Automatic Scale Selection as a Pre-Processing Stage for Interpreting the Visual World[1]

Tony Lindeberg

Department of Numerical Analysis and Computing Science
KTH, S-100 44 Stockholm, Sweden
tony@nada.kth.se, http://www.nada.kth.se/~tony

Abstract:

This paper reviews a systematic methodology for formulating mechanisms for automatic scale selection when performing feature detection. An important property of the proposed approach is that the notion of scale is included already in the definition of image features.

1 Introduction

Computer vision algorithms for interpreting image data usually involve a feature detection step. The need for performing early feature detection is usually motivated by the desire of condensing the rich intensity pattern to a more compact representation for further processing. If a proper abstraction of shape primitives can be computed, certain invariance properties can also be expected with respect to changes in view direction and illumination variations.

The earliest works in this direction were concerned with the edge detection [22, 21]. While edge detection may at first to be a rather simple task, it was empirically observed that it can be very hard to extract edge descriptors reliably. Usually, this was explained as a noise sensitivity that could be reduced by pre-smoothing the image data before applying the edge detector [25]. Later, a deeper understanding was developed that these difficulties originate from the more fundamental aspect of image structure, namely that real-world objects (in contrast to idealized mathematical entities such as points and lines) usually consist of different types of structures at different scales [27, 11]. Motivated by the multi-scale nature of real-world images, multi-scale representations such as pyramids [4] and scale-space representation [27, 11, 16] were constructed. Theories were also formed concerning what types of image features should be extracted from any scale level in a multi-scale representation [13, 8, 16, 7].

The most common way of applying multi-scale representations in practice has been by selecting one or a few scale levels in advance, and then extracting image features at each scale level more or less independently. This approach can be sufficient under simplified conditions, where only a few natural scale levels are involved and provided that the image features a stable over large ranges of scales. Typically, this is the case when extracting edges of man-made objects viewed under controlled imaging conditions. In other cases, however, there may be a need for adapting scale levels individually to each image feature, or even to adapt the scale levels along an extended image feature, such as a connected edge. Typically, this occurs when detecting ridges (which turn out to be much more scale sensitive than edges) and when applying an edge detector to a diffuse edge for which the degree of diffuseness varies along the edge.

To handle these effects in general cases, we argue that it is natural to complement feature detection modules by explicit mechanisms for automatic scale selection, so as to automatically adapt the scale levels to the image features under study. The purpose of this article is to present such a framework for automatic scale selection, which is generally applicable to a rich variety of image features, and has been successfully tested by integration with other visual modules. For references to the original sources, see [18, 17, 19] and the references therein.

An attractive property of the proposed scale selection mechanism is that in addition to automatic tuning of the scale parameter, it induces the computation of natural abstractions (groupings) of image shape. In this respect, the proposed methodology constitutes a natural pre-processing stage for subsequent interpretation of visual scenes.

2 The need for a scale-selection mechanism for feature detection

To demonstrate the need for an automatic scale selection mechanism, let us consider the problems of detecting edges and ridges, respectively, from image data. Figure 1 shows two images, from which scale-space representations have been computed by convolution with Gaussian kernels, *i.e.* given an image $f: \mathbb{R}^D \to \mathbb{R}$, its scale-space representation $L: \mathbb{R}^D \times \mathbb{R}_+ \to \mathbb{R}$ is

$$L(x;\ t) = \int_{\xi \in \mathbb{R}^N} f(x - \xi)\, g(\xi)\, d\xi. \tag{1}$$

where $g: \mathbb{R}^N \times \mathbb{R}_+ \to \mathbb{R}$ denotes the Gaussian kernel

$$g(x;\ t) = \frac{1}{(2\pi\sigma^2)^{D/2}}\, e^{-(x_1^2 + \cdots + x_D^2)/2t} \tag{2}$$

and the variance t of this kernel is referred to as the scale parameter.

Edge detection. At each scale level, edges are defined from points at which the gradient magnitude assumes a local maximum in the gradient direction [5, 15]. In terms of local

directional derivatives, where ∂_v denotes a directional derivative in the gradient direction, this edge definition can be written

