tag{18}
$$

To summarize, the drift velocity in the normal direction of a *(curved) edge* in scale-space is (with $\alpha_{\bar{u}}$ and $\alpha_{\bar{u}}$ according to (18))

$$(\partial_t u, \partial_t v) = -\frac{L_{\bar{v}}(L_{\bar{u}\bar{u}\bar{v}\bar{v}} + L_{\bar{v}\bar{v}\bar{v}\bar{v}}) + 2 L_{\bar{u}\bar{v}}(L_{\bar{u}\bar{u}\bar{u}} + L_{\bar{u}\bar{v}\bar{v}})}{2((L_{\bar{v}} L_{\bar{u}\bar{v}\bar{v}} + 2 L_{\bar{u}\bar{v}} L_{\bar{u}\bar{u}})^2 + (L_{\bar{v}} L_{\bar{v}\bar{v}\bar{v}} + 2 L_{\bar{u}\bar{v}}^2)^2)}\left(\frac{\alpha_{\bar{u}}}{L_{\bar{v}}}, \frac{\alpha_{\bar{v}}}{L_{\bar{v}}}\right) \quad . \tag{19}$$

Unfortunately, this expression cannot be further simplified unless additional constraints are posed on L. For a *straight edge*, however, where all partial derivatives with respect to u are zero, it reduces to

$$(\partial_t u, \partial_t v) = -\frac{1}{2}\frac{L_{\bar{v}\bar{v}\bar{v}\bar{v}}}{L_{\bar{v}\bar{v}\bar{v}}}(0, 1) \tag{20}$$

(see also [11, 14]). For a curve given by the *zero-crossings of the Laplacean* we have

$$(\partial_t u, \partial_t v) = -\frac{\nabla^2(\nabla^2 L)}{2((\nabla^2 L_{\bar{u}})^2 + (\nabla^2 L_{\bar{v}})^2)}(\nabla^2 L_{\bar{u}}, \nabla^2 L_{\bar{v}}) \quad , \tag{21}$$

which also simplifies to (20) if all directional derivatives in the u-direction are set to zero. Similarly, for a *parabolic curve*, given by $\det(\mathcal{H}L) = L_{xx} L_{yy} - L_{xy}^2 = 0$, the drift velocity in the normal direction is

$$
\begin{aligned}
(\partial_t x, \partial_t y) = \quad &- \quad \frac{L_{yy} L_{xxxx} - 2 L_{xy} L_{xxxy} + (L_{xx} + L_{yy}) L_{xxyy} - 2 L_{xy} L_{xyyy} + L_{xx} L_{yyyy}}{2((L_{xx} L_{xyy} - 2 L_{xy} L_{xxy} + L_{yy} L_{xxx})^2 + (L_{xx} L_{yyy} - 2 L_{xy} L_{xyy} + L_{yy} L_{xxy})^2)} \\
&(L_{xx} L_{xyy} - 2 L_{xy} L_{xxy} + L_{yy} L_{xxx}, L_{xx} L_{yyy} - 2 L_{xy} L_{xyy} + L_{yy} L_{xxy}).
\end{aligned}
$$

This expression simplifies somewhat in a *pq*-coordinate system, with the *p*- and *q*-axes aligned to the principal axes of curvature so that the mixed second-order directional derivative $L_{\bar{p}\bar{q}}$ is zero.

4 Invariance Properties of the Scale-Space Primal Sketch

A particular type of representation constructed in this way is the scale-space primal sketch [10, 14]. It is a tree-like multi-scale representation aimed at *making explicit* blobs in scale-space as well as the relations between blobs at different scales. It is constructed by first defining one type of blobs, called *grey-level blobs*, at all levels of scale. The definition of this concept should be obvious from Fig. 2. Every local extremum is associated with a grey-level blob, whose extent
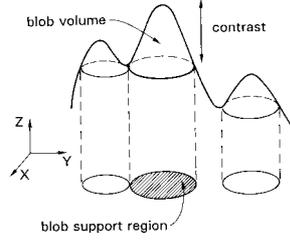
Fig. 2. Grey-level blob definition for bright blobs of a two-dimensional signal. In two dimensions a grey-level blob is generically given by a local extremum and the level curve through a specific saddle point, denoted *delimiting saddle point.*

