JUNCTION DETECTION WITH AUTOMATIC SELECTION OF DETECTION SCALES AND LOCALIZATION SCALES

Tony Lindeberg

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

ABSTRACT

The subject of scale selection is essential to many aspects of multi-scale and multi-resolution processing of image data. This article shows how a general heuristic principle for scale selection can be applied to the problem of detecting and localizing junctions. In a first uncommitted processing step initial hypotheses about interesting scale levels (and regions of interest) are generated from scales where normalized differential invariants assume maxima over scales (and space). Then, based on this scale (and region) information, a more refined processing stage is invoked tuned to the task at hand. The resulting method is the first junction detector with automatic scale selection.

Whereas this article deals with the specific problem of junction detection, the underlying ideas apply also to other types of differential feature detectors, such as blob detectors, edge detectors, and ridge detectors.

1. INTRODUCTION

A basic problem when extracting information from measured data, such as images, originates from the fact that objects in the world, and hence image structures, exist as meaningful entities only over certain ranges of scale. In essence, it manifests itself as follows:

To derive any information from data, it is necessary to interact with it using operators. The type of information that can be obtained by this is to a large extent determined by the relationship between the size of the actual structures in the data and the size of the probes. In contrast to certain ideal mathematical entities, such as "point" or "line", which appear in a qualitatively similar way independent of the scale of observation, a description of real world data may vary substantially depending on the scale at which it is computed. This problem is especially acute for a vision system with the task to analyse unknown scenes about which no or very little *a priori* information is available.

A general framework for analysing structures at different scales is provided by *scale-space theory* (Witkin 1983; Koenderink 1984; Yuille and Poggio 1986; Florack *et al.* 1992; Lindeberg 1994a, 1994b). It is based on the general idea that in situations when no information is available about appropriate scales for analysis, the only reasonable approach is to consider descriptions at all scales. The essence of the results obtained from this theory is that if one assumes that the first stages of visual processing should be as uncommitted as possible and have no particular bias, then convolution with Gaussian kernels and their derivatives is singled out as a canonical class of low-level operators. The output from these operators can then be used as a basis for expressing a large number of early visual operations at multiple scales, such as feature detection, matching, and computation of shape cues.

This article deals with the problem of *selecting* interesting structures and appropriate scales from the scale-space representation. One approach concerning blob-like image structures was developed in (Lindeberg 1993a) based on the idea of selecting scale levels where a certain blob measure assumes maxima over scales. Here, a generalization of that idea (Lindeberg 1993b) is applied to the problem of detecting junctions. This results in a genuine two-stage approach with detection at coarse scales followed by localization to finer scales. Methods are presented for selecting both these scale levels automatically.

2. METHOD FOR SCALE SELECTION

The scale-space representation of an N-dimensional continuous signal $f : \mathbb{R}^N \to \mathbb{R}$ is the one-parameter family of functions $L : \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$

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

obtained by convolving f with the one-parameter family of Gaussian kernels

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

where $x = (x_1, ..., x_N)^T \in \mathbb{R}^N$ and $t \in \mathbb{R}_+$ is the scale parameter. A well-known property of this representation is that the amplitude of spatial derivatives

$$L_{x^\alpha}(\cdot;\ t) = \partial_{x^\alpha} L(\cdot;\ t) = \partial_{x_1^{\alpha_1}} \ldots \partial_{x_N^{\alpha_N}} L(\cdot;\ t) \quad (3)$$

in general *decrease with scale.* This can be seen, for example, for a one-dimensional sine wave, $f(x) = \sin \omega_0 x$, for which the scale-space representation is $L(x;\ t) = e^{-\omega_0^2 t/2} \sin \omega_0 x$, and the amplitude of any spatial derivative decreases exponentially with scale

$$L_{x^m, max}(t) = \omega_0^m\, e^{-\omega_0^2 t/2}. \qquad (4)$$

A natural way to reparametrize the spatial coordinates in scale-space is in terms of *normalized* (dimensionless) *coordinates*, $\xi = x/\sigma = x/\sqrt{t}$, (Florack *et al.* 1992). The corresponding *normalized derivative operator* is

$$\partial_\xi = \sqrt{t}\, \partial_x. \qquad (5)$$

For the sinusoidal signal the amplitude of any normalized derivative is

$$L_{\xi^m, max}(t) = t^{m/2}\, \omega_0^m\, e^{-\omega_0^2 t/2}, \qquad (6)$$

i.e., it first increases and then decreases. Introducing $\lambda_0 = 2\pi/\omega_0$ shows that the σ-value ($\sigma = \sqrt{t}$) for which $L_{\xi^m, max}(t)$ assumes its maximum over scales is *proportional* to the wavelength, λ_0, of the signal:

$$\sigma_{max, L_{\xi^m}} = \frac{\sqrt{m}}{2\pi} \lambda_0. \qquad (7)$$

Note that the maximum value $L_{\xi^m, max}(t_{max, L_{\xi^m}}) = m^{m/2} e^{-m/2}$ is *independent* of ω_0. Hence, all frequencies are treated in a similar manner. Observe the similarity to a local Fourier transform, although this method allows for local measurements of frequency contents without explicit setting of window size.
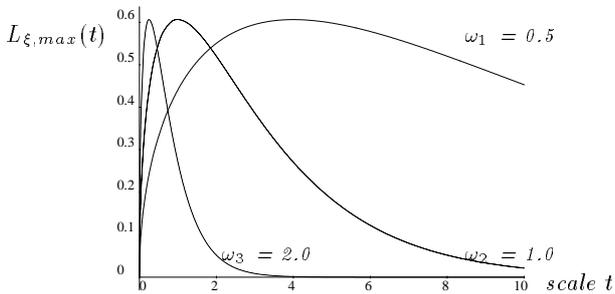

Figure 1: The amplitude of first order *normalized derivatives* as function of scale for sinusoidal input signals of different frequencies ($\omega_1 = 0.5$, $\omega_2 = 1.0$ and $\omega_3 = 2.0$).

2.1. Generalization: Homogeneous polynomial combinations of normalized derivatives

Whereas we have so far been concerned with a specific signal, it can be shown (Lindeberg 1994c) that for a large class of differential invariants (which can be used for expressing different types of feature detectors) it holds that such maxima over scales have a nice behaviour under rescalings of the input signal. If a normalized differential invariant $\mathcal{D}_{norm} L$ assumes a maximum over scales at a certain point $(x_0;\ t_0)$ in scale-space, then if a rescaled signal f' is defined by $f'(sx) = f(x)$, a scale-space maximum in the corresponding normalized differential entity $\mathcal{D}_{norm} L'$ is assumed at $(sx_0;\ s^2 t_0)$.

This means that if a differential invariant can be found that assumes a maximum over scales for a desired class of signals, this maximum will follow any size variations in the image data, and can be used for size measurements. Here, we shall apply this idea to the problem of junction detection, by selecting junctions candidates and scales for treating those from points where a (specific) normalized differential invariant simultaneously assumes maxima with respect to scale and space. Such points are called *scale-space maxima.*

This is a natural generalization of the common approach of taking a spatial maximum of an operator response as the estimate of the location of a feature.

3. SCALE SELECTION IN JUNCTION DETECTION

It is well-known that junctions provide important cues to three-dimensional structure. For example, T-junctions generically indicate interposition and hence depth discontinuities. 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 a polynomial expression,

$$\tilde{\kappa} = L_{x_2}^2 L_{x_1 x_1} - 2 L_{x_1} L_{x_2} L_{x_1 x_2} + L_{x_1}^2 L_{x_2 x_2}. \quad (8)$$

Fig. 2 shows the result of detecting scale-space maxima in the corresponding normalized entity $\tilde{\kappa}_{norm} = t^2 \tilde{\kappa}$ (obtained by replacing all derivative operators ∂_{x_i} by $\sqrt{t}\, \partial_{x_i}$). Each scale-space maximum is illustrated by a circle with the size (area) equal to the detection scale (measured as $t = \sigma^2$). Observe that a set of junction candidates is generated with reasonable interpretation in the scene. Moreover, the circles give natural regions of interest around the candidate junctions.
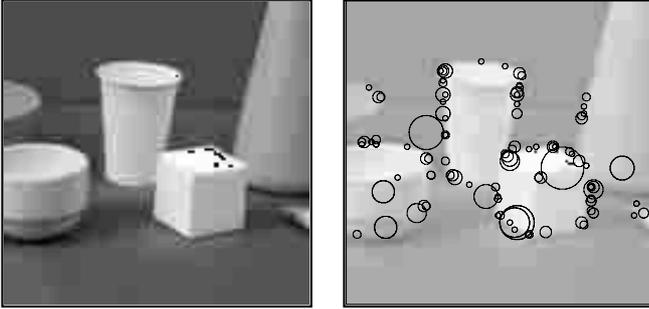
Figure 2: Junction candidates detected from an indoor scene by selecting the 100 scale-space maxima having the strongest maximum normalized response.
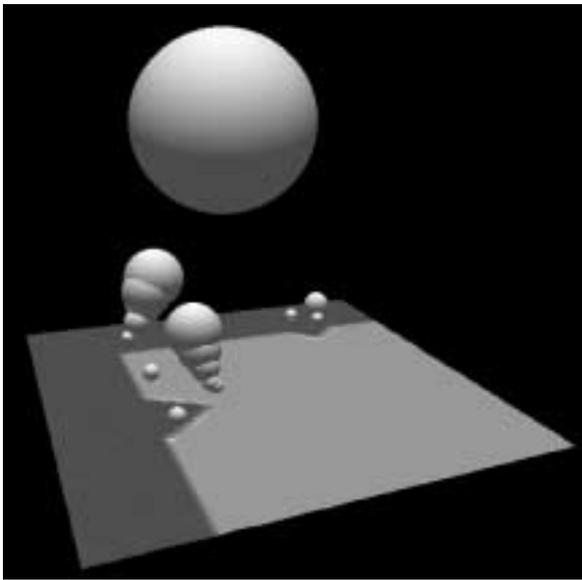