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

Such edges at three scales are shown in the left column in figure 1. As can be seen, sharp edge structures corresponding to object boundaries give rise to edge curves at both fine and coarse scales. At fine scales, the localization of object edges is better, while the number of spurious edge responses is larger. Coarser scales are on the other hand necessary to capture the shadow edge, while the localization of $e.g.$ the finger tip is poor at coarse scales.
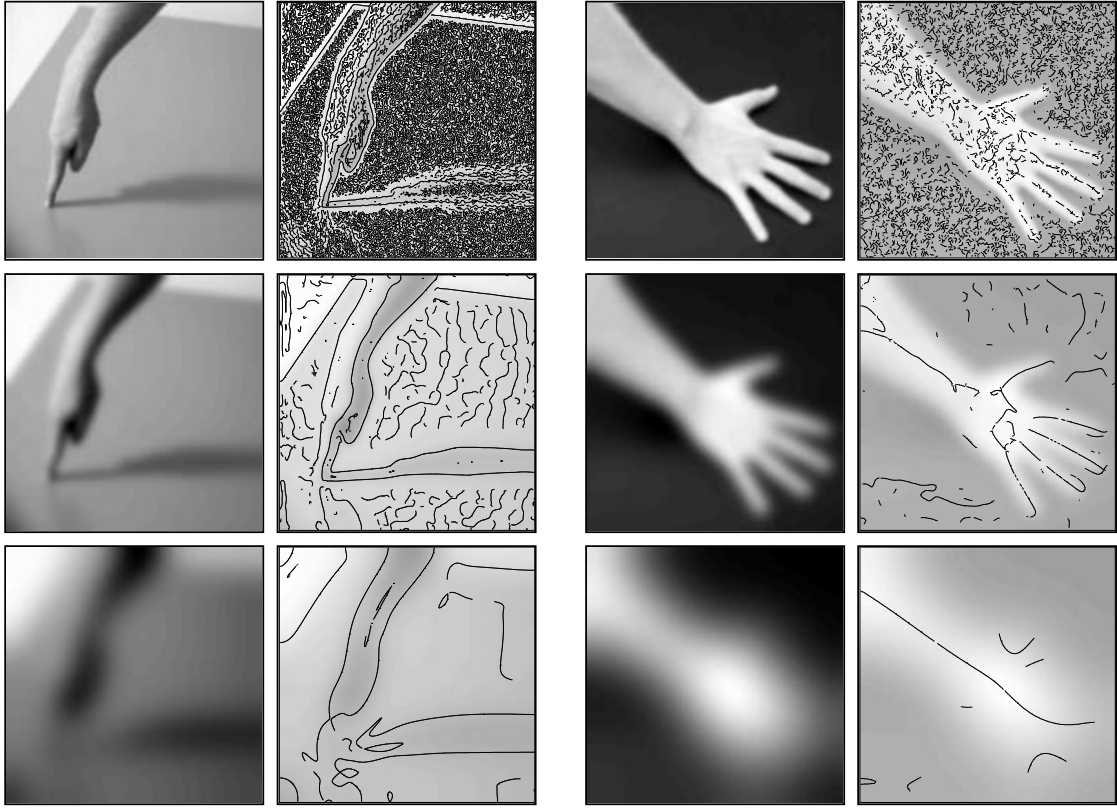

Figure 1: Edges and bright ridges detected at scale levels $t = 1.0$, 16.0 **and** 256.0, **respectively.**

Ridge detection. The right column in figure 1 show corresponding results of multi-scale ridge extraction. A (bright) ridge point is defined as a point where the intensity assumes a local maximum in the main eigendirection of the Hessian matrix [9, 14]. In terms of local (p, q)-coordinates with the mixed directional derivative $L_{pq} = 0$, this ridge definition can be written

$$\begin{cases} L_p = 0, \\ L_{pp} < 0, \\ |L_{pp}| \geq |L_{qq}|, \end{cases} \quad \text{or} \quad \begin{cases} L_q = 0, \\ L_{qq} < 0, \\ |L_{qq}| \geq |L_{pp}|, \end{cases} \tag{4}$$

while in terms of a local (u, v)-system with the v-direction parallel to the gradient direction and the u-direction perpendicular, the ridge definition assumes the form

$$\begin{cases} L_{uv} & = L_x L_y \left(L_{xx} - L_{yy}\right) - (L_x^2 - L_y^2)L_{xy} & = 0, \\ L_{uu}^2 - L_{vv}^2 & = \left(L_y^2 - L_x^2\right)\left(L_{xx} - L_{yy}\right) - 4L_x L_y L_{xy} & > 0. \end{cases} \tag{5}$$

As can be seen, the types of ridge curves that are obtained are strongly scale dependent. At fine scales, the ridge detector mainly responds to spurious noise structures. Then, it gives rise to ridge curves corresponding to the fingers at $t = 16$, and a ridge curve corresponding to the arm as a whole at $t = 256$. Notably, these ridge descriptors are much more sensitive to the choice of scale levels than the edge features in figure 1(a). In particular, no single scale level is appropriate for describing the dominant ridge structures in this image.

3   Proposed scale selection mechanism

The experimental results in figure 1 emphasize the need for adapting the scale levels for feature detection to the local image structures. How should such an adaptation be performed without *a priori* information about what image information is important? The subject of this section is to give an intuitive motivation of how *size estimation* can be performed, by studying the *evolution properties over scales of scale-normalized derivatives*. The basic idea is as follows: At any scale level, we define a normalized derivative operator by multiplying each spatial derivative operator ∂_x by the scale parameter t raised to a (so far free) parameter $\gamma/2$:

$$\partial_\xi = t^{\gamma/2} \, \partial_x \tag{6}$$

Then, we propose that *automatic scale selection can be performed by detecting the scales at which γ-normalized differential entities assume local maxima with respect to scale.* Intuitively, this approach corresponds to selecting the scales at which the operator response is as strongest.