is determined by the level curve through a specific saddle point, called *delimiting saddle point*. Formally, grey-level blobs can be defined by a water-shed analogy: Given any differentiable signal $f : \mathbb{R} \to \mathbb{R}$ consider any pair of maxima, A and B. They are connected by an infinite set of paths $P_{A,B}$. On each path, $p_{A,B}$, the grey-level function assumes a minimum. To reach another maximum from A, one must at least descend to the grey-level

$$z_{base}(A) = \sup_{B \in M} \; \sup_{p_{A,B} \in P_{A,B}} \; \inf_{(\xi,\eta) \in p_{A,B}} f(\xi,\eta) \; , \qquad (22)$$

where M is the set of all local maxima; $z_{base}(A)$ is the grey-level value of the delimiting saddle point associated with the local maximum A. The support region $D_{supp}(A)$ of the blob is the region

$$D_{supp}(A) = \{ \mathbf{r} \in \mathbb{R}^2 : \sup_{p_{A,r} \in P_{A,r}} \; \inf_{(\xi,\eta) \in p_{A,r}} f(\xi,\eta) \geq z_{base}(A) \} \; . \qquad (23)$$

Finally, the grey-level blob associated with A is the (three-dimensional) set

$$G_{blob}(A) = \{ (x,y,z) \in \mathbb{R}^2 \times \mathbb{R} : ((x,y) \in D_{supp}(A)) \wedge (z_{base}(A) \leq z \leq f(x,y)) \} \; . \qquad (24)$$

In general, to every grey-level blob existing at some level of scale there will correspond a similar blob both at a finer scale and a coarser scale. This notion can

be made precise by applying the implicit function theorem to the extremum and saddle point associated with each blob. Grey-level blobs along an *extremum path* across scales are linked as long as neither the extremum point nor its delimiting saddle point is involved in any bifurcation, that is, as long as the Hessian remains non-degenerate. The resulting (four-dimensional) objects, which have extent in both space (x, y), grey-level z, and scale t, are called *scale-space blobs* (see [10, 11, 14] for a detailed description). From the bifurcations between these objects, which can be of four types; annihilation, merge, split, and creation, a tree-like data structure can be constructed with the scale-space blobs as primitives and the bifurcations as arcs between them.

Earlier work has shown that this representation can be used for extracting blob-like image structures from grey-level images without any prior information about the contents of the image. A significance measure is postulated as the (four-dimensional) volume the scale-space blobs occupy in scale-space,

$$S_{vol,norm}(r) = \int_{t=t'_{min}}^{t'_{max}} V_{trans}(G_{volume}(r(t));\ t)\ d(\tau_{eff}(t))\ , \tag{25}$$

however normalized in order to enable uniform treatment of structures at different scales. $V_{trans}(G_{volume}(r(t));\ t)$ denotes a transformed volume of a grey-level blob along an extremum path r delimited by two scale values t'_{min} and t'_{max}, while $\tau_{eff} : \mathbb{R} \to \mathbb{R}$ is a transformation mapping the ordinary scale parameter into a transformed scale parameter called effective scale. For continuous signals τ_{eff} is given by $\tau_{eff}(t) = C_1 + C_2 \log t$ for some constants, C_1 and C_2 (see [12]).

Since the scale-space primal sketch is defined solely in terms of local extrema, level curves through saddle points, and bifurcations between critical points, it inherits the invariance properties listed in Sect. 2.2. This means that the topological relations in the tree-like data structure are preserved under *translations*, *rotations*, and (uniform) *rescalings* in space as well as *affine intensity transformations*. The relative ranking on significance obeys the following properties:

Invariance with respect to translations and rotations is trivial, since the scale-space representation and volumes are invariant to such operations. Concerning affine intensity transformations, it is obvious that the grey-level blob volumes are insensitive to the grey-level offset. Invariance with respect to linear stretching is achieved by dividing the measured grey-level volumes by the variation level of the input image in the transformation function V_{trans}. Because of the invariance of the scale-space primal sketch with respect to coordinate rescaling, the only way an extremum path is affected by this operation is by moving it so that the scale values t'_{min} and t'_{max} are multiplied by a constant factor. Clearly the logarithmic measure τ_{eff} is invariant to this, since it corresponds to a translation of the integration domain, which affects all scale-space blobs in the same way. The intention with the transformation function V_{trans} is that the integrand should also be well-behaved under this operation.