Figure 3: Corresponding results for a synthetic image with junctions of different spatial extent. (Here, each junction candidate is illustrated by a sphere centered at the scale-space maximum of $\tilde{\kappa}_{norm}$ with the radius proportional to the detection scale.)

3.1. Properties of the scale selection method

It can be shown that for a *diffuse junction model* defined as the product of two error functions along the coordinate directions $(f(x_1, x_2) = \Phi(x_1;\ t_0)\,\Phi(x_2;\ t_0)$ where $\Phi(x_i;\ t_0) = \int_{x'=-\infty}^{x_i} g(x';\ t_0)\,dx')$ the magnitude of $\tilde{\kappa}_{norm}$ *increases* monotonically with scale.

On the other hand, for a *non-symmetric Gaussian blob*, $(f(x_1, x_2) = g(x_1;\ t_1)\,g(x_2;\ t_2))$ the magnitude of $\tilde{\kappa}_{norm}$ *decreases* with scale.

Hence, $\tilde{\kappa}_{norm}$ can be expected to increase with scales when an infinite extent junction model constitutes a

reasonable approximation, and to decrease with scales when so much smoothing is applied that the shape distortions are substantial. Therefore, *selecting scale levels (and spatial points) from scale-space maxima of $\tilde{\kappa}_{norm}$ can be expected to give rise to larger scale values for junctions having larger spatial extent.*

These effects are illustrated in fig. 4, which shows *scale-space signatures* of $\tilde{\kappa}_{norm}$ for junctions of different spatial extent (obtained by tracking spatial maxima in $\tilde{\kappa}_{norm}$ across scales and registering the value of $\tilde{\kappa}_{norm}$ at each scale). The peak is assumed at coarser scales for the junction having larger spatial extent. (This property is also evident in fig. 3.)
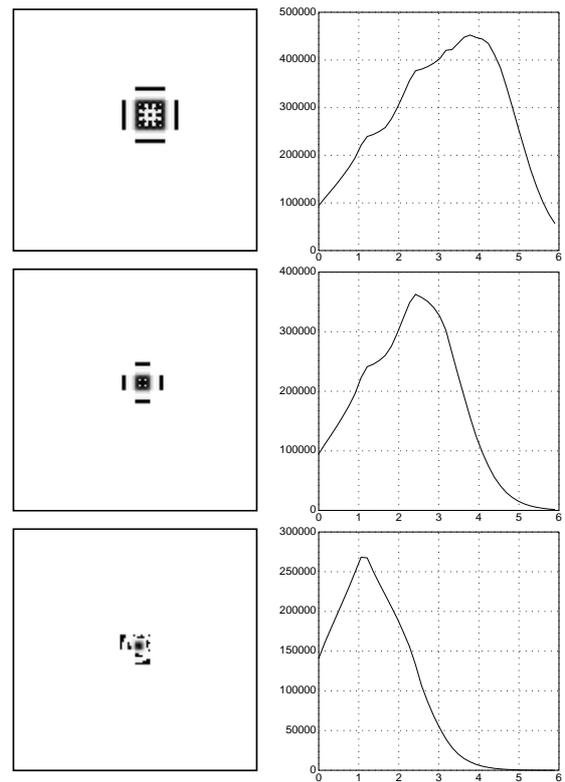

Figure 4: Scale-space signatures of $\tilde{\kappa}_{norm}$ for diffuse L-junctions ($t_0 = 1.0$) of different spatial extent. (Horizontal axis: Effective scale $\approx \log t$.)