Local frequency estimation I: For a sine wave

$$f(x) = \sin \omega_0 x. \tag{7}$$

the scale-space representation is given by

$$L(x; \; t) = e^{-\omega_0^2 t/2} \sin \omega_0 x. \tag{8}$$

and the amplitude of the mth-order normalized derivative operator is

$$L_{\xi^m, max}(t) = t^{m\gamma/2} \, \omega_0^m \, e^{-\omega_0^2 t/2}. \tag{9}$$

This function assumes a unique maximum over scales at

$$t_{max, L_{\xi^m}} = \frac{\gamma \, m}{\omega_0^2}, \tag{10}$$

implying that the corresponding σ-value ($\sigma = \sqrt{t}$) is *proportional* to the wavelength $\lambda = 2\pi/\omega_0$ of the signal. In other words, the wavelength of the signal can be detected from the maximum over scales in the scale-space signature of the signal (see figure 2). In this respect, the scale selection approach has similar properties as a local Fourier analysis, with the difference that there is no need for explicitly determining a window size for computing the Fourier transform.
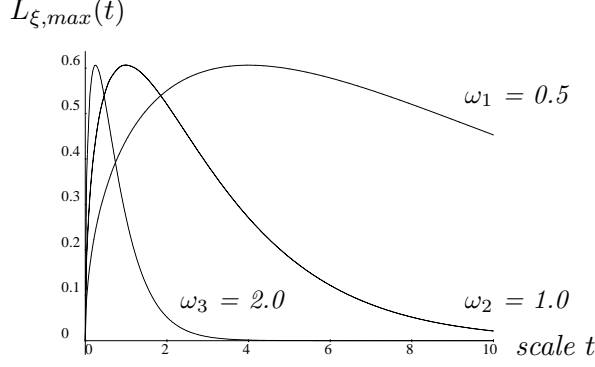

Figure 2: The amplitude of first-order normalized derivatives as function of scale for sinusoidal input signals of different frequency ($\omega_1 = 0.5$, $\omega_2 = 1.0$ and $\omega_3 = 2.0$).

Scale invariance property of the scale selection mechanism.   If a local maximum over scales in the normalized differential expression $\mathcal{D}_{\gamma-norm}L$ is detected at the position x_0 and the scale t_0 in the scale-space representation L of a signal f, then for a signal f' rescaled by a scaling factor s such that

$$f(x) = f'(sx) \tag{11}$$

the corresponding local maximum over scales is assumed at

$$(x_0', t_0') = (sx_0, s^2 t_0). \tag{12}$$

This property shows that the selected scales follow any size variations in the image data, and this property holds for all homogeneous polynomial differential invariants (see [18]).

Necessity of the γ-normalization.   In view of the abovementioned scale invariance result, one may ask the following. Imagine that we take the idea of performing local scale selection by local maximization of some sort of normalized derivatives (not specified yet). Moreover, let us impose the requirement that the scale levels selected by this scale selection mechanism should commute with size variations in the image domain according to equation (11) and (12). Then, what types of scale normalizations are possible? Interestingly, it can then be shown that the form of the γ-normalized derivative normalization (6) arises by *necessity*, *i.e.*, with the free parameter γ it spans all possible reasonable scale normalizations (see [18] for a proof).

Interpretation of normalized Gaussian derivative operators. The idea is that the normalized scale-space derivatives will be used as a basis for expressing a large class of image operations, formulated in terms of normalized differential entities. Equivalently, such derivatives can be computed by applying γ-normalized Gaussian derivative operators

$$g_{x_i^m, \gamma-norm}(x;\ t) = t^{m\,\gamma/2} \partial_{x_i} \left(\frac{1}{(2\pi\sigma^2)^{D/2}} e^{-(x_1^2 + \cdots + x_D^2)/2t} \right) \tag{13}$$

to the original D-dimensional image. It is straightforward to show that the L_p-norm of such a γ-normalized Gaussian derivative kernel is

$$\|h_{\xi^m}(\cdot;\ t)\|_p = \sqrt{t}^{\,m(\gamma-1)+D(1/p-1)} \|h_{\xi^m}(\cdot;\ t)\|_p, \tag{14}$$

which means that the γ-normalized derivative concept can be interpreted as a *normalization to constant L_p-norm over scales*, with p given by

$$p = \frac{1}{1 + \frac{m}{D}(1 - \gamma)}. \tag{15}$$

The special case $\gamma = 1$ corresponds to L_1-normalization for all orders m.

Another interesting interpretation can be made with respect to image data $f\colon \mathbb{R}^2 \to \mathbb{R}$ having self-similar power spectra of the form

$$S_f(\omega) = S_f(\omega_1, \ldots, \omega_D) = (\hat{f}\hat{f}^*)(\omega) = |\omega|^{-2\beta} = (\omega_1^2 + \cdots + \omega_D^2)^{-\beta}. \tag{16}$$

Let us consider the following class of energy measures, measuring the amount of information in the mth order γ-normalized Gaussian derivatives

$$E_m = \int_{x \in \mathbb{R}^D} \sum_{|\alpha|=m} t^{m\gamma} |L_{x^\alpha}|^2 \, dx. \tag{17}$$

In the two-dimensional case, this class includes the following differential energy measures:

$$E_0 = \int_{x \in \mathbb{R}^2} L(x;\ t)^2 \, dx, \tag{18}$$

$$E_1 = \int_{x \in \mathbb{R}^2} t^\gamma (L_\xi^2 + L_\eta^2) \, dx, \tag{19}$$

$$E_2 = \int_{x \in \mathbb{R}^2} t^{2\gamma} (L_{\xi\xi}^2 + 2L_{\xi\eta}^2 + L_{\eta\eta}^2) \, dx, \tag{20}$$

$$E_3 = \int_{x \in \mathbb{R}^2} t^{3\gamma} (L_{\xi\xi\xi}^2 + 3L_{\xi\xi\eta}^2 + 3L_{\xi\eta\eta}^2 + L_{\eta\eta\eta}^2) \, dx. \tag{21}$$

It can be shown that the variation over scales of these energy measures is given by

$$E_m(\cdot;\ t) \sim t^{\beta - D/2 - m(1-\gamma)}, \tag{22}$$

and this expression is scale independent if and only if

$$\beta = \frac{D}{2} + m(1 - \gamma). \tag{23}$$

Hence, *the normalized derivative model is neutral with respect to power spectra of the form*

$$S_f(\omega) = |\omega|^{-D-2m(1-\gamma)}. \tag{24}$$

Empirical studies on natural images often show a qualitative behaviour similar to this [6].