The scale-space primal sketch satisfies the following properties, which are essential for a low-level image representation: (i) it is based on the underlying *topology*, since it is defined from families of level curves; (ii) it is *hierarchical* in

the sense that the primitives are related through a tree-like data structure, and there is a natural ranking of events in order of significance; (iii) it is *local* in the sense that the primitives of the representation have finite support and only influence their nearest neighbours (in fact, it can be used for delimiting regions in space (and intervals in scale) for further processing); (iv) it is *continuous* in the sense that the topology of the overall representation is preserved as long as the relations in the underlying images remain the same; (v) it is *invariant* to transformations such as rotations and translations in space; (vi) it is *compatible with rescalings* of both the spatial coordinates and the grey-level intensity. For discrete signals, the rotations must be multiples of 90 degrees, the translations must be pixel-wise, and the spatial rescaling factor must be an integer in order to preserve the invariance.

5 Summary and Discussion

It has been shown how the formulation of feature detectors in terms of differential singularities makes it theoretically simple to analyze their behaviour over scales. Even though further work may be needed before implementing the drift velocity estimates derived for the different feature detectors, these expressions completely describe the theoretical evolution properties of such non-linear combinations defined in scale-space. The discretization of the scheme in Fig. 1 is treated in [13, 14], where experimental results are also presented.

References

1. Babaud, J., Witkin, A.P., Baudin, M., Duda, R.O. (1986). Uniqueness of the Gaussian kernel for scale-space filtering, IEEE Trans. Patt. Anal. Mach. Intell. 8 (1), pp. 26–33.
2. Blostein D., Ahuja N. (1987). Representation and three-dimensional interpretation of image texture: An integrated approach. In: Proc. 1st Int. Conf. Comp. Vision, London, England, Jun. 8-11, pp. 444–449.
3. Brunnström K., Lindeberg T.P, Eklundh J.-O. (1992). Active detection and classification of junctions by foveation with a head-eye system guided by the scale-space primal sketch. In: Proc. 2nd European Conf. Comp. Vision, Santa Margherita Ligure, Italy, pp. 701–709.
4. Canny J. (1986). A computational approach to edge detection, IEEE Trans. Patt. Anal. Machine Intell. 8 (6), pp. 679–698.
5. Florack L.M.J., ter Haar Romeny B.M., Koenderink J.J., Viergever M.A. (1991). General intensity transformations and second order invariants, Proc. 7th Scand. Conf. Image Analysis, Aalborg, Denmark, Aug 13-16, pp. 338–345.
6. Kitchen, L., Rosenfeld, R., (1982). Gray-level corner detection, Patt. Recogn. Lett. 1 (2), pp. 95–102.
7. Koenderink J.J. (1984). The structure of images, Biol. Cyb. 50, pp. 363–370.
8. Koenderink J.J., van Doorn A.J. (1987). Representation of local geometry in the visual system, Biol. Cyb. 55, pp. 367–375.
9. Koenderink J.J., Richards W. (1988). Two-dimensional curvature operators, J. Opt. Soc. Am. 5 (7), pp. 1136–1141.

10. Lindeberg T.P., Eklundh J.-O. (1992). The scale-space primal sketch: Construction and experiments, Image and Vision Comp, 10, pp. 3–18.
11. Lindeberg T.P. (1992). Scale-space behaviour of local extrema and blobs, J. Math. Imaging Vision 1, pp65-99.
12. Lindeberg T.P. (1993). Effective scale: A natural unit for measuring scale-space lifetime, IEEE Trans. Pattern Anal. Machine Intell., in press.
13. Lindeberg, T.P. (1993). Discrete derivative approximations with scale-space properties, J. Math. Imaging and Vision, in press.
14. Lindeberg, T.P. (1993). Scale-Space Theory in Early Vision, Kluwer Academic Publishers, Boston, to appear.
15. Marr D. (1982). Vision, Freeman, San Francisco.
16. Witkin, A.P. (1983). Scale-space filtering, In: Proc. 8th Int. Joint Conf. Art. Intell., Karlsruhe, Germany, Aug. 8-12, pp. 1019–1022.