Scale Detection and Region Extraction from a Scale-Space Primal Sketch

Tony Lindeberg, Jan-Olof Eklundh

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

Abstract

We present: (1) A multi-scale representation of grey-level shape, called scale-space primal sketch, which makes explicit both features in scale-space and the relations between features at different levels of scale. (2) A theory for extraction of significant image structure from this representation. (3) Applications to edge detection, histogram analysis and junction classification demonstrating how the proposed method can be used for guiding later stage processing. The representation gives a qualitative description of the image structure that allows for detection of stable scales and regions of interest in a solely bottom-up data-driven way. In other words, it generates coarse segmentation cues and can be hence seen as preceding further processing, which can then be properly tuned. We argue that once such information is available many other processing tasks can become much simpler. Experiments on real imagery demonstrate that the proposed theory gives perceptually intuitive results.

1 Introduction

Scale-space theory [30, 15] provides a well-founded framework for dealing with image structures, which naturally occur at different scales. According to this theory one can from a given signal generate a family of derived signals by successively removing features when moving from fine to coarse scale. In contrast to other multi-scale or multi-resolution representations, scale-space is based on a precise mathematical definition of causality, and the behaviour of structure as scale changes can be analytically described. However, the information in the scale-space embedding is only *implicit* in the grey-level values. The smoothed images in the raw scale-space representation contain no explicit information about the features or the relations between features at different levels of scale.

The goal of this paper is to present such an explicit representation, called the scale-space primal sketch, and to demonstrate that it enables extraction of significant image structure in such a way that the output can be used for guiding later stage processing. We shall treat intensity images, that is the grey-level landscape, and the chosen objects will be blobs, i.e. bright regions on dark backgrounds or vice versa. However, our theory applies to any bounded function and is therefore useful in many tasks occurring in computer vision, such as the study of level curves and spatial derivatives in general, depth maps etc, and also, histograms, point clustering and grouping in one or several variables.

1.1 Scale and Segmentation

Many methods in image analysis implicitly assume that the problems of scale detection and segmentation have been solved. An illustrative example is for instance the well-known trade-off question arising in gradient calculations: A small step size leads to a small truncation error but the noise sensitivity might be severe. Conversely, a large step size reduces the noise sensitivity at the cost of an increasing truncation error. In the worst case we may even miss the slope of interest and get meaningless results if the difference quotient is formed over a wider distance than the object to be considered in the image. The problem falls back on the basic scale problem, namely that objects in the world and features in an image make sense only over a limited range of scale.

A commonly used technique to improve the results obtained in computer vision and other branches of applied numerical analysis is by pre-processing the input data with some amount of smoothing and/or careful tuning of the operator size or some other parameters. In some situations the output result may depend strongly on these processing steps. For some algorithms these so-called tuning parameters can be estimated; in other cases they are set manually. A robust image analysis method, intended to work in an autonomous robot situation, must however be able to make such decisions. How should this be done? We contend that these problems are in many situations nothing but disguised scale problems.

Also, in order to apply a refined mathematical model like a differential equation or some kind of deformable template it is necessary to have some kind of qualitative starting information, i.e., a domain where the differential equation is (assumed to be) valid or an initial region for application of the raw deformable template. How should we select such regions *automatically*? Many methods cannot be used unless this non-trivial part of the problem has been solved.

How do we detect appropriate scales and regions of interest when there is no a priori information available, i.e., how to determine the scale of an object and where to search for it before knowing what kind of object we are studying and before knowing where it is located. This problem which arises implicitly in many kinds of processes, e.g. dealing with texture, contours etc, seems to boil down to an impossible chicken-in-the-egg problem. The solution of the recognition problem requires the solution of the scale and region problems and vice versa. In this work, however, we will show that such *scale and region determination actually can be performed computationally from raw image data by early low-level processing*, at least for regions which stand out from the surrounding. The basic tools for the analysis will be the scale-space theory and a variational principle where we vary the scale parameter systematically to find locally stable states.

We argue that once the scale information is available and once we have extracted "regions of interest" the remaining processing tasks can be much simpler. We support this claim by experiments on edge detection and classification based on local features.

1.2 Detection of Image Structure

The main features arising in the scale-space representation are smooth regions which stand out from their surroundings. We will call them blobs (a precise definition will be given later). The purpose of the suggested scale-space primal sketch representation is to make these blobs as well as their relations across scales explicit. The idea is also that this representation should reflect the intrinsic shape of the grey-level landscape and not be an effect of some externally chosen criteria or tuning parameters. The theory should in a bottom-up fashion allow for a data-driven detection of significant structures, their relations and the scales at which they occur. We will experimentally show that the proposed representation gives perceptually reasonable results, in which salient structures are (coarsely) segmented out. Hence, this representation can serve as a guide to subsequent, more finely tuned processing, that requires knowledge about the scales at which structures occur. In this respect it can serve as a mechanism for focus-of-attention.

Since the representation tries to capture *all* the important structures with a small set of primitives, it bears some similarity to Marr's primal sketch, even though fewer primitives are used. However, the central issue here is to explicitly represent also the scale at which different events occur. In this respect our work addresses problems similar to those studied by Bischof & Caelli [4]. They try to parse the scale-space by defining a measure of stability. However, their work focuses on zero-crossings of the Laplacian. Moreover, they overlook the fact that in measuring significance or stability of structures we must treat the scale parameter properly. The behaviour of structures over scale will be analyzed to give the basis of such measurements. Of course, several other representations of the grey-level landscape have been proposed without relying on scale-space theory. Let us also note that Pizer and his co-workers, see e.g. [24], have performed studies of the behaviour of local extrema in scale-space. However, we will defer discussing the relations to other work, see Section 6.1, until we have described our own methodology.

The idea of scale-space representation of images, suggested by Witkin [30] has, in particular, been developed by Koenderink & van Doorn [15, 16]. Our work is aimed at complementing this work by considering the *computational aspects* and by adding means of *making significant structures and scales explicit.* The main idea behind our approach is to *link* features at different levels of scale into higher order objects, called scale-space blobs, and to extract significant image features based on their appearance and stability in scale-space.

2 The Scale-Space Primal Sketch

From experiments one can (visually and subjectively) observe that the main features arising in the scale-space representation seem to be blob-like, i.e., they are regions either brighter or darker than the background. In particular regions, which appear to stand out from the surroundings in the original image, seem to be further enhanced by scale-space smoothing. In the suggested scale-space primal sketch we focus on this aspect of image structure with the purpose of building a representation to make such information in scale-space explicit. Therefore, there is a need to formalize what should be meant by a "blob".

2.1 Grey-Level Blobs

It is clear that a blob should be a region associated with a (at least one) local extremum point. However, it is essential to define the spatial extent of the blob region around the extremum. Ehrich & Lai [10] considered this problem. They allowed peaks to extend to valleys, a definition that will give unintuitive results e.g. for small peaks on large slopes. Koenderink & van Doorn [15] briefly touch upon the problem and our definition is inspired by their argument.

The blob definition we base this work on should be evident from Figure 1. The basic idea is to let the blob extend "until

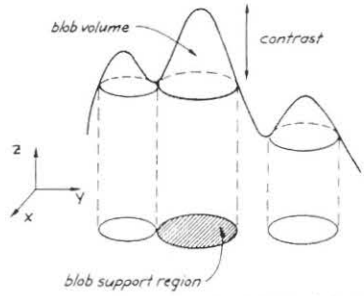

Figure 1: *Illustration of our grey-level blob definition for a two-dimensional signal. (This figure shows bright blobs on dark background).*

it would merge with another blob". To intuitively illustrate this notion consider a grey-level image at a fixed level of scale, and study the case with bright blobs on dark background. Imagine the image function as a flooded grey-level landscape. If the water level sinks gradually, peaks will appear. At some instances two different peaks become connected. The corresponding elevation levels or grey-levels we call the *base-levels* of the blobs, and are used to the delimit the spatial extent of the blobs. The *support region* of a blob is defined to consist of those points that have a grey-level exceeding the base-level and can be reached from the local maximum point without descending below the base-level of the blob. In this sense the blob definition can be regarded as rather conservative, since no attempt is made to include points in other directions. Finally, we define the *grey-level blob* to be the 3D volume delimited by the grey-level surface and the base-level. Its magnitude comprises both the amplitude and the spatial extent of the blob. Precise mathematical definitions of these quantities are given in [19].