4 The automatic scale selection mechanism in operation

The results presented so far apply generally to a large class of image descriptors formulated in terms of differential entities derived from a multi-scale representation. The idea is that the differential entity $\mathcal{D}$ used for automatic scale selection, together with its associated normalization parameter γ should be determined for the task at hand. In this section, we shall present several examples of how this scale selection mechanism can be expressed in practice for various types of feature detectors.

Edge detection. Let us first turn to the problem of edge detection, using the differential definition of edges expressed in equation (3). A natural measure of edge strength that can be associated with this edge definition is given by normalized gradient magnitude

$$\mathcal{E}_{\gamma-norm} = t^{\gamma} L_v^2 = t^{\gamma}(L_x^2 + L_y^2). \tag{25}$$

If we apply the edge definition (3) at all scales, we will sweep out an edge surface in scale-space. On this edge surface, we can define a *scale-space edge* as a curve where the edge strength measure assumes a local maximum over scales

$$\begin{cases} \partial_t(\mathcal{E}_{\gamma-norm}L(x, y;\ t)) = 0, \\ \partial_{tt}(\mathcal{E}_{\gamma-norm}L(x, y;\ t)) < 0, \\ L_{vv}(x, y;\ t) = 0, \\ L_{vvv}(x, y;\ t) < 0. \end{cases} \tag{26}$$

To determine the normalization parameter γ, we can consider an idealized edge model in the form of a diffuse step edge

$$f(x, y) = \int_{x'=-\infty}^{x} h(x';\ t_0)\, dx', \tag{27}$$

It is straightforward to show that the edge strength measure is maximized at

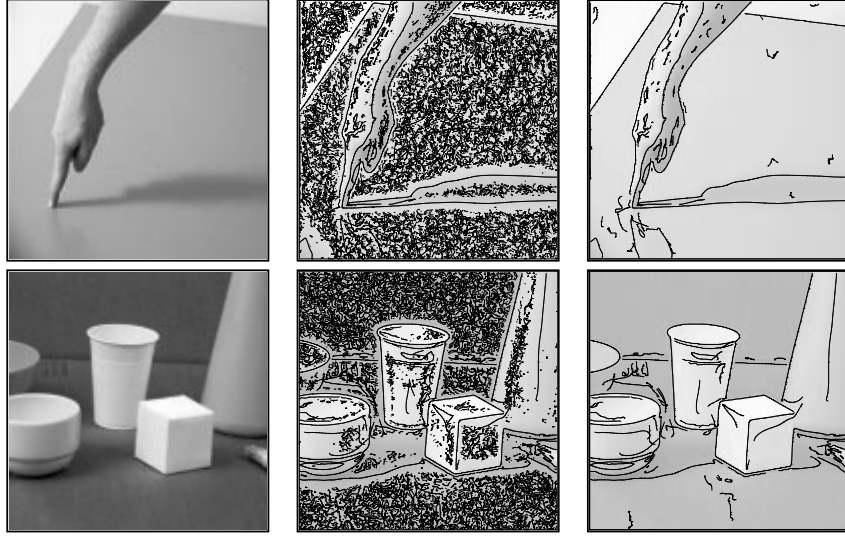$$t_{\mathcal{E}_{\gamma-norm}} = \frac{\gamma}{1 - \gamma}\, t_0. \tag{28}$$

Figure 3: The result of edge detection with automatic scale selection based on local maxima over scales of the first order edge strength measure $\mathcal{E}L$ 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.
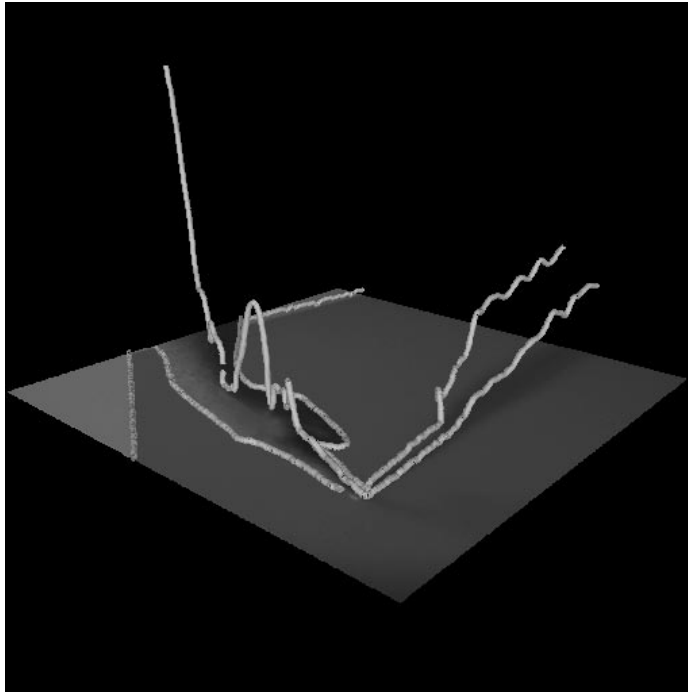

Figure 4: Three-dimensional view of the 10 most significant scale-space edges extracted from the arm image. From the vertical dimension representing the selected scale measured in dimension length (in units of $\sqrt{t}$), it can be seen how coarse scales are selected for the diffuse edge structures (due to illumination effects) and that finer scales are selected for the sharp edge structures (the object boundaries).