4. SELECTION OF LOCALIZATION SCALE

Whereas the junction detection method presented so far is conceptually clean, it can certainly lead to poor localization, since shape distortions may be substantial at coarse scales in scale-space. A straightforward way to improve the location estimate is by determining the point x that minimizes the (perpendicular) distance to all edge tangents in a neighbourhood of the junction

candidate x_0. By defining these edge tangents with the gradient vectors as normals and by weighting each distance by the pointwise gradient magnitude, this problem can be given a standard least squares formulation (Förstner and Gülch 1987),

$$\min_{x \in \mathbb{R}^2} x^T A\, x - 2\, x^T b + c \quad \Longleftrightarrow \quad A\, x = b, \quad (9)$$

where $x = (x_1, x_2)^T$, w_{x_0} is a window function, and A, b, and c are entities determined by the local statistics of the gradient directions in a neighbourhood of x_0,

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

$$b = \int_{x' \in \mathbb{R}^2} (\nabla L)(x')\, (\nabla L)^T(x')\, x'\, w_{x_0}(x')\, dx',$$

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

4.1. Window function and localization scale.

To select the window function in (9) a Gaussian kernel with scale value equal to the detection scale is proposed as the first uncommitted choice. To determine the *localization scale* for computing the gradient vectors ∇L, a new heuristic principle is proposed that the location estimate should be computed at the scale that minimizes the *normalized residual*,

$$\tilde{d}_{min} = \min_{x \in \mathbb{R}^2} \frac{x^T A\, x - 2\, x^T b + c}{\text{trace}\, A} = \frac{c - b^T A^{-1} b}{\text{trace}\, A}.$$

The motivation for this choice is that at very fine scales, where a large amount of noise and other fine-scale structures can be expected to be present, the first-order derivative operator will mainly respond to such structures. Hence, the gradient directions can be expected to be roughly randomly distributed, and the normalized residual will in general be large. When the amount of smoothing is increased, the fine-scale structures will be suppressed, and the locally computed gradient directions will be better aligned to the underlying corner structure. On the other hand, if too much smoothing is applied, the shape distortions due to scale-space smoothing will be dominant, and the residual will increase. Hence, selecting the minimum gives a natural trade-off between these two effects.

Fig. 5 shows these effects for a synthetic T-junction. Note that coarser scale levels are selected when the noise level is increased. (Observe that for an ideal step junction the selected localization scale will always be zero.) Numerical values of the selected scale $t_{d_{min}}$ and the minimum residual $\tilde{d}_{min}$ are given in table 1.
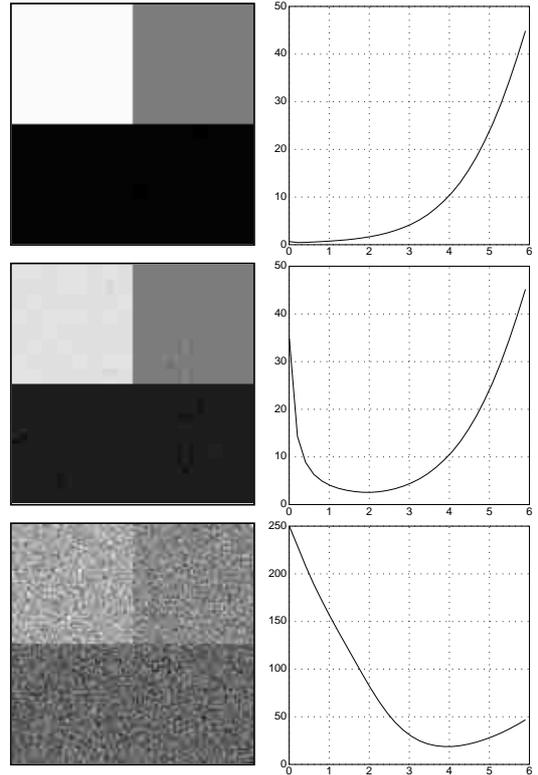

Figure 5: Scale-space signatures of the normalized residual $\tilde{d}_{min}$ at a synthetic sharp T-junction ($t_0 = 0.0$) at noise levels 1 %, 10 %, and 100 % (from top to bottom).

sharp T-junction (90°, $t_0 = 0.0$)			
noise level	$t_{d_{min}}$	$\tilde{d}_{min}$	error
0.0	0.0	0.2	0.05
1.0	0.3	0.4	0.07
3.0	0.8	0.7	0.12
10.0	2.0	1.7	0.23
30.0	5.0	4.2	0.56
100.0	13.0	9.7	1.34

Table 1: The scale levels $t_{d_{min}}$ where the minima over scales in $\tilde{d}_{min}$ are assumed for the T-junction in figure 5. (These data are median values of 21 experiments.) The last column gives the distance (measured in units of pixel lengths) between the estimated and the true position.