Local minima can be treated analogously and every local minimum point will give rise to a dark blob on bright background. Hence, we get *separate* systems for bright blobs on dark background and dark blobs on bright background. This implies that some points will be left unclassified. Consequently, the given definition will, in contrast to, e.g. the sign of the Laplacian of the Gaussian, only attempt to make a partial (and hopefully safer) classification of the grey-level landscape.

Blobs are not purely local features, as are extrema, but *regional.* An inherent property of the stated definition is that it leads to a *competition between parts,* and that things manifest themselves only compared to their background. Also, the blobs are directly determined from *geometric properties* of the grey-level function since the base-level is attained at a saddle point. These aspects reflect important principles of the approach.

2.2 Scale-Space Blobs

In general, a grey-level blob existing at one level of scale in scale-space will have a similar blob both at a finer and a coarser level of scale. By *linking* together such grey-level blobs across scales we obtain 4D objects with extent both in space, scale and grey-level that we call *scale-space blobs.* At some levels of scale in scale-space it may be impossible to accomplish a single link between a grey-level blob at the current level of scale to a similar grey-level blob at a coarser or finer scale — a catastrophe affecting the connectivity of the blobs has occurred. According

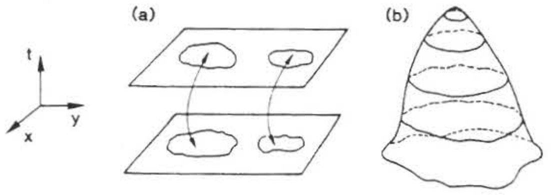

Figure 2: *(a) By linking together similar grey-level blobs at adjacent levels of scale we obtain (b) scale-space blobs, which are objects having extent both in space, scale and grey-level. (In this figure we have omitted the grey-level coordinate.)*

to a classification[1] by Koenderink & van Doorn [16] three possible types of (generic) blob events may occur when the scale parameter increases: *annihilation*, *merge* and *split*. The scale

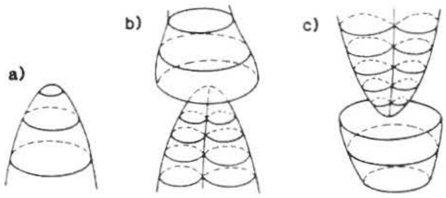

Figure 3: *Common blob events in scale-space: (a) annihilation (b) merge (c) split.*

levels where these singularities take place are used for delimiting the extent in the scale direction of the scale-space blobs. Consequently, every scale-space blob will be associated with a minimum scale and a maximum scale, denoted the *appearance scale* and the *disappearance scale* respectively, see Figure 2. The difference[2] between these values yields the *scale-space lifetime* of the blob.

2.2.1 Motivation and Experimental Results

It is easy to realize that the concept of a grey-level blob at a single level of scale is not powerful enough to allow for extraction of relevant image structures. It leads to an extreme degree of noise sensitivity, since two closely situated local extrema will neutralize each other. This means that a large peak distorted by a few superimposed low-amplitude local extrema will be hard to detect as one unit, because on a first attempt only the fine scale blobs will be found. To some extent this kind of problem can be circumvented by scale-space blurring. But, it is certainly a far from trivial problem to determine a proper amount of smoothing, *automatically*, based on existing conventional methods.

In Figure 4 we give an example with a toy block image showing how (the support regions of) the extracted grey-level blobs behave with increasing scale together with the raw grey-level images in the scale-space representation. We see that at fine levels of scale mainly small blobs, due to noise and surface texture, are detected. When the scale increases the noise blobs disappear gradually, although much faster in regions near steep gradients. Notable in this context is that blobs due to noise can survive for a long time in scale-space if located in regions with slowly varying grey-level intensity. This observation shows that *scale-space lifetime alone is not appropriate as a significance measure*,

[1] The derivation given by Koenderink, van Doorn is based on a different blob definition than the definition used in this paper. However, since both these blob concepts are directly determined by local extrema it follows that the possible types of blob events in scale-space will be the same.

[2] It turns out that some transformation of the scale parameter is necessary in order for the difference between scale values to capture the concept of scale-space lifetime "properly". This issue is developed in Section 2.3.

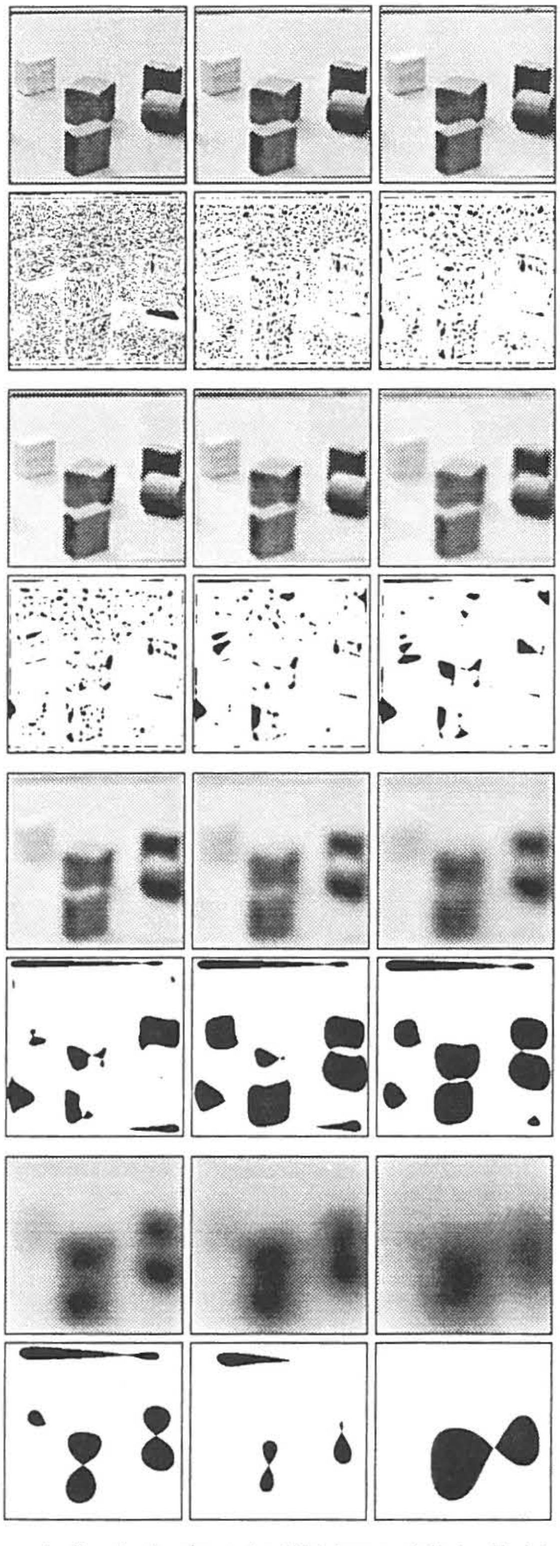

Figure 4: *Grey-level and grey-level blob images of the toy block image at scale levels $t = 0, 1, 2, 4, 8, 16, 32, 64, 128, 256, 512$ and 1024 (from top left to bottom right). The grey-level images have been histogram equalized in order to increase the contrast.*

since then the significance of such blobs due to noise would be substantially overestimated[3]. At coarse levels of scale, the toy blocks appear as single blob objects. Finally, at very coarse levels of scale, adjacent blocks become grouped into larger entities.