If we require that this maximum is assumed at t_0, implying that we use a similar derivative filter for detecting the edge as the shape of the differentiated edge, then we obtain $\gamma = 1/2$.

Figure 3 shows the results of detecting edges from two images in this way. The middle column shows all scale-space edges that satisfy the definition (26), while the right column shows the result of selecting the most significant edges by computing a significance measure as the integrated normalized edge strength measure along each connected edge curve

$$E(\Gamma) = \int_{(x;\ t)\in\Gamma} \sqrt{(\mathcal{E}L)(x;\ t)}\, ds. \tag{29}$$

Figure 4 shows a three-dimensional view of the 10 most significant scale-space edges from the hand image, with the selected scales illustrated by the height over the image plane. Observe that fine scales are selected for the edges corresponding to object boundaries. This result is consistent with the empirical finding that rather fine scales are usually appropriate for extracting object edges. For the shadow edges on the other hand, successively coarser scales are selected with increasing degree of diffuseness, in agreement with the analysis of the idealized edge model in (28).

Ridge detection. Let us next turn to the problem of ridge detection, and sweep out a ridge surface in scale-space by applying the ridge definition (4) at all scales. Then, given the following ridge strength measure

$$\mathcal{R}_{norm}L = \mathcal{A}_{\gamma-norm}L = (L_{pp,\gamma-norm} - L_{qq,\gamma-norm})^2 = t^{2\gamma}\left((L_{xx} - L_{yy})^2 + 4\,L_{xy}^2\right). \tag{30}$$

which is the square difference between the eigenvalues $L_{pp,\gamma-norm}$ and $L_{qq,\gamma-norm}$ of the normalized Hessian matrix, let us define a *scale-space ridge* as a curve on the ridge surface where the normalized ridge strength measure assumes local maxima with respect to scale

$$\begin{cases} \partial_t(\mathcal{R}_{norm}L(x,y;\ t)) = 0, \\ \partial_{tt}(\mathcal{R}_{norm}L(x,y;\ t)) < 0, \\ L_p(x,y;\ t) = 0, \\ L_{pp}(x,y;\ t) < 0. \end{cases} \tag{31}$$

To determine the normalization parameter γ, let us consider a Gaussian ridge

$$f(x,y) = g(x;\ t_0). \tag{32}$$

The maximum over scales in $\mathcal{R}_{\gamma-norm}L$ is assumed at

$$t_{\mathcal{R}_{\gamma-norm}} = \frac{2\,\gamma}{3-2\,\gamma}\,t_0, \tag{33}$$

and by requiring this scale value to be equal to t_0 (implying that a similar rotationally aligned Gaussian derivative filter is used for detecting the ridge as the shape of the second derivative of the Gaussian ridge) we obtain $\gamma = 3/4$.

original grey-level image100 strongest bright ridges10 strongest bright ridges
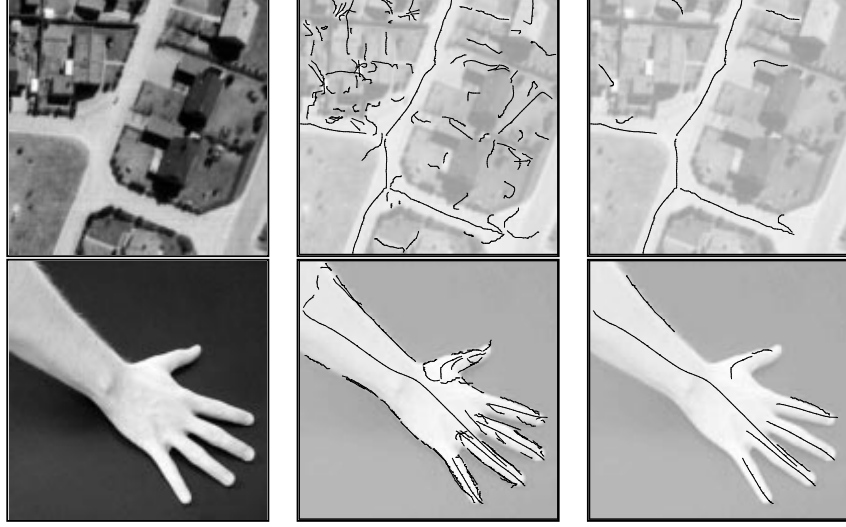
Figure 5: The 100 and 10 strongest bright ridges respectively extracted using scale selection based on local maxima over scales of $\mathcal{A}_{\gamma-norm}$ (with $\gamma = \frac{3}{4}$). Image size: 128×128 pixels in the top row, and 140×140 pixels in the bottom row.

