A New Approach For Line Extraction and its Integration in a Multi-Scale, Multi-Abstraction-Level Road Extraction System *

Helmut Mayer and Carsten Steger

Lehrstuhl für Photogrammetrie und Fernerkundung
Technische Universität München
Arcisstraße 21, 80290 München, Germany
e-mail: helmut@photo.verm.tu-muenchen.de

Lehrstuhl für Bildanalyse und wissensbasierte Systeme
Technische Universität München
Orleansstraße 34, 81667 München, Germany
e-mail: stegerc@informatik.tu-muenchen.de

Abstract: *Road extraction is an active area of research. Because roads are elongated objects they can be described symbolically by means of lines. This paper shows, theoretically as well as in practice, why and under which circumstances lines extracted from an image are an abstract representation of roads. Especially the importance of scale is emphasized and the connection between multiple scales and multiple abstraction levels by means of so-called scale-space events is derived. The investigations are based on a sophisticated model for line extraction for which the scales where events occur can be determined analytically. Results of a multi-scale multi-abstraction-level road extraction system, which is based on the theoretical and practical investigations, are presented.*

1. Introduction

A lot of work exists in the area of the extraction of roads from satellite and aerial imagery [1, 9, 3, 25]. Of special importance for road extraction in small-scale imagery is line extraction, for which a new approach is presented. It is integrated in a multi-scale multi-abstraction-level road extraction system [3, 25]. For the link of multiple scales and multiple abstraction levels theoretical considerations [18] which are based on findings about scale-space events [16] are given together with practical examples

*This work was supported by Deutsche Forschungsgemeinschaft under grant No. Eb 74/8–2

for the integrated system. The presented system can be used for aerial imagery as well as for forthcoming high resolution satellites with down to 0.8 m ground resolution [8]. The goal of this paper is to give some theoretical as well as practical considerations which support the use of multiple scales.

When one thinks about the world one uses abstract concepts which describe a complex physical world. The concept road, for instance, is associated with different types of road (e.g., highway or country road), cars, concrete, the freedom of speed, etc. This shows that concepts like road are very hard to define in a consistent and formal way which can be implemented as a computer program (for problems due to imprecise concepts cf. [2]). The central issue for a formal definition is knowledge representation [22]. Another aspect is the knowledge which has to be represented. It has to come from an application (e.g., Photogrammetry, Remote Sensing, or GIS). Issues which have been addressed only recently are the importance of context [26] or the combination of different kinds of information (information fusion). A special case is the fusion of different scales (resolutions). This can aid the interpretation, as has been shown in [3, 25]. Scale in the context of image sequences was treated in [4, 21]. It will be shown how the semantics of objects is linked to scale, i.e., how scale can help to define concepts.

In knowledge representation as well as in psychology [14] distinctions are made between the description of single objects and their spatial relations as well as between class and instance processing. Single objects can be used together with their spatial relations as substructure, i.e., parts, of more complex objects. Besides the fact that this constitutes a hierarchy of objects based on the part-of relation, there also is an *abstraction* linked to this. A settlement has, for instance, a substructure made of buildings, roads, etc. But in addition to this it also has a new, more abstract, quality, e.g., its own size or characteristics (shopping, recreation, etc.).

Opposed to abstraction, which deals with symbols, scale-space theory [12, 16] is concerned with subsymbolic signal (here: image) information. The standard procedure is that a scale-space is constructed by smoothing the original image with Gaussian kernels of successively increasing width. A property of scale-space theory is that, additionally to the continuously evolving smoothing of the image, events (*scale-space events*) occur. These events are annihilation, merge, split, and creation of extrema. Because most structure in an image, like points, edges, or lines, are related to extrema, this means that also the structure is changed significantly.

One of the interesting properties of the human visual system is that it appears to represent information on multiple scales [14]. In many cases there is a close interaction of larger and smaller scale. An interesting question in this context is how these scale-space events are related to the abstraction of symbols describing objects in the image. By and large, two things can happen simultaneously when an image is transformed by means of smoothing from a larger to a smaller scale:

- The information content of the image is reduced by eliminating regions and edges: Noise as well as meaningful information are removed due to scale-space events.
- The removal of meaningful information is synonymous with the elimination of substructure

(parts) of objects and results in simplification. The interesting feature is that the loss of substructure often emphasizes the objects. If abstraction is defined as the increase of the level of simplification and emphasis, this is synonymous with an abstraction of objects.

This abstraction by means of smoothing will be studied using the extraction of lines to recognize roads as an example. In recent years progress has been made to extract (curved) lines from digital images [13]. In this paper a new approach [23, 24] is presented. It is based on differential geometric properties of the image function. For each pixel, the second order Taylor polynomial is computed by convolving the image with the derivatives of a Gaussian smoothing kernel. Line points are required to have a vanishing gradient and a high curvature in the direction perpendicular to the line. The use of the Taylor polynomial and the Gaussian kernels leads to a single response of the filter to each line. Furthermore, the line position can be determined with sub-pixel accuracy and the position of the edges can be determined reliable even when the brightness on both sides of the line is different. An analysis about the scale-space behavior of the most common type of line profile, the bar-shaped profile, is given. Furthermore, two types of profiles modeling various common interactions of objects and roads are analyzed in scale-space. From this analysis, useful conditions for scale-space events, e.g., the merging of two lines, can be derived.