The aim with the suggested blob linking between scales is to determine which structures in an image can be regarded as significant, *without any a priori information* about neither scale, spatial location nor the shape of the primitives. As we will see later, the output from the linking procedure also enables determination of a relevant scale for each blob, that is a suitable amount of blurring for dealing with that *individual*[4] blob.

The idea behind this combination of grey-level blobs and scale-space smoothing is that instead of trying to design "an intelligent blob detector" to handle difficult situations as the case above with superimposed local extrema, we establish a *simple blob definition based on distinct topographical properties of the signal.* Then, we use *the scale-space embedding to integrate local properties into regional descriptors*, and to make the hierarchical relations between features at different levels of scale explicit. Another aspect of this definition of a scale-space blob is that we treat *the scale parameter as equally important as the spatial and grey-level coordinates*, and *the primitives of our representation are objects having extent not only in grey-level and space, but also in scale.*

2.3 Measuring Significant Image Structure

Since the ultimate goal of the analysis is to extract important regions in the image based on the significance of the scale-space blobs in the scale-space representation, there is an absolute need for some methodology for comparing blob significance *between* different levels of scale. In other words, what we actually desire is a mechanism to judge if a blob, existing only at coarse scales, can be regarded as more significant or less significant than a blob present primarily at fine scales. The approach we propose is to use the volume of the scale-space blobs in scale-space. We suggest that it is a useful quantity for such a significance measure, since it comprises both the grey-level blob volume, which is a combination of the contrast, spatial extent and lifetime of the blob in scale-space, see also Section 3.1.

However, if one is to base a significance measure on the scale-space blob volume, it is of crucial importance that the scale parameter and the grey-level blob volume are measured in proper units. For instance, measuring scale-space lifetime just by a simple difference or a difference of the logarithms between the disappearance and appearance scales will not work, since then either the importance of coarse scale or fine scale blobs would be substantially overestimated.

Consequently, it is necessary to introduce a transformed scale parameter, which we will term *effective scale* τ, that captures the scale-space lifetime properly. From a requirement that the amount of structure, which is destroyed if the effective scale parameter is increased with a small increment, should be independent of the current scale and the current amount of structure in the image we have shown [19] that this concept can be defined in essentially one way only. One obtains

$$\tau(t) = \log \frac{p(0)}{p(t)} \tag{1}$$

where the quantity $p(t)$ is the "expected density of local extrema at scale t". For implementational purpose $p(t)$ is estimated from simulation results for point noise images, see Figure 5.

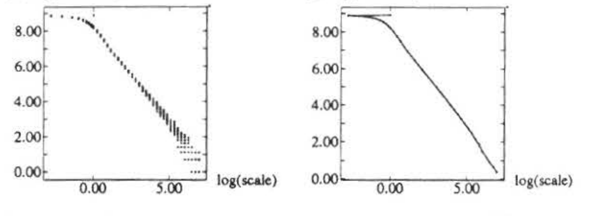

Figure 5: *Experimental results showing the number of local extrema as function of the scale parameter t in log-log scale (a) measured values (b) accumulated mean values. The dashed line indicates the value at $t = 0$. Note that a straight-line approximation is valid only in the interior part of the scale interval. In the left part we have interference with the inner scale, given by the sampling density, and the right part there is interference with the outer scale, given by the size of the image.*

Similarly, the grey-level blob volumes need to be transformed, since their expected magnitude will vary substantially with scale. A natural normalization to perform is to subtract the mean value, $V_m(t)$, and divide by the standard deviation, $V_\sigma(t)$.

$$V_{transf,prel}(t) = \frac{V(t) - V_m(t)}{V_\sigma(t)} \tag{2}$$

However, such a quantity is not suited for integration (which is a necessary step in the computation of the scale-space blob volumes), since it will assume negative values. Therefore, in the current implementation, we have chosen to define the transformed blob volume as

$$V_{transf}(t) = \begin{cases} 1 + V_{transf,\,prel} & \text{if } V_{transf,\,prel} \geq 0 \\ e^{V_{transf,\,prel}} & \text{otherwise} \end{cases} \tag{3}$$

which empirically turns out to give reasonable results. As in the previous case, $V_m(t)$ and $V_\sigma(t)$ are determined from simulation results for point noise images, i.e. we estimate to which extent accidental groupings take place in scale-space and compensate for this phenomenon.

In this context it is of considerable importance to take the discrete aspects of implementation into actual account. Direct application of results obtained from the scale-space theory for continuous signals could give a severe bias at finer levels of scale where the effects from the spatial sampling can be strong. For instance, in continuous theory (for normal processes) it holds that the density of local extrema varies with scale approximately as t^α [19], which means that the transformation function $\tau(t)$ would be logarithmic (and the graph in Figure 5 would be a straight line). Then, blobs existing at $t = 0$ would be given infinite lifetime and thereby also infinite significance. By using a scale-space concept especially designed for discrete signals, see [18], and by simulating the relevant quantities, we try to overcome such problems.

2.4 Resulting Representation

To summarize, the data structure we propose is a tree-like multi-scale representation of blobs at all levels of scale in scale-space *including* the relations between blobs at different levels of scales.

[3] Of course, the contrast of such noise blobs decreases, but it is far from clear that it is possible to set a globally valid threshold on objective ground.

[4] We emphasize the word individual since we believe that *stable scales when they exist are in local properties associated with objects — not with entire images*. However, the assumption about a globally stable scale is sometimes used implicitly in computer vision algorithms, for instance when edge detection is performed with uniform smoothing all over an image. Instead we believe that better performance can be accomplished if the scale levels are adapted to the local image structure, see Section 4 for an example.

Grey-level blobs should be extracted at all scales, the bifurcations occurring in scale-space be explicitly registered and grey-level blobs stable over scales be linked across scales into the higher-order objects called scale-space blobs.

Since this representation tries to capture the significant features and events occurring in scale-space with a small set of primitives we call it a *scale-space primal sketch.* Every scale-space blob contains explicit information about which grey-level blobs it consists of. The grey-level blobs are given at (sampled) scale levels obtained from an adaptive scale linking and refinement procedure outlined in Section A.2. Further, the (normalized) scale-space blob volume, the appearance scale, the disappearance scale and the scale-space lifetime have been computed (using straightforward numerical techniques). The scale-space blobs "know" about the type of bifurcations (annihilation, split, merge) that have taken place at the appearance and disappearance scales. They also have links to the other scale-space blobs involved in the bifurcation processes. Hence, the data structure we have computed explicitly describes the hierarchical relations between blobs at different levels of scale.

In the next section we will show how some directly available information from this representation can be used for extraction of significant image structure. Further applications for tuning later stage processing and guiding the focus-of-attention are given in Sections 4-5 and [20].

3 Detecting Significant Structures and Their Scales

One motivation for this research was to investigate if the scale-space model really allows for determination and detection of global and stable phenomena. In this section we will demonstrate that this is indeed possible and that the presented representation can be used for extracting stable scales and regions of interest from an image, in a solely data-driven way. The treatment is based on the assumption that:

- *Features, which are significant in scale-space, correspond to relevant details in the image.*

More precisely, since the primitives we intend to use are scale-space blobs we formulate it as follows:

Assumption 1 *A scale-space blob having a large scale-space volume in scale-space corresponds to a relevant region in the image.*

A scale-space blob will in general exist over some scale interval in scale-space. When there is a need to reduce the amount of data represented and to select an appropriate scale and a spatial region for a scale-space blob we make use of the following postulates:

Assumption 2 *The scale-level, at which a scale-space blob assumes its maximum grey-level blob volume, is a relevant scale for treating that individual blob.*

Assumption 3 *The spatial extent of a scale-space blob can be represented by the blob support region corresponding to its grey-level blob at the relevant scale.*

Below, we will give experimental results showing that these statements, combined with a careful computational treatment of the scale-space, segment out perceptually relevant image regions.