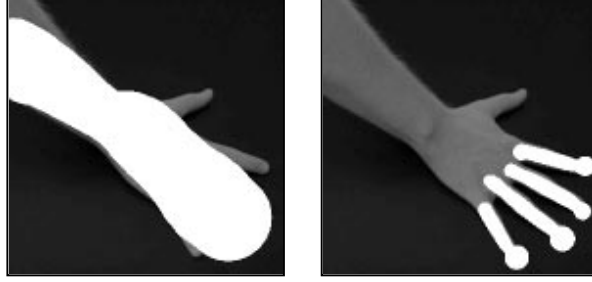*backprojection of ridge 1 backprojection of ridges 2–5*

Figure 6: Alternative illustration of the five strongest scale-space ridges extracted from the image of the arm in figure 5. 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.

Figure 5 shows the result of applying such a ridge detector to an image of an arm and an aerial image of a suburb, respectively. The ridges have been ranked on significance, by integrating the normalized ridge strength measure along each connected ridge curve.

$$R(\Gamma) = \int_{(x;\ t)\in\Gamma} \sqrt{(\mathcal{R}L)(x;\ t)}\, ds, \tag{34}$$

Observe that descriptors corresponding to the roads are selected from the aerial image. Moreover, for the arm image, a coarse-scale descriptor is extracted for the arm as a whole, whereas the individual fingers give rise to ridge curves at finer scales.

Blob detection.   The Laplacian operator $\nabla^2 L = L_{xx} + L_{yy}$ is a commonly used entity for blob detection, since it gives a strong response at the center of blob-like image structures. 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

$$\nabla^2_{norm} L = t(L_{xx} + L_{yy}) \tag{35}$$

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

For a Gaussian blob model defined by

$$f(x, y) = g(x, y; \; t_0) = \frac{1}{2\pi t_0} e^{-(x^2+y^2)/2t_0} \tag{36}$$

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

$$\partial_t (\nabla^2_{norm} L)(0, 0; \; t) = 0 \quad \Longleftrightarrow \quad t_{\nabla^2 L} = t_0. \tag{37}$$

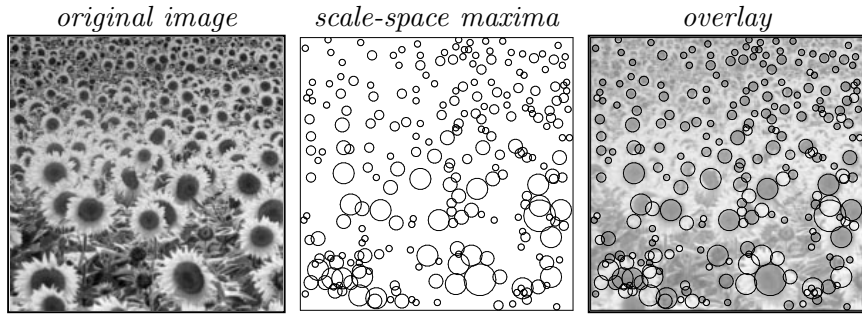*original image*   *scale-space maxima*   *overlay*

Figure 7: 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 $(\nabla_{norm} L)^2$ having the strongest normalized response. (c) Circles overlayed on image.
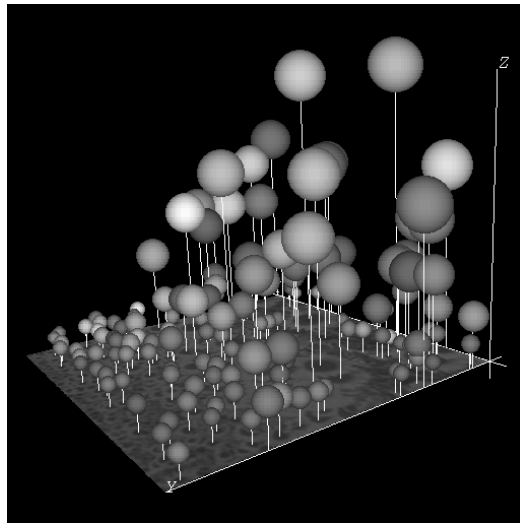

Figure 8: Three-dimensional view of the 150 strongest scale-space maxima of the square of the normalized Laplacian of the Gaussian computed from the sunflower image.

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

Figures 7–8 show the result of applying this blob detector to an image of a sunflower field. In figure 7, each blob feature detected as a scale-space maximum is illustrated by a circle, with its radius proportional to the selected scale. Figure 8 shows a three-dimensional illustration of the same data set, by marking the scale-space extrema by spheres in scale-space. Observe how well the size variations in the image are captured by this structurally very simple operation.

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 [10, 12]. 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

$$\tilde{\kappa}_{norm} = t^{2\gamma} L_v^2 L_{uu}. \tag{38}$$

Figure 9 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 (see figure 10 for results on real-world data).
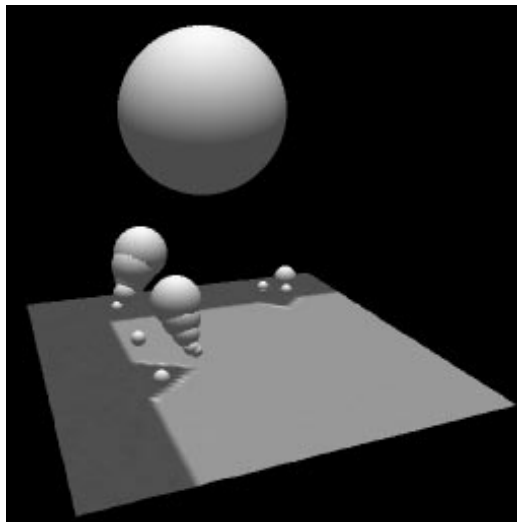

Figure 9: Three-dimensional view of scale-space maxima of $\tilde{\kappa}_{norm}^2$ computed for a large scale corner with superimposed corner structures at finer scales.