After an introduction into abstraction and scale-space a sophisticated approach for the detection of lines and the determination of their width is presented. The scale-space behavior of the line extraction with special focus on scale-space events is analyzed and a model for roads represented as a semantic net is condensed from it. Results for a multi-scale, multi-abstraction-level road extraction system, which is based on these considerations, are presented, and conclusions are given.

2. Abstraction and Scale-Space Events

2.1. Abstraction and Models

There are a lot of definitions for the term abstraction (e.g., [6]). In this paper a special notion of abstraction is used. It is defined in the context of image understanding where symbols are mapped to portions of images. The description by means of the symbols has to be structured. Additionally it has to be simplified, emphasis has to be laid on important things, and others have to be neglected. *Abstraction* is therefore defined as the increase of the degree of simplification and emphasis.

As has been shown in the introduction, this is also connected with parts which construct the *substructure* of an object. Because the notion of a term has to be defined to enable a sound reasoning, the part-of relation as well as the specialization and the concrete relation are defined in this paper in terms of semantic networks following [20] or [19].

Simplification and emphasis are important characteristics of models which are used to achieve the mapping of symbols and image data. This means that abstraction is also an implicit but integral part of models, and models are the critical basis of image understanding. They can be considered as the "theory" part of the theoretical framework of [17] as well as the conceptual level of the levels of knowledge representation of [5]. Explicit models have to be the foundation for every project in image understanding, because they can give reasons for deficits of an approach. Without an explicit model, i.e., if a system is only based on heuristics, no sound analysis of errors is possible and therefore the further development is hampered. The typical development will start with constructing a model from experience. The model is implemented and tested, and according to the arising problems the model is improved. This is done iteratively.

2.2. Scale-Space Events

Images are analog representations, representing non- or subsymbolic information by means of a homomorphism: The represented facts are contained in the representation. Relations between objects of the real world are transfered without loss of structure into relations of the representation.

The things which can be seen in an image are dependent on the scale (physical resolution). In a Landsat-TM image it is impossible to recognize a single human being on the ground whereas in an aerial image of scale $1:4000$ this is easy.

Recently tools have been created for handling the concept *scale* in a formal manner. The main idea is the creation of a multi-scale representation by a one-parameter family of derived signals, where fine-scale information is successively suppressed [16]. Data is systematically simplified and finer-scale details, i.e., high-frequency information is removed. The *scale parameter* $\sigma \in \mathbf{R}_+$ is intended to describe the current level of scale.

The representation at coarser scales are often given by a convolution of the given signal with Gaussian kernels of successively increasing width

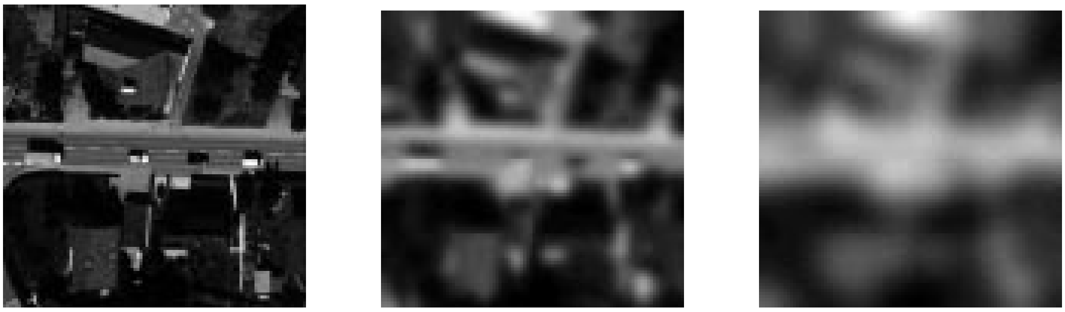
$$L_\sigma(x) = g_\sigma(x) * f(x), \tag{1}$$

where $g : \mathbf{R} \times \mathbf{R}_+ \setminus \{0\} \to \mathbf{R}$ is the (one dimensional) Gaussian kernel

$$g_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad . \tag{2}$$

Another valid way to define a scale-space is in terms of mathematical morphology [10, 15]. This scale-space has different characteristics than the Gaussian scale-space. A comparison and integration of both scale-spaces was given for the simplified application of closed contours in [11]. In order to not further complicate matters, the remainder of the paper is restricted to the Gaussian scale-space.

To describe the structure in an image so-called *blobs* as the (zero order) scale-space features can be used [16]. Blobs are smooth regions which are brighter or darker than the background, stand out from the surroundings and are therefore closely linked to extrema in the image.



(a) (b) (c)

Figure 1: Car on road in Gaussian scale-space — Images: (a) original image, (b) $\sigma = 1.2\,\mathrm{m}$, (c) $\sigma = 3\,\mathrm{m}$

In the process of smoothing the image four different discrete events can happen to a blob: annihilation, merge, split, and creation. Whereas split and creation are not too likely to occur (examples are given in [16]), merge and annihilation of blobs are quite common. But blobs are only one means to represent the information content of an image. More commonly used representations are regions, edges, and lines. In a first approximation most of the events which can happen to blobs will happen to regions, their delimiting edges, or lines as well. Taking all this into account the term *scale-space event* is used for the remainder of this paper referring to events of regions, edges and lines.