3.1 Motivation for the Assumptions

A central issue in low-level vision concerns what should be meant by "image structure". In other words, which features in an image should be regarded as important, and which ones can be rejected as noise. Notably, Lowe [21] defines structure based on non-accidentalness. However, such an approach requires a probabilistic model of the situation. It is well known that it is difficult to find a statistical model generally valid for the image formation process.

In this work we take an alternative viewpoint and suggest *a definition of structure based on features, which are stable with respect to (appropriately selected) transformations and/or parameter variations.* For this specific treatment the transformation family of interest is the semi-group of convolution transformations associated with the scale-space smoothing. The parameter we vary is the scale parameter. We think that features stable, or invariant, with respect to variations in scale can be regarded as significant. In more general situations one could also imagine the probing transformation as given by variations in viewing distance (focusing), spatial resolution, regularization parameters etc.

One can motivate such a standpoint by a pragmatic argument. If a feature is to be useful for recognition it must necessarily be stable with respect to small disturbances. Otherwise it can hardly be practically useful, since then, it inherently cannot be computed accurately. This definition of structure in terms of *transformational invariance* also induces a straightforward and general method for detecting significant image features, namely by *subjecting the image to systematic parameter variations.* In line with that idea we believe that those features, that are the most stable ones during such a parameter variation process, can be regarded a strong candidates for being useful for later processing and possibly recognition.

Of course, the reverse statement does not hold. There are many other sources of information, e.g., lines in line-drawings, that are not captured by a blob concept and scale-space smoothing. In this work we focus mainly on one aspect of image structure, namely regions that are brighter or darker than the background and stand out from the surrounding.

The approach is closely related to Witkin's [30] observation about correspondence between stability in scale-space and perceptual salience. However, here we base the stability measure on the scale-space blob volumes instead of the scale-space lifetime. The intention is that this choice also should reflect the size of the blobs and how strongly they manifest themselves with respect to the background. As mentioned in Section 2.2.1 we have observed that small blobs due to noise can survive over large range of scales if they are located in regions with slowly varying grey-level.

Because of complexity arguments, the entire parameter variation information from the low-level modules cannot not be transferred to modules intended to perform higher-level processing tasks. Instead we think that low-level modules working after this paradigm should be able to extract stable intervals, and that it should suffice to determine a representative descriptor for each important stability region.

The second and third assumptions express such a desire to represent a scale-space blob with a grey-level blob at a single level of scale in order to have a more compressed[5] representation — an *abstraction* for further processing. We believe that a relevant scale of a scale-space blob should be a scale where the grey-level blob manifests itself "as its best", i.e., it should be the scale level where the blob response "is maximally strong". Empirically, we have found that this suggested scale value will give a good description of the situation. It turns out that it often will be

[5] Worth noting is also that Assumption 2 implies a projection from 4D scale-space blob to a 3D grey-level blob and Assumption 3 a projection from the 3D grey-level blob to its 2D blob support region.

close to the appearance scale of the scale-space blob.

3.2 Basic Extraction Method for Image Structure

The basic methodology, in our suggested algorithm for extraction of important image structure, should be obvious from the previous presentation.

- Generate the suggested multi-scale representation, where blobs are extracted at all levels of scale and linked across scales into scale-space blobs.
- Compute the volume of each scale-space blob based on the notions of effective scale and transformed grey-level blob volumes.
- For each scale-space blob determine the scale where it assumes its maximum grey-level blob volume, and extract the blob support region of the grey-level blob at that scale.
- Sort the scale-space blob in descending significance order, i.e., with respect to their scale-space blob volumes.

3.3 Experimental Results

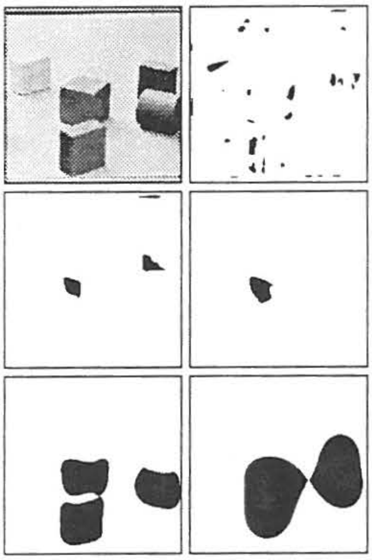

Figure 6: *The 50 most significant dark blobs from a toy block image. Note how these images have been produced — they are not just blob images at a few levels of scale. Instead every blob has been marked at its representative scale, that is the previously mentioned scale where the scale-space blob assumes its maximum grey-level blob volume. Finally, (the support regions of) the blobs have been drawn in different images as to avoid overlap.*

In Figures 6-8 we show the results for three different images with toy blocks, a telephone and a calculator and a dot pattern. The reader is encouraged to study them carefully. For display purpose we have extracted the N scale-space blobs having the largest blob volumes. As we see, the blocks are extracted from the toy block scene. Also, the groups of adjacent blocks and the imperfections in the image acquisition near the boundaries are pointed out. In the telephone scene the buttons, the keyboard, the calculator, the cord and the receiver are detected as single units. Finally, in the dot pattern image the algorithm finds all dots and performs those groupings we find perceptually reasonable. In order to show the spatial relations between the blobs at the various levels of scale we have also drawn the blob boundaries from these examples in Figure 9.

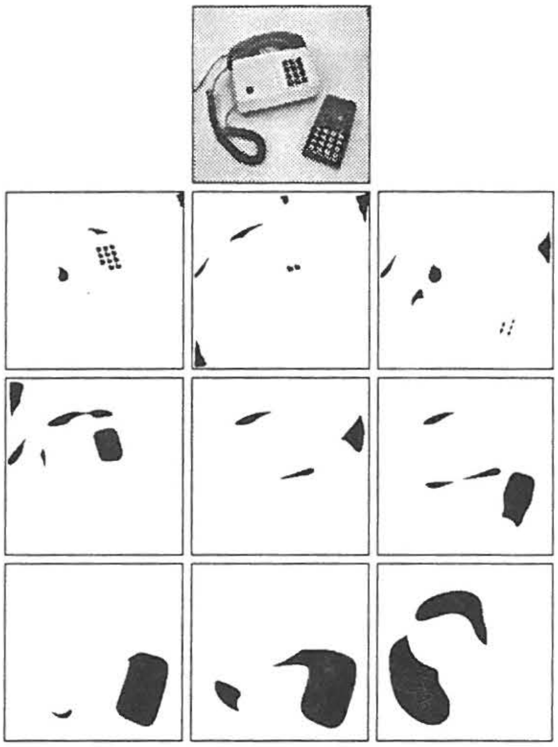

Figure 7: *The 50 most significant dark blobs from a telephone and calculator image.*

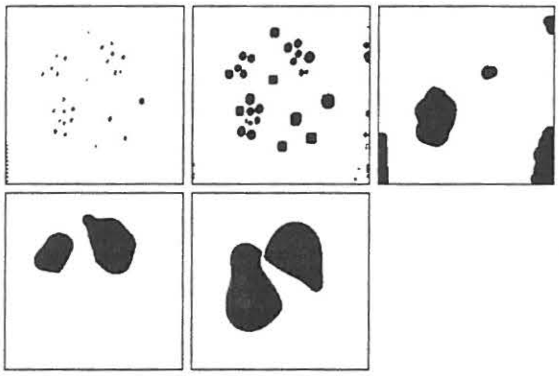

Figure 8: *The 50 most significant dark blobs from a dot pattern image.*

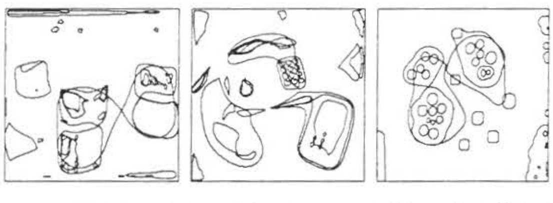

Figure 9: *Blob boundaries of the 50 most significant dark blobs in (a) the toy block image (b) the telephone and calculator image and (c) the dot pattern image.*

Let us conclude by stressing that we *extract the intrinsic shape of the grey-level landscape in a completely bottom-up data-driven way without any assumptions about the shape of the primitives* (except for the fact that the scale-space smoothing favours blob-like objects, since it is equivalent to correlation with a Gaussian-shaped kernel). The segmentation we get is coarse in the sense that the localization of object boundaries may be poor, due to the natural shape distortions which occur in scale-space. However, the segmentation is safe in the sense that those regions, who are given by the scale-space blobs with large scale-space volume, really serve as landmarks of significant structure in the image, with information about

- the *approximate location and extent* of relevant regions in the image.
- an *appropriate scale* for treating those regions.