5. COMPOSED TWO-STAGE METHOD

To summarize, the composed scheme is based on the following heuristic principles:

1. *Detection.* In absence of other evidence, detect junction candidates at the scales (and spatial points) where the normalized rescaled level curve curvature assumes maxima over scale (and space).

2. *Localization.* In absence of further information, compute localization estimates at the scales that minimize the normalized residual $\tilde{d}_{min}$ over scale. The new localization estimate is $x = A^{-1}b$.

3. *Iterations.* Optionally, repeat these steps until the increment is sufficiently small.

The left image in fig. 6 shows the result of applying this procedure (with five iterations in the last step), and graphically illustrating each junction candidate by a circle with the area equal to the localization scale. The right image shows the result of suppressing those points for which the procedure did not converge after five iterations, and also suppressing overlapping blobs. Fig. 7 gives the result of applying the composed procedure to another image.
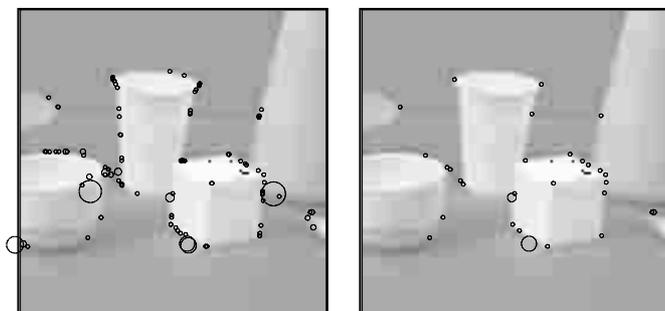

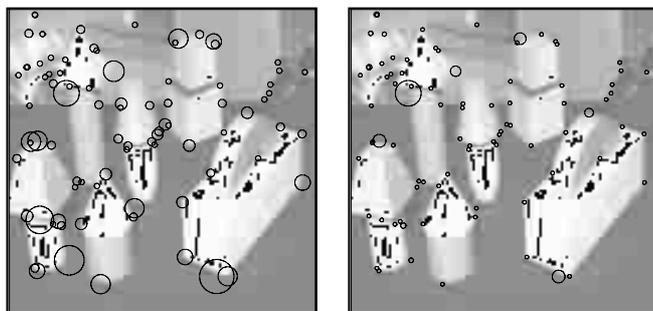**Figure 6:** Improved estimates for the candidates in fig. 2.

Figure 7: Composed junction detection: In the left image, the size of the circles reflect the detection scales, and in the right image the size reflects the localization scales.

6. SUMMARY AND DISCUSSION

The subject of scale selection is essential to many problems in computer vision and automated image analysis. This article has shown how this general problem can be approached in junction detection and junction localization by studying the evolution properties over scales of certain non-linear combinations of normalized Gaussian derivatives. The presented method, which is the first junction detector with automatic scale selection, is basically free from tuning parameters. The only essential parameter is the number of junction candidates to be treated in the initial step.

Further experiments (Lindeberg 1994c) demonstrate that for junctions with sufficiently large opening angles true subpixel accuracy is obtained even in the presence of moderately high noise levels (typically 10–100 % added Gaussian noise related to the contrast).

Of course, the task of selecting "the best scale" for handling real-world image data (about which usually no or very little *a priori* information is available) is intractable if treated as a pure mathematical problem. Therefore, the proposed heuristic principles should not be interpreted as "optimal solutions", but rather as systematic methods for generating initial hypotheses in situations where no or very little information is available about what can be expected to be in the scene.

7. REFERENCES

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

M.A. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centers of circular features. In *ISPRS*, 1987.

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

J. J. Koenderink and W. Richards. Two-dimensional curvature operators. *JOSA*, 5:7:1136–1141, 1988.

J. J. Koenderink. The structure of images. *Biol. Cyb.*, 50:363–370, 1984.

T. Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *IJCV*, 11:283–318, 1993a.

T. Lindeberg. On scale selection for differential operators. In *8th SCIA*, pages 857–866, 1993b.

T. Lindeberg. *Scale-Space Theory in Computer Vision.* Kluwer Academic Publishers, 1994a.

T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *J. Appl. Stat.*, 21(2):223–261, 1994b. (In press).

T. Lindeberg. Scale selection for differential operators. ISRN KTH/NA/P--94/03--SE, Royal Inst. Tech., 1994c.

A. P. Witkin. Scale-space filtering. In *8th IJCAI*, pages 1019–1022, 1983.

A. L. Yuille and T. A. Poggio. Scaling theorems for zero-crossings. *IEEE-PAMI*, 8:15–25, 1986.