Figure 1 (original image) shows a car on the road, The image is gradually smoothed (from left to right) until the car cannot be recognized any more. The level of smoothing where this happens depends on the level of noise in the image as well as on the closeness to other objects. Linked to this phenomenon are the *inner scale* and *outer scale* [12]. The outer scale is the (minimum) size of a window which completely contains the object while the inner scale is the scale at which substructures of an object begin to appear.

How abstraction can occur by means of change of scale, how this is linked to scale-space events, and how this can be used in practical applications is shown in the following using the extraction of roads as line-type objects from aerial imagery as an example. For line extraction a sophisticated approach is presented, which was described in detail in [23, 24]. Therefore only a brief summary is given. However, details that pertain to the scale-space analysis and to object abstraction are described more in-depth.

3. Detection of Lines and Line Width

3.1. Detection of Line Profiles in 1D

By far the most commonly occurring line type in aerial images is a line with a bar-shaped profile. The ideal line profile of width $2w$ and height h is given by

$$f_b(x) = \begin{cases} h, & |x| \le w \\ 0, & |x| > w \ . \end{cases} \tag{3}$$

This type of line occurs for many "interesting" objects in aerial images, e.g., roads or rivers, since these structures often exhibit a relatively flat cross section compared to the height of their edges.

In order to detect lines with a profile given by (3) the image should be convolved with the derivatives of the Gaussian kernel (2). This leads to a description of the behavior of the derivatives of the bar-shaped profile in scale-space:

$$r_b(x,\sigma,w,h) = h(\phi_\sigma(x+w) - \phi_\sigma(x-w)) \tag{4}$$

$$r'_b(x,\sigma,w,h) = h(g_\sigma(x+w) - g_\sigma(x-w)) \tag{5}$$

$$r''_b(x,\sigma,w,h) = h(g'_\sigma(x+w) - g'_\sigma(x-w)) \ . \tag{6}$$

In order to detect line points it is sufficient to determine the points where $r'_b(x,\sigma,w,h)$ vanishes. However, it is usually convenient to select only salient lines. A useful criterion for salient lines is the magnitude of the second derivative $r''_b(x,\sigma,w,h)$ in the point where $r'_b(x,\sigma,w,h) = 0$. Bright lines on a dark background will have $r''_b(x,\sigma,w,h) \ll 0$ while dark lines on a bright background will have $r''_b(x,\sigma,w,h) \gg 0$. When the scale-space behavior of $r''_b(x,\sigma,w,h)$ is analyzed it can be seen that the second derivative will not take on its maximum negative value for small σ. Furthermore, there will be two distinct minima in the interval $[-w, w]$. It is, however, desirable for $r''_b(x,\sigma,w,h)$ to exhibit a clearly defined minimum at $x = 0$ since the selection of salient lines is based on this value. It can be shown that

$$\sigma \ge w/\sqrt{3} \tag{7}$$

has to hold for $r''_b(x,\sigma,w,h)$ to exhibit a unique minimum. Furthermore, it can be shown that $r''_b(x,\sigma,w,h)$ will have its maximum negative response in scale-space for $\sigma = w/\sqrt{3}$ (see [23]).

From (3) and (5) it is also evident that a line is bounded by an edge on each side of the line. Hence, to detect the line width the edge points to the left and right of the line point need to be extracted. Their position is given by the solutions of $r''_b(x,\sigma,w,h) = 0$, where $r'''_b(x,\sigma,w,h) < 0$, i.e., the points that exhibit a maximum in the absolute value of the gradient.

Figure 2 shows the behavior of the line and edge positions for a line with $w = 1$ and $h = 1$ for $x \in [-4, 4]$ and $\sigma \in [0, 4]$. It can be seen that the line position will always be in the correct place for all σ. The edge position, i.e., the line width, will be extracted most accurately for small σ; it will be extracted too large if σ is chosen too large for the desired w. However, the map that describes the relationship of the extracted line width and the true line width can be inverted. Therefore, it is possible to extract the true line width for all choices of σ.

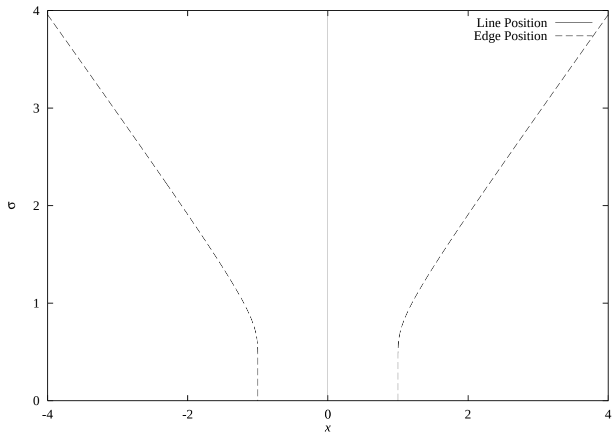


Figure 2: Scale-space behavior of the line and edge positions of the bar-shaped line with $w = 1$ and $h = 1$ for $x \in [-4, 4]$ and $\sigma \in [0, 4]$.