This is exactly the kind of coarse information[6] that is necessary for many higher-level processes, see e.g. the application to edge detection in Section 4.

3.4 Further Treatment of the Generated Hypothesis

The number of scale-space blobs N selected for display above is, of course, rather arbitrary. However, note that there is a well-defined ranking between the blobs. If one studies their significance values, see [19], one can observe that those blobs we regard as the most important ones have significance values standing out from the other ones. Hence, it seems plausible that a few image regions could be extracted just based on this observation. In more general situations there is a need for feed-back or reasoning.

The output from this algorithm should not be over-estimated. Since the results come from a low-level processing module they should be interpreted as such, namely *as indicators signalling that "there might be something there of about that size — now some other module should take a closer look"*. From this viewpoint it can be noted how well the extracted regions describe the images in the previous examples, considering that the blobs have been extracted almost without any a priori information.

In principle we think that a reasoning process, working on this output could operate in either of two possible modes:

1. Use a threshold on the significance measure. In a real system such a threshold could in some applications be set from given context information and expectations.
2. Evaluate the generated hypothesis in decreasing order of significance, i.e., try to analyze them one by one in a feed-back loop. Continue as long as the hypotheses deliver meaningful interpretations for the higher-level modules.

Another inherent property with this representation is that it does not have any limiting requirement that there is just one possible interpretation of a situation. Instead, given some region in space, several hypotheses may be generated and active for it (or parts of it) concerning structures at different levels of scale.

4 Application to Edge Detection

As one application of the suggested representation we present an integration of the output from the scale-space primal sketch with an edge detection method known as edge focusing, developed by Bergholm [2, 3]. The leading idea is to use the output scale information to guide an *edge detection scheme working at an adaptively determined level of scale.* We demonstrate that this task can be relatively easy and that there is no need for thresholding on gradient magnitude, since the image has been subjected to an appropriately selected amount of blurring. Hence, the detection step will be safe, but the localization could certainly be poor due to the natural shape distortions, that occur at coarser levels of scale in scale-space. However, the localization can be improved using the edge focusing method, which traces the safely detected edges at coarse levels of scale to corresponding and better localized edges at finer levels of scale. Hence, the resulting method will achieve a good compromise between the two conflicting goals in edge detection, namely eliminating the noise without distorting the localization of the edges.

We do not maintain that this part of the presentation describes any "optimal way" to solve every occurring subproblem. Instead the intention is to illustrate the idea about how a connection between the scale-space primal sketch with other modules can be done. The application supports the claim we make, that *if the image contains significant structures, which stand out from the surrounding, then they are extracted in such a way that the output information from the scale-space primal sketch is useful for further processing.*

4.1 Edge Detection at a Proper Scale

By intention we use a simple edge detector, since we want to illustrate how this problem becomes easier once the earlier mentioned scale and region information is available. The image is smoothed to the scale level given by a significant blob from the scale-space primal sketch. Then a non-maximum-suppression step is performed on Sobel gradients to get thin edges. To suppress spurious noise points at the finest levels of scale we accept only edge segments of length exceeding 2 pixels.

4.2 Matching Blobs to Edges

Our matching procedure for associating blobs with edges is based on three conditions about spatial coincidence:

1. Geometric coincidence. The edge segment should "encircle" or be "included" in the blob. Let B be the set of pixels contained in the support region of a blob and let E be the set of pixels covered by an edge segment. Further, given any region R define the quantities x_{min}, x_{max} by
$$x_{min}(R) = \min_{(x,y)\in R} x; \quad x_{max}(R) = \max_{(x,y)\in R} x; \tag{4}$$
and the quantities y_{min}, y_{max} analogously. An edge segment E will be regarded as a matching candidate of a blob B if
$$\begin{aligned} x_{min}(E) \le x_{max}(B); \quad x_{max}(E) \ge x_{min}(B) \\ y_{min}(E) \le y_{max}(B); \quad y_{max}(E) \ge y_{min}(B) \end{aligned} \tag{5}$$
In order to reduce the directional sensitivity of this criterion it is suitable to require that similar conditions hold also in a coordinate system rotated by 45 degrees.
2. The edge segment should not be too far from the blob boundary. In other words, it should comprise at least some pixel located near the boundary of the blob. We state the condition
$$\min_{(x_E,y_E)\in E;\,(x_B,y_B)\in B} \sqrt{(x_E-x_B)^2+(y_E-y_B)^2} \le \frac{d(t)}{2} \tag{6}$$
where $d(t)$ is a typical spatial length at the current level scale, here, set to the square root of an experimentally determined blob area, $A_m(t)$, at that scale, see [19] for further details.

[6] The scale-space primal sketch contains much more information than is presented in this rudimentary output. For instance we have not illustrated the registered blob bifurcations in scale-space. Nor have we shown or made use of the hierarchical relations between blobs at different levels of scale induced by the blob events. This information is however explicit in the computed representation.

3. An edge segment closely related to one particular blob should not be associated with other blobs. We compute a Voronoi diagram of the grey-level blob image at the given level of scale and require the edge segment to have at least one pixel in common with the Voronoi region associated with the grey-level blob.

For an edge segment to be accepted as a matching candidate of a blob it must satisfy all these criteria. Hence, the matching is relatively restrictive. But again, the situation is improved by the fact that it is performed at a proper scale. Once we know that a spatial region has given rise to a large blob at some level of scale, it seems very improbable that conflicting edges should appear in the edge image at the same level of scale, since most interfering fine-scale structures ought to be suppressed by the scale-space smoothing. The main problem with this matching procedure is that it does not comprise any mechanism for breaking up long edge segments into shorter ones. This means that the edge segments at coarser levels of scale may be very long, and spread far from the boundary of the actual blob, see the example in Figure 10.

4.3 Blob-Initiated Edge Focusing

The basic principle in edge focusing [2, 3] is to detect edges at a coarse level of scale in scale-space and to trace them to finer level of scale. In this application we initiate the focusing procedure from several scale levels, since the significant blobs from the scale-space primal sketch manifest themselves at different scales. Hence, we pre-sort the blobs in decreasing scale order. We start with the coarsest scale blob, detect edges at that level of scale and match the obtained edges to the blob. This gives the input for the focusing procedure, that follows these edges to the scale given by the second blob. The edge detection and matching steps are repeated at this new level of scale and the resulting edges are added to the output from the previous focusing step. This new edge image, serves as input for another focusing procedure, tracing the edges to the next finer level of scale etc.

Figure 10 illustrates some steps from this composed procedure applied on the telephone and calculator image. In order to reduce the number of blob hypotheses treated we have used a threshold on the significance value. The final result is shown in the lower right corner of Figure 10. Observe that this method, which we call *blob-initiated edge focusing*, is not just another edge detector, but that the resulting edge segments are more meaningful entities, since a coarse grouping has already been performed and the edges are associated with significant dark blobs and explicit scale information. Note that label information for the edge segments can be easily inherited during the edge focusing process.