5 Application to feature tracking and human computer interaction

We argue that a scale selection mechanism is an essential tool whenever our aim is to automatically interpret the image data that arise from observations of a dynamic world. For

example, if we are tracking features in the image domain, then it is essential that the scale levels are adapted to the size variation that may occur over time.

Figure 10 shows a comparison between a feature tracker with automatic scale selection [2] and a corresponding feature tracker operating at fixed scales. (Both feature trackers are based on corner detection from local maxima of the corner strength measure (38), followed by a localization stage [18] and a multi-cue verification [2].) As can be seen from the results in figure 10, three out of the ten features are lost by the fixed scale feature tracker compared to the adaptive scale tracker.

Initial frame with 14 detected corners   *Tracked features with automatic scale selection*
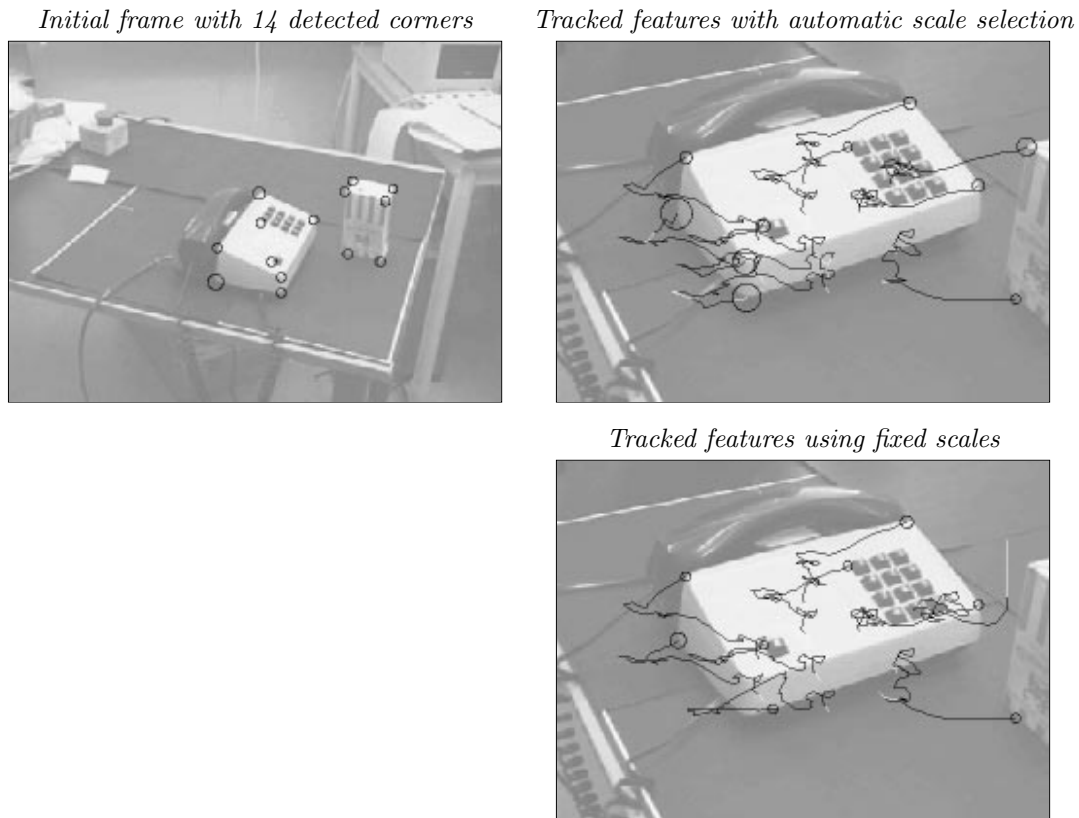

Tracked features using fixed scales

Figure 10: Comparison between feature tracker with automatic scale selection and a feature tracker operating at fixed scale. The left column shows a set of corner features in the initial frame, and the right column gives a snapshot after 65 frames.

A brief explanation of this phenomenon is that if we use a standard algorithm for feature detection at a fixed scale followed by hypothesis evaluation using a fixed size window for correlation, then the feature tracker will after a few frames fail to detect some of the features. The reason why this occurs is simply the fact that the corner feature no longer exists at the predetermined scale. In practice, this usually occurs for blunt corners.

An attractive property of a feature detector with automatic scale selection is that it allows us to capture less distinct features than those that occur on man-made objects. Specifically, we

have demonstrated how it makes it possible to capture features associated with human actions. Figure 11 illustrates one idea we have been working on in the area of visually guided human-computer-interaction. The idea is to have a camera that monitors the motion of a human hand. At each frame blob and ridge features are extracted corresponding to the fingers and the finger tips. Assuming rigidity, the motion of the image features allow us to estimate the three-dimensional rotation of the hand [3]. These motion estimates can in turn be used for controlling other computerized equipment; thus serving as a "3-D hand mouse" [20].