3.2. Detection of Lines in 2D

Curvilinear structures in 2D can be modeled as curves $s(t)$ that exhibit the characteristic 1D line profile f_b in the direction perpendicular to the line, i.e., perpendicular to $s'(t)$. Let this direction be $n(t)$. This means that the first directional derivative in the direction $n(t)$ should vanish and the second directional derivative should be of large absolute value [23].

The only problem that remains is to compute the direction of the line locally for each image point. In order to do this, the partial derivatives r_x, r_y, r_{xx}, r_{xy}, and r_{yy} of the image will have to be estimated. This can be done by convolving the image with the appropriate 2D Gaussian kernels. The direction in which the second directional derivative of $r(x, y)$ takes on its maximum absolute value will be used as the direction $n(t)$. This direction can be determined by calculating the eigenvalues and eigenvectors of the Hessian matrix

$$H(x, y) = \begin{pmatrix} r_{xx} & r_{xy} \\ r_{xy} & r_{yy} \end{pmatrix} \quad . \tag{8}$$

The eigenvector corresponding to the eigenvalue of maximum absolute value gives the direction perpendicular to the line. As in the 1D case, the second directional derivative along n, i.e., the maximum eigenvalue, can be used to select salient lines. In the case of discrete images this procedure must be modified to compute the derivatives by discrete convolution, sub-pixel calculation of the zero crossing of r' along n, and checking whether this location lies within the current pixel's boundaries. Details are given in [23].

To detect the width of the line, for each line point the closest points in the image to the left and to the right of the line point, i.e., along $-n$ and n, where the absolute value of the gradient takes on its maximum value need to be determined. Equation (7) shows that it is sensible to search for edges only in a restricted neighborhood of the line. Ideally, the maximum distance for the search would be $\sqrt{3}\sigma$.

In order to ensure that most of the edge points are detected, the current implementation uses a slightly larger maximum distance of 2.5σ.

3.3. Scale-Space Events and Abstraction

Scale-space events and their link to abstraction can be analyzed in the same manner as the bar-shaped profile. Two lines that are close and approximately parallel to each other frequently occur in aerial images as the two road segments of a divided highway. The corresponding profile is given by

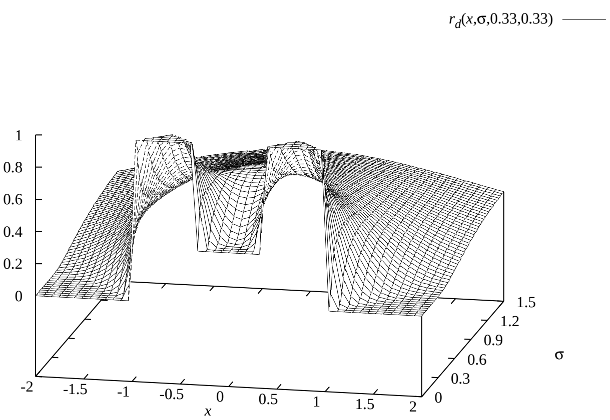
$$f_d(x) = \begin{cases} 1, & v < |x| \leq 1 \\ h, & 0 \leq |x| \leq v \\ 0, & |x| > 1 \ . \end{cases} \tag{9}$$

It models two lines of the same intensity 1 and width $1 - v$ separated by a gap of width $2v$ and intensity h. This results in no loss of generality since lines of arbitrary width can always be analyzed by multiplying the results obtained with this profile with the real width w. The same holds for any real intensity of the lines.

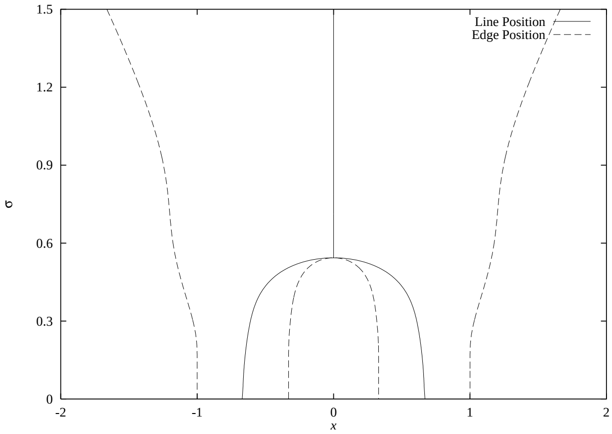
Figure 3 displays the scale-space behavior of the line and edge positions for a line of this type with $v = 0.33$ and $h = 0.33$, i.e., the lines and the gap have the same width and the gap is relatively dark. From Fig. 3(a) it can be seen that the gap between the lines will vanish as σ increases. Figure 3(b) shows the line and edge positions of this smoothed profile, while in Fig. 3(c) they are drawn onto the profiles to make their location more intuitively clear. It is apparent that as σ increases the two distinct bright lines and the dark line in the middle will merge into a single bright line. At the same position the edge points between the two lines must vanish. For large σ only a single bright line will be extracted. It can be shown that the two lines will merge for