With this experiment using the output from the scale-space primal sketch to control an edge focusing procedure, we have eliminated two of the tuning parameters in the edge focusing algorithm, namely the initial scale for edge detection and the threshold on gradient magnitude. What remains undetermined is the stop scale down to which the edge focusing should be performed. In this work it has been throughout set to $t = 1$, a scale where the discretization effects start to become important, see also [18]. It seems plausible that some further guidance for this selection could be obtained by studying the behaviour of the focused edges in scale-space, compare with the classification of diffuse edges in [3].

This integration of the two algorithms exemplifies the previously mentioned guidance of the focus-of-attention. Note that the processing initiated by the scale-space primal sketch is performed only for a small subset of the image data. Hence, the

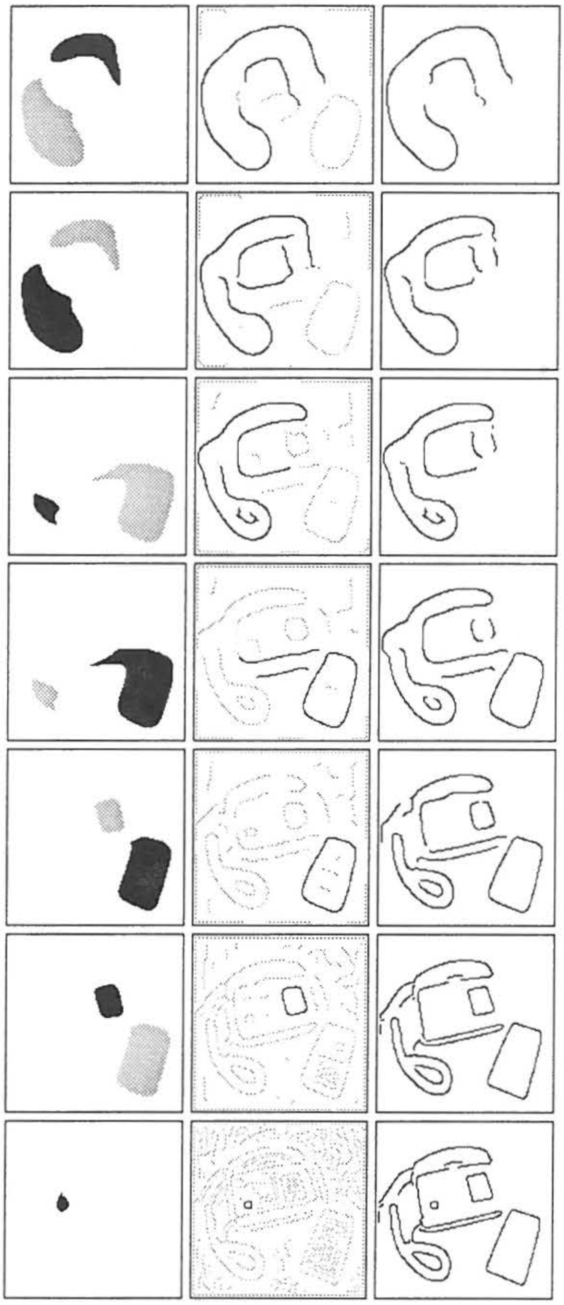

Figure 10: *Illustration of the composed blob-edge focusing procedure for the telephone and calculator image. The left column shows the active blob hypothesis. Its blob support region has been marked with black. The middle column shows the edge image at the scale level given by the previous blob. The matching edge segments have been drawn black, while the other ones are grey. The right column shows the result after focusing, just before a new blob is considered. The image in the lower right corner shows the output result, i.e. edges related to the dark blobs in the image. The scale and significance values for the different blobs are from top to bottom (101.6, 14.1), (50.8, 252.8), (32.0, 11.4), (25.4, 660.9), (14.3, 40.8), (6.4, 63.6) and (1.3, 13.2) respectively.*

resulting method closely follows the idea of a "focused beam", derived by Tsotsos [27] from complexity arguments.

5 Histogram Analysis and Junction Classification

The scale-space primal sketch is well suited for automated cluster detection, since it is designed for detection of bright blobs on dark background and vice versa. Hence, it lends itself as a natural module for peak detection in algorithms based on histogramming techniques. Although it is well-known that histogram-based segmentation hardly can be expected to work globally on entire images, such methods can often give useful results *locally* in small windows, where only a few regions of different characteristics are present. In Figures 11-12 we illustrate how it can constitute a helpful tool in such histogram modality analysis of multispectral data. We have accumulated histograms[7] of the chroma information and used the scale-space primal sketch to detect peaks and clusters in the histograms. We see that the extracted blobs induce a meaningful partitioning of the histogram corresponding to regions in the image with distinctly different colours.

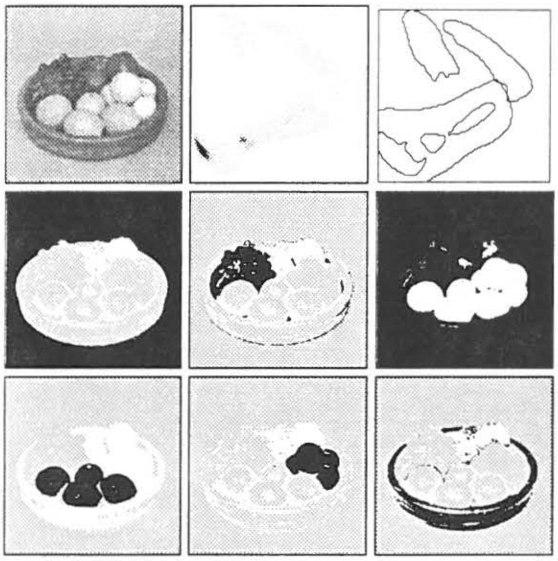

Figure 11: *Histogram-based colour segmentation of a fruit bowl image: (a) Grey-level image. (b) Histogram over the chroma information. (c) Boundaries of the 6 most significant blobs detected by the scale-space primal sketch. (d)-(i) Backprojections of the different histogram blobs to the original image (in decreasing order of significance). The pixels corresponding to the various blobs have been marked in black. (The region in Figure (f) is the union of the regions in Figures (d), (e) and (i)). The significance values of the accepted blobs were 42.6 (background), 8.3 (grapes), 3.6 (oranges), 3.1 (apples), 3.0 (bowl) and for the rejected blobs 2.0 and less (2.0, 1.9, 1.8, 1.4, 1.3, 1.1, 1.1, 1.1, ...).*

Of course, there is a decision finally to be made about which peaks in the histogram should be considered significant. However, we believe that the significance values given by the scale-space blob volumes reflect the situation in a manner useful for such reasoning. In these examples (single) thresholds were set manually in "gaps" in the sequences of significance values, see the captions of Figures 11-12.

[7] The colour images have been converted from the usual RGB format to the CIEu^*v^* 1976 format, which separates the intensity and the chroma information. The histogram is formed only over the chroma information, ignoring the intensity information.