Controlling hand motion   *Detected ridges and blobs*   *Controlled object*
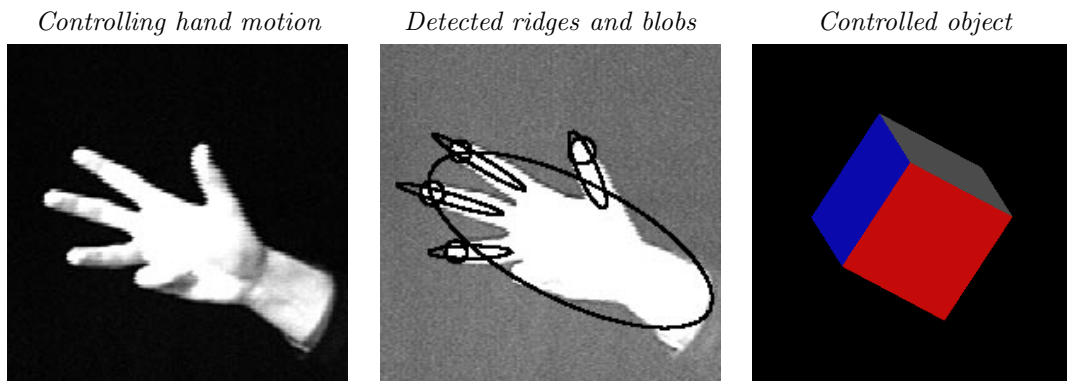
Figure 11: Illustration of the concept of a "3-D hand mouse". The idea is to monitor the motion of a human hand (here, via a set of tracked image features) and to use estimates of the hand motion for controlling other computerized equipment (here, the visualization of a cube).

6 Summary

We have presented a general framework for automatic scale selection as well as examples of how this scale selection mechanism can be integrated with other feature modules. The experiments demonstrate how *abstractions* of the image data can be computed in a conceptually very simple way, by analysing the *behaviour* of image features *over scales* (sometimes referred to as "deep structure"). For applications in other areas as well as related works, see [18, 17, 19] and [1, 26].

References

[1] A. Almansa and T. Lindeberg. Fingerprint enhancement by shape adaptation of scale-space operators with automatic scale-selection. Technical Report ISRN KTH/NA/P--99/01--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, Jan. 1999. (Submitted).

[2] L. Bretzner and T. Lindeberg. Feature tracking with automatic selection of spatial scales. *Computer Vision and Image Understanding*, 71(3):385–392, Sep. 1998.

[3] L. Bretzner and T. Lindeberg. Use your hand as a 3-D mouse or relative orientation from extended sequences of sparse point and line correspondances using the affine trifocal tensor. In H. Burkhardt and B. Neumann, editors, *Proc. 5th European Conference on Computer Vision*, volume 1406 of *Lecture Notes in Computer Science*, pages 141–157, Freiburg, Germany, June 1998. Springer Verlag, Berlin.

[4] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. Communications*, 9:4:532–540, 1983.

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

[6] 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.

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

[8] 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.

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

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

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

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

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

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

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

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

[17] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *Int. J. of Computer Vision*, 30(2):117–154, 1998.

[18] T. Lindeberg. Feature detection with automatic scale selection. *Int. J. of Computer Vision*, 30(2):77–116, 1998.

[19] T. Lindeberg. Principles for automatic scale selection. In B. Jähne (et al.), editor, *Handbook on Computer Vision and Applications*, pages 239–274. Academic Press, Boston, USA, 1999.

[20] T. Lindeberg and L. Bretzner. Förfarande och anordning för överföring av information genom rörelsedetektering, samt användning av anordningen. Patent pending., 1998.

[21] J. M. S. Prewitt. Object enhancement and extraction. In A. Rosenfeld and B. S. Lipkin, editors, *Picture Processing and Psychophysics*, pages 75–149. Academic Press, New York, 1970.

[22] L. G. Roberts. Machine perception of three-dimensional solids. In J. T. Tippet *et al.*, editor, *Optical and Electro-Optical Information Processing*, pages 159–197. MIT Press, Cambridge, Massachusetts, 1965.

[23] 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.

[24] B. M. ter Haar Romeny, L. M. J. Florack, J. J. Koenderink, and M. A. Viergever, editors. *Scale-Space Theory in Computer Vision: Proc. First Int. Conf. Scale-Space'97.* Lecture Notes in Computer Science. Springer Verlag, New York, Utrecht, Netherlands, Jul. 1997.

[25] V. Torre and T. A. Poggio. On edge detection. *IEEE Trans. Pattern Analysis and Machine Intell.*, 8(2):147–163, 1980.

[26] K. Wiltschi, A. Pinz, and T. Lindeberg. An automatic assessment scheme for steel quality inspection. Technical Report ISRN KTH/NA/P--98/20--SE, Dept. of Numerical Analysis and Computing Science, KTH, Stockholm, Sweden, Nov. 1998. (Submitted).

[27] A. P. Witkin. Scale-space filtering. In *Proc. 8th Int. Joint Conf. Art. Intell.*, pages 1019–1022, Karlsruhe, West Germany, Aug. 1983.