$$\sigma = \sqrt{\frac{1}{2} \frac{v^2 - 1}{\ln(v(1 - h))}} \ . \tag{10}$$

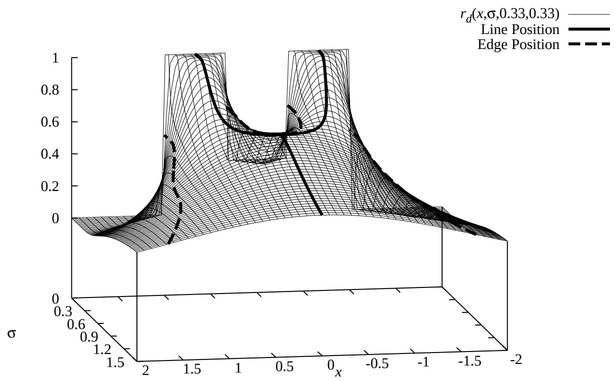
This is important in two ways. Firstly, this means that a scale-space event has occurred and (10) gives the exact σ when the merge will happen. Secondly, it means that merely by increasing σ one can abstract two road segments into a more abstract highway. The location of the abstracted object will automatically reflect the center axis of the total highway. (10) can be used to select an appropriate scale to ensure the merge happens or to prevent it from happening, depending on the application. For example, in the German topographic medium scale GIS ATKIS highways must be stored as two separate road segments. A typical three lane highway has a total width of 37.5 m with a median strip of width 4 m, which is usually relatively dark. If (10) is to be used, one has to set $v = 0.106666$ and can select $h = 0.2$, for example. This means that the two lanes will merge at $\sigma = 0.448154$, which corresponds to $\sigma \approx 8.40\,\mathrm{m}$. Therefore, in this case σ must be chosen from the interval $[4.84\,\mathrm{m}, 8.40\,\mathrm{m}]$,



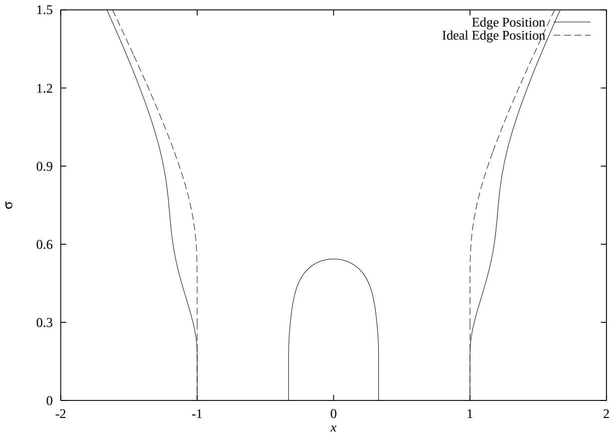
(a) Scale-space behavior



(b) Line and edge position



(c) Line and edge position in scale-space



(d) Extracted and ideal edge position

Figure 3: Scale-space behavior of two parallel lines.

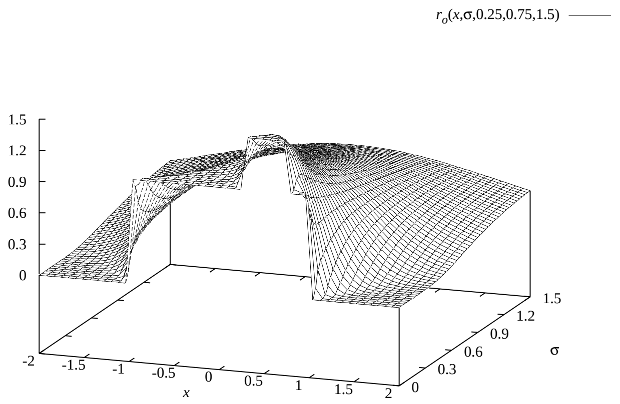
where the lower bound is obtained from (7). To conclude this example, Fig. 3(d) compares the edge positions of this line type with those of a bar-shaped profile of the same width. It can be seen that the main difference is the outward bulge that develops in the area where the two lines merge. Therefore, the width correction mentioned in Sect. 3.1 will still yield meaningful results.

An object on the line obscuring a part of it happens, for example, for cars on a road or ships on a river. The object is typically not in a location symmetrical to the center axis of the line. This can be modeled by the following profile:

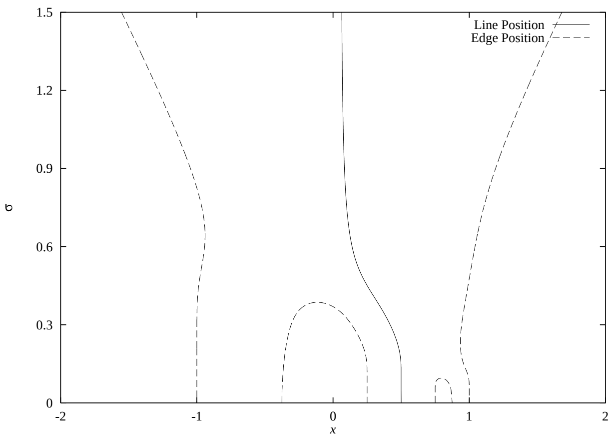
$$f_o(x) = \begin{cases} 1, & -1 \le x < l \vee r < x \le 1 \\ h, & l \le x \le r \\ 0, & |x| > 1 \ . \end{cases} \tag{11}$$

For $h < 1$ the object is darker than the line, and for $h > 1$ it is brighter. Because the scale-space behavior is somewhat simpler, the latter case is considered first.