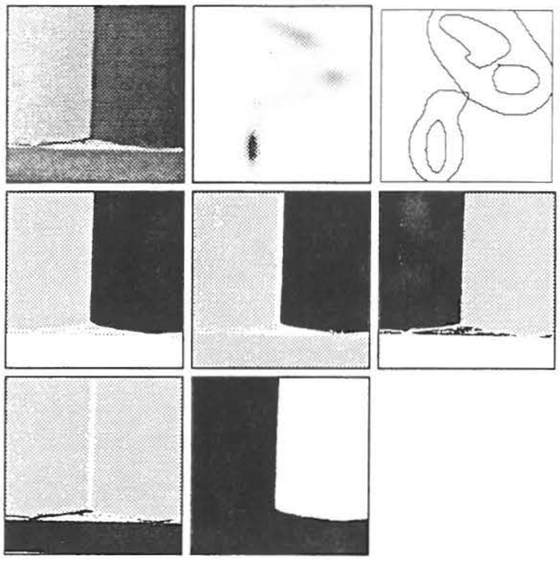

Figure 12: *Similar histogram-based colour segmentation of a detail from an office scene. The image shows a small window from a bookcase with two binders (yellow and blue) on a shelf made of wood. (a) Grey-level image. (b) Chroma histogram. (c) Boundaries of the 5 most significant blobs. (d)-(h) Backprojections. The displayed blobs have significance 187.9 (blue binder, large blob), 173.7 (blue binder, small blob), 170.1 (yellow binder), 80.6 (shelf) and 66.7 (yellow binder and shelf). As we see, two blobs corresponding to the blue binder have been detected. This is a common phenomenon in the scale-space primal sketch, that arises because a large blob merges with a small (insignificant) blob and forms a new scale-space blob. Two such duplicate blobs corresponding to the yellow binder (significance 18.0) and the shelf (significance 17.9) have been suppressed. The remaining blobs had significance 2.5, 2.0, 2.0, 1.2, 1.2, 1.2, 1.1 and less.*

It can also be noted that this peak detection concept will be less sensitive to quantization effects in the histogram acquisition than many traditional peak detection methods. The problems due to too fine a quantization in the accumulator space will be substantially reduced, since the scale-space blurring will lead to a propagation of information between different cells. Thus, even though the original histogram may have been acquired using "too many and too small" accumulators, large scale peaks will be detected anyway, since the contents of their cells will merge to large scale blobs in scale-space after sufficient amounts of blurring.

Finding peaks in histograms is a problem arising in many contexts. Brunnström et al. [6] have shown that a reliable classification of junctions can be performed by analysis of the modality of the intensity and directional histograms, combined with an active focusing procedure, which varies the resolution and the size of the window around the candidate junction point over which the histograms are formed. A critical step in this method is to determine the modality of the appearing local histograms. We believe that a similar scale-space primal sketch for one-dimensional signals would be a useful tool in such analysis.

The junction classification algorithm also has to deal with an indirect scale problem, since one needs to determine the initial window size for the focusing procedure. We are currently investigating how a combination of regions extracted from the scale-space primal sketch of the grey-level image can be used together with interest points produced by some other method. A matching step has to be performed. Then we propose that an approximate window size can be set given that we know the size of the blob and have the scale information explicit. The precise window size can finally be determined from a stability analysis, where the window size is varied around the initial value. We are also investigating if and how the blob information can assist already in the phase of detecting interest points. Some preliminary results about this integration of the scale-space primal sketch as to control the junction classification method are reported in [6, 20].

6 Summary and Discussion

The representation that we build is similar to the primal sketch suggested by Marr [22, 23], in the sense that it is a two-dimensional representation of the significant grey-level structures in the image. It is also computed under extremely weak assumptions. However, besides that it is a region-based and not an edge-based representation it is more *qualitative*, without strong assumptions about the shape of the primitives. Moreover, the proposed representation consists of coarse features like blobs represented at multiple scales, and allows for

- Automatic detection of salient (stable) scales, if they exist.
- Ranking of events in order of significance.
- Generation of hypotheses for grouping and segmentation.

This implies that candidate regions for further processing are generated, as well as information about the scale. We see that the representation gives clues to subsequent analysis and, hence, it can guide focus-of-attention mechanisms.

We have also tried to demonstrate the effects of one as we believe very promising methodology, namely that *simple methods and qualitative reasoning can perform surprisingly well if the treatment is performed at a proper scale and over an appropriately selected region in space*, provided that the sampling density is sufficient to clearly resolve the phenomena we are studying[8]. For instance, the primitives (grey-level blobs, scale-space blobs and edges from non-maximum suppression) used for extracting image structure were defined solely in terms of singularities and geometric properties in scale-space. These entities can be very noise sensitive when considered only at one single (arbitrarily selected) level of scale. However, here we have shown that they can give robust results if combined with a careful treatment of the scale issue.

The underlying principle we use when extracting image structures is that structure should be invariant under transformations in parameter space. Our method consists of three steps: (1) Vary the parameters systematically. (2) Detect (locally) stable states (intervals). (3) Choose a representative descriptor as an abstraction of each stable interval and pass only this information on to the higher level modules. In this specific case the parameter we vary is the scale parameter in the scale-space representation. However, we believe that this kind of methodology could be applicable also in other types of situations.

6.1 Relations to Previous Work

There are, of course, earlier attempts to derive similar representations of the grey-level landscape. Rosenfeld and his co-workers, see e.g. [11, 26] have studied blob detection in pyramids e.g. using relaxation methods. Blostein & Ahuja [5] detect texture elements based on zero-crossings and use multiple scales and a significance measure based on a background noise assumption. There is also the wealth of literature on pyramids, see e.g. Levine [17], Crowley & Sanderson [8], Crowley & Parker [9] and Burt & Adelson [7]. The texton theory proposed by Julesz, see [13, 14] and [28], essentially also treats the blob detection problem. There are finally a number of representations based on intensity changes, e.g. Marr [22], Bergholm [2, 3] and Watt [29] and approaches working at higher levels, like Saund's [25] token based symbolic grouping. Of interest is also the approach by Haralick et.al. [12], which allows a more detailed representation, but only at a single spatial scale.

Our approach differs from these in three important aspects. First, our representation can be seen as preceding e.g. the edge-based schemes in that it selects the appropriate scales and regions, intrinsically defined by the image itself, in a complementary data-driven manner. Secondly, it is a hierarchic representation of the structures at *all scales* in the image with explicit information about their significance and relations. Finally, it is derived in a formal way using the well-defined notions of scale-space, which allow a precise analysis of the behaviour of structure. Hence, we can study how events at different scales can be related in a well-defined manner.

6.2 Scale-Space Experiences

Let us point at a few aspects of scale-space representations that have been given little or insufficient attention in the literature and that have to be dealt with in creating a representation of the sort we want. First, it is noteworthy, that the amount of noise in real images usually leads to a large number of local extrema. These extrema may disappear rather early, provided that they are subsumed by some more prominent extremum. However, if they occur in a region with smoothly varying grey-levels, then they will exist over a large range of scale. This effect is alleviated, but not remedied, by annihilation between nearby noise extrema. Even though their amplitudes decrease rapidly it is not clear that one can set a threshold on objective grounds. This problem is related to the issue of estimating the noise level in an image, which hardly can be addressed without any constraining assumptions, like e.g. in Voorhees & Poggio [28].

Another property is that images of scenes of even moderate complexity rarely have a global scale, at which all structure above the noise level is present. This aspect is explicitly dealt with in our representation. Bischof & Caelli [4] treat the same question for zero-crossings of the Laplacian. However, their measure of stability seems to be more arbitrary.

In this context one can ask more generally what is the relation between our representation and zero-crossings of the second derivative. We suggest that our representation, with extrema and their extents, captures important structure. The zero-crossings will not always be localized in the same places and, therefore, not represent the same structure. Watt [29], in fact argues that the extrema of the second derivative, and not the zero-crossings, should be used to pick up information about intensity discontinuities. We feel that this question should be investigated further.

7 Conclusions

We have presented a multi-scale representation of grey-level image structure similar to the primal sketch idea. It can be used for extraction of important regions from an image in a solely bottom-up data-driven way, without any a priori assumptions about the shape of the primitives. The representation, which is essentially free from tuning parameters and ad hoc error criteria, gives a qualitative description of the grey-level landscape with information about *approximate location*, *spatial extent* and

[8] In this work we have assumed that the images have been acquired with sufficient resolution. How this issue can be further coped with is developed in [6], where we perform active focusing to acquire new images of higher resolution with the purpose of simplifying the task of classifying junctions.

an *appropriate scale* for relevant regions in the image. In other words, it generates coarse but safe segmentation cues, and can be used as a hypothesis generator for higher-level processes. We have demonstrated how such information can serve as a guide to an edge detection scheme working at a locally adapted level of scale and that it is applicable for automatic cluster detection and modality analysis of histograms. More generally, we find this approach useful for guiding the focus-of-attention and tuning other low-level processes.

8 Acknowledgments

We would like to thank Kjell Brunnström for the cooperation with the work on junction classification, Fredrik Bergholm for useful comments when implementing the edge focusing algorithm, Harald Winroth for many discussions, and Hans Tråvén for providing the fruit bowl image.

This work was performed partially under the Esprit-BRA project Insight. The support from the National Swedish Board for Technical Development, STU, is gratefully acknowledged.

A Appendix: Algorithmic Aspects

When building the scale-space primal sketch representation of an image there are a few computational aspects that need to be taken into account. The algorithm for creating this suggested data structure consists of two major modules; a grey-level blob detection algorithm and an *adaptive* scale linking and refinement procedure. In this appendix we will just attempt to sketch on those parts. A more detailed description is given in [19].

A.1 Grey-Level Blob Detection

From the definition of a grey-level blob one easily realizes that the following basic properties hold in the classification of the bright blobs of a discrete signal. To simplify the presentation let the notation "higher-neighbour" stand for "neighbour pixel having a higher grey-level value". Further, the concept "background" will mean a pixel which has been classified as not belonging to a blob.

1. If a pixel has no higher-neighbour then it is a local maximum and will be the seed of a blob.
2. Else, if it has at least one higher-neighbour which is background then it cannot be part of a blob and must be background.
3. Else, if it has more than one higher-neighbour and if those higher-neighbours are parts of different blobs then it cannot be a part of a blob, but must be background.
4. Else, it has one or more higher-neighbours, which are all parts of the same blob. Then it must also be a part of that blob.

Starting from these properties sequential or parallel blob detection algorithms can be easily constructed, see [19] Section 4.

A.2 Blob Linking Strategy

Linking blobs across scales could be a potential source to difficult matching problems, since blobs may move, disappear, merge or split when the scale parameter changes. However, the notion of a scale-space with a continuous scale parameter gives us a simple way to circumvent these problems in many cases, since the scale step may be varied at will. If one is confronted with a problematic matching situation then the matching difficulties can often be avoided by refinement of the scale sampling. If the scale step is adaptively made just fine enough it should be trivial to judge which grey-level blobs belong to the same scale-space blob. Based on this idea the blob linking between two levels of scale can be performed based on spatial coincidence. A straightforward strategy is to start with a relatively fine initial sampling in scale[9] and then for each pair of scale levels traverse all pixels and for each point investigate if it is included in a blob both at the lower scale and at the higher scale. If so, the lower blob is registered as a match candidate of the higher blob, and the higher blob is registered as a match candidate of the lower blob. What remains to decide is when a blob match should be accepted. In our current implementation we perform a refinement in scale each time an unclear matching situation occurs, and accept matches in principle only when all blob events between the two scale levels can be classified as either of the cases: single link, blob annihilation, blob merge or blob split. Further details about this algorithm can be found in [19], Section 5.

[9] Corresponding to about $\frac{1}{3}$ - $\frac{1}{2}$ octave in t at coarse levels of scale

References

[1] Bajcsy R., Rosenthal D.A. (1980) "Visual and Conceptual Focus of Attention", in Tanimoto S., Klinger A. (Eds), *Structured Computer Vision*, Academic Press, N.Y., 133-147.

[2] Bergholm F. (1987) "Edge Focusing", *IEEE-PAMI*, **9:6**, 726-741.

[3] Bergholm F. (1989) "On the Content of Information in Edges and Optical Flow", *TRITA-NA-P8904*, Royal Inst. of Technology, Stockholm.

[4] Bischof W.F., Caelli T (1988) "Parsing Scale-Space and Spatial Stability Analysis", *CVGIP*, **42**, 192-205.

[5] Blostein D., Ahuja N. (1987) "Representation and Three-Dimensional Interpretation of Image Texture", *1:st ICCV*, 444-449.

[6] Brunnström K., Eklundh J.O., Lindeberg T.P. (1990) "On Scale and Resolution in the Analysis of Local Image Structure", *1:st ECCV*, 3-12.

[7] Burt P.J., Adelson E.H. (1983) "The Laplacian Pyramid as a Compact Image Code", *IEEE-COM*, **9:4**, 532-540.

[8] Crowley J., Sanderson A. (1987) "Multiple Resolution Representation and Probabilistic Matching of 2D Gray-Scale Shape", *IEEE-PAMI*, **9:1**, 113-121.

[9] Crowley J., Parker A. (1984) "A Representation of Shape Based on Peaks and Ridges in the Difference of Low-Pass Transform", *IEEE-PAMI*, **6:2**, 156-169.

[10] Ehrich R.W., Lai P.F. (1978) "Elements of a Structural Model of Texture", *PRIP*, IEEE CS Press, 319-326.

[11] Gross A.D. (1986) "Multiresolution Object Detection and Delineation", *TR-1613*, Computer Vision Lab., Univ. of Maryland.

[12] Haralick R.M., Watson L.T., Laffey T.J. (1983) "The Topographic Primal Sketch", *Int. J. Rob. Res*, **2**, 50-72.

[13] Julesz B., Bergen J. (1983) "Textons, The Fundamental Elements in Preattentive Vision and Perception of Textures", *Bell Syst. Tech. J.*, **62:6**, 1619-1645.

[14] Julesz B. (1986) "Texton Gradients: The Texton Theory Revisited", *Biol. Cyb.*, **54**, 245-251.

[15] Koenderink J.J., van Doorn A.J. (1984) "The Structure of Images", *Biol. Cyb.*, **50**, 363-370.

[16] Koenderink J.J., van Doorn A.J. (1986) "Dynamic Shape", *Biol. Cyb.*, **53**, 383-396.

[17] Levine M.D. (1980) "Region Analysis Using a Pyramid Data Structure", in Tanimoto S., Klinger A. (Eds), *Structured Computer Vision*, Academic Press, N.Y., 57-100.

[18] Lindeberg T.P. (1990) "Scale-Space for Discrete Signals", *IEEE-PAMI*, **12:3**, 234-254.

[19] Lindeberg T.P., Eklundh J-O. (1990) "Computation of a Scale-Space Primal Sketch", *TRITA-NA-P9005, Submitted.*

[20] Lindeberg T.P., Eklundh J-O., (1990) "Guiding Early Visual Processing with a Scale-Space Primal Sketch", *TRITA-NA-P9013, Submitted.*

[21] Lowe D.G. (1985) *Perceptual Organization and Visual Recognition*, Kluwer Academic Publishers, Boston.

[22] Marr D. (1976) "Early Processing of Visual Information", *Phil. Trans. of the Royal Society (B)*, 27S, 483-524.

[23] Marr D. (1982) *Vision*, Freeman, San Francisco.

[24] Pizer S.M., Gauch J.M., Lifshitz L.M., Oliver W.R. (1988) "Image Description via Annihilation of Essential Structures", *Tech Rep 88-001*, Comp. Sci. and Radiology, Univ. of North Carolina, Chapel Hill.

[25] Saund E. (1989) "Adding Scale to the Primal Sketch", *CVPR-89*, 70-78.

[26] Sher A.C., Rosenfeld A. (1987) "Detecting and Extracting Compact Textured Objects Using Pyramids", *TR-1789*, Computer Vision Lab., Univ. of Maryland.

[27] Tsotsos J.K. (1989) "Analyzing Vision at the Complexity Level", *to appear in Behavioural and Brain Sciences.*

[28] Voorhees H., Poggio T. (1987) "Detecting Textons and Texture Boundaries in Natural Images", *1:st ICCV*, 250-258.

[29] Watt R. (1988) *Visual Processing: Computational, Psychophysical and Cognitive Research*, Lawrence Erlbaum Associates, London.

[30] Witkin A.P. (1983) "Scale-Space Filtering", *7:th IJCAI*, 1019-1022.