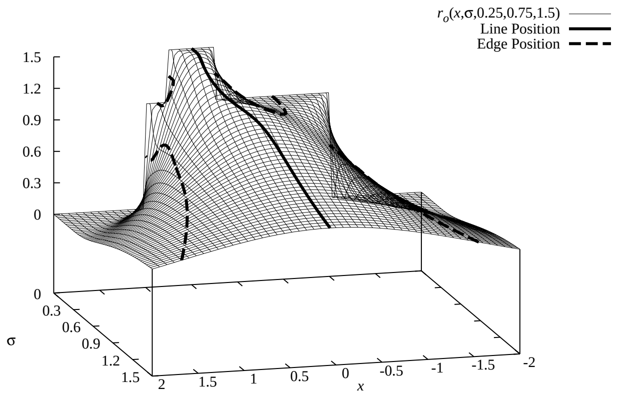
Figure 4 shows this behavior for $l = 0.25$, $r = 0.75$, and $h = 1.5$. This models, for example, a bright



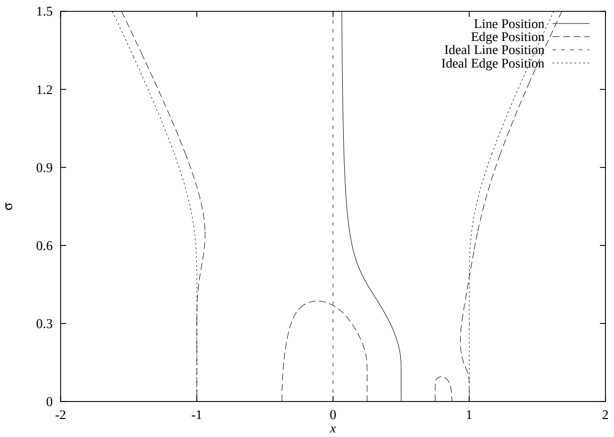
(a) Scale-space behavior



(b) Line and edge position



(c) Line and edge position in scale-space

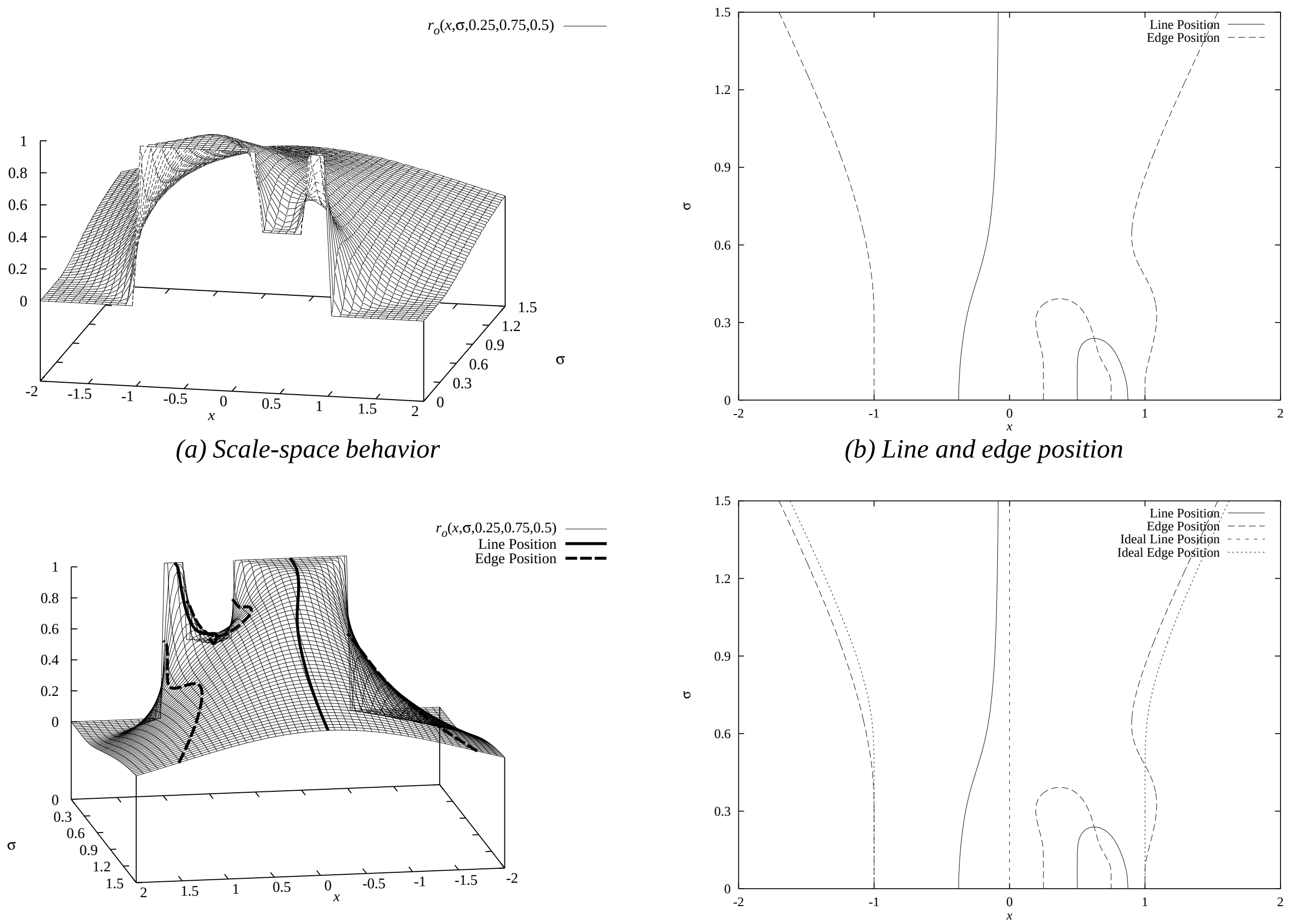


(d) Extracted and ideal line and edge position

Figure 4: Scale-space behavior of a line with a bright disturbance on it.

car on the right lane of a road. It is intuitively clear that only one line should be detected for all σ, and Figs. 4(a) and (b) show that this is indeed the case. Again, Fig. 4(c) displays the line and edge positions mapped onto the smoothed profiles, while Fig. 4(d) compares them to the corresponding positions of an undisturbed profile. For small σ the extracted line position will be the one of the bright object, while for large σ it will correspond to the center axis of the line. Therefore, by increasing σ one can eliminate the car from the road (scale-space event). It can also be seen that the two edges corresponding to the bright object will vanish along with the flat inflection points on the undisturbed part of the line. In general, they will have vanished if σ is chosen according to (7). This means that the extracted line width will correspond to the true line width, and not the disturbance. However, as Fig. 4(d) shows, the line and edge positions will be asymptotically slightly shifted because of the disturbance. Nevertheless, the width correction will yield the true value of w.

Finally, the case of a line with a dark disturbance on it, e.g., a road with a dark car is examined. Figure 5 shows the scale-space behavior of this line type for $l = 0.25$, $r = 0.75$, and $h = 0.5$. It is apparent that the dominant left part of the line is extracted for all σ, while the narrow dark and bright lines on the left side will annihilate one another. The interesting thing to note in this case is that the



(a) Scale-space behavior (b) Line and edge position

(c) Line and edge position in scale-space (d) Extracted and ideal line and edge position

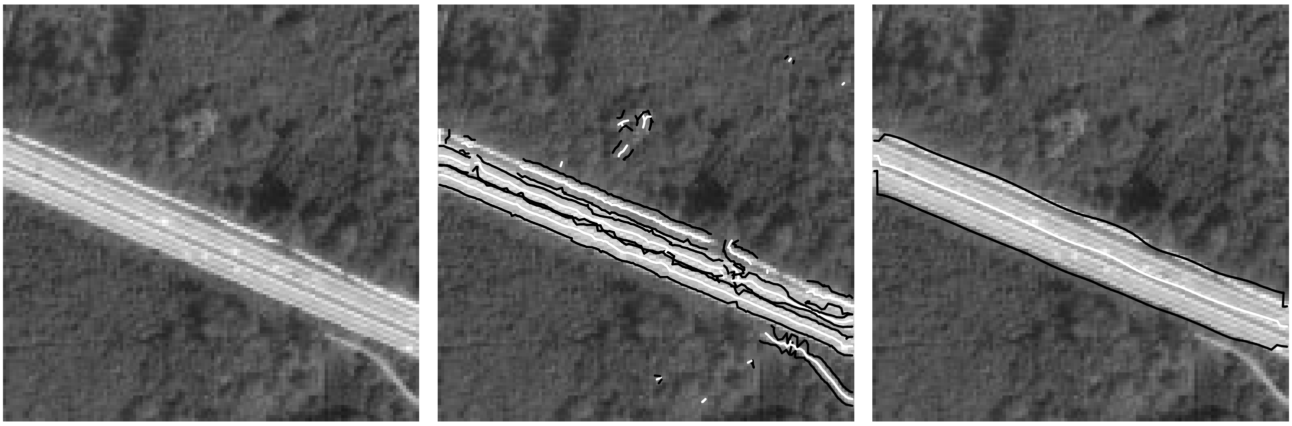
Figure 5: Scale-space behavior of a line with a dark disturbance on it.

corresponding edges exist for much larger σ before they vanish (scale-space event). Furthermore, the edge location crosses the line location at the point where the lines merge. This means that the edge must change its polarity at this point, and thus exhibit a flat inflection there. Again, for σ chosen according to (7) the unwanted substructure will in general already have vanished. Analogously to the previous case there will be an asymptotic shift in line and edge positions, but this time to the other side of the line.

In summary, scale-space events often correspond directly to abstraction on a symbolic processing level. The locations of the abstracted objects will correspond in a semantically meaningful way to the locations of the real objects. The analytical analysis of these model profiles allows the derivation of interesting and important scales (in terms of the sizes of objects) at which such an abstraction will take place.

4. Examples on Aerial Images

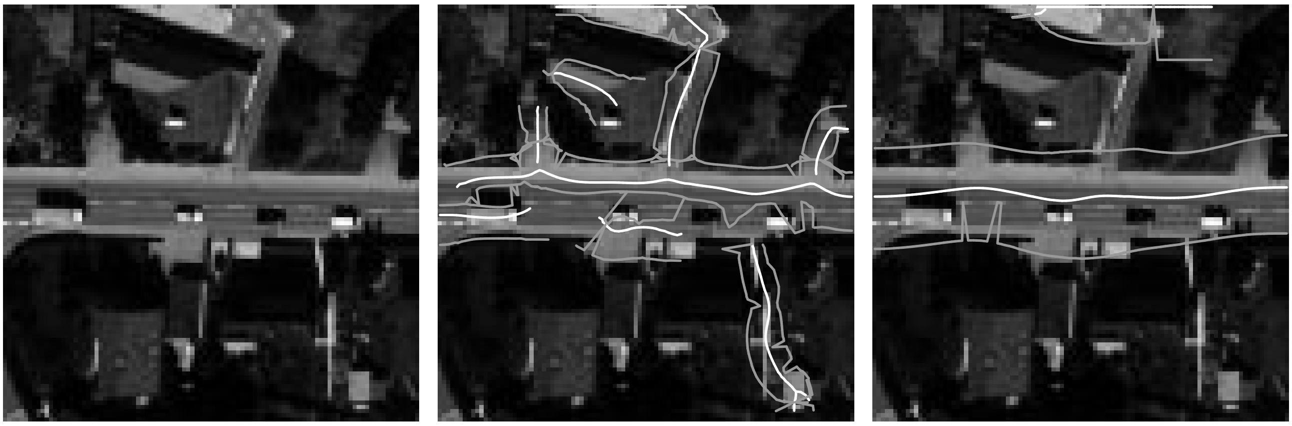
In this Section, examples of the scale-space events discussed above on aerial images are presented to demonstrate that they are valid and meaningful in practice as well as in theory.



(a) *Input image, resolution* $= 4\,\mathrm{m}$ (b) *Result with* $\sigma = 5.6\,\mathrm{m}$ (c) *Result with* $\sigma = 18\,\mathrm{m}$

Figure 6: Result of extracting lines and line width with different σ in an aerial image showing a highway.

Figure 6(a) shows an aerial image with a ground resolution of $\approx 4\,\mathrm{m}$ containing a highway. Figure 6(b) displays the results of extracting lines and their width with $\sigma = 5.6\,\mathrm{m}$. This results in the two lanes of the highway being extracted as two separate objects. A higher level vision module would have to reason that the two lines are road segments, and would have to group them together as a single object highway. In contrast, Fig. 6(c) shows the results obtained with $\sigma = 18\,\mathrm{m}$. As can be seen, only the center line of the highway is extracted. Thus, no further grouping is necessary to create the more abstract highway.



(a) *Input image, resolution* $= 0.6\,\mathrm{m}$ (b) *Result with* $\sigma = 2.4\,\mathrm{m}$ (c) *Result with* $\sigma = 5.4\,\mathrm{m}$

Figure 7: Result of extracting lines and line width with different σ in an aerial image showing a road with several cars.

A different type of scale-space events is exemplified in Fig. 7(a). Here, a road segment is disturbed by substructure. Several cars, some of them brighter than the pavement, and some of them darker, but all possessing large shadows, conceal the pavement. Moreover, the sidewalks are brighter than the pavement. This gives rise to a very complex scale-space behavior. In Fig. 7(b) the results of extracting lines with a scale of $\sigma = 2.4\,\mathrm{m}$ are displayed. This results in the extraction of several parallel and sometimes broken lines. Most notably, the van on the left side of the image gives rise to a separate line. Furthermore, the sidewalks are extracted as separate lines since they form the dominant lines at this scale. If σ is increased to $5.6\,\mathrm{m}$ only a single line is extracted for the whole road. Because of the complex configuration of objects on the road the extracted position and width of the road is not perfect, but still very useful, as can be seen from Fig. 7(c).

5. Model

From Sect. 4 it is evident that information is lost when going from large scale to small scale by eliminating regions and edges. In other words: The complex object highway of large scale is changed into a more abstract and stable linear road segment in smaller scale. Abstraction has occurred by means of elimination of substructure (annihilation) and merge of different parts separated by effects of noise and scale. In Fig. 8 the knowledge about the highway and the road segments is integrated into one model. The model is split into three levels. The *real world* level consists of the objects and their relations on a natural language level. In the large scale a road segment is constructed of a pavement and the markings (solid or dashed) and cars drive or park on it.

The objects in the *real world* level are connected to the objects in the *geometry and material* level by means of the *concrete* relation which connects *concepts* describing the same object on different levels, i.e., from different points of view. The *geometry and material* level is an intermediate level which represents the three-dimensional shapes of objects as well as their material [27]. This level has the advantage that it represents objects independent of sensor characteristics, in contrast to the *image* level.

In the medium scale the highway consists of road segments and the median strip. The road segments are linked to mostly straight bright lines in the *image* level via the mostly straight concrete or asphalt lines in the *geometry and material* level. The median strip is linked to a dark line via the grass strip.

The highway is in small scale linked to a mostly straight bright line in the *image* level which consists of the combined signal of the road segments and the median strip (cf. Fig. 6).

In contrast to this, the pavement of the large scale is linked to the elongated bright area in the *image* level by way of the elongated flat concrete of asphalt area in the *geometry and material* level. The markings are related to bright lines via colored lines and the car is a bright or dark region as the concrete of a painted box made of glass and metal.

Conceptually two things have happened here:

- The type of the geometrical representation has changed. Three lines have collapsed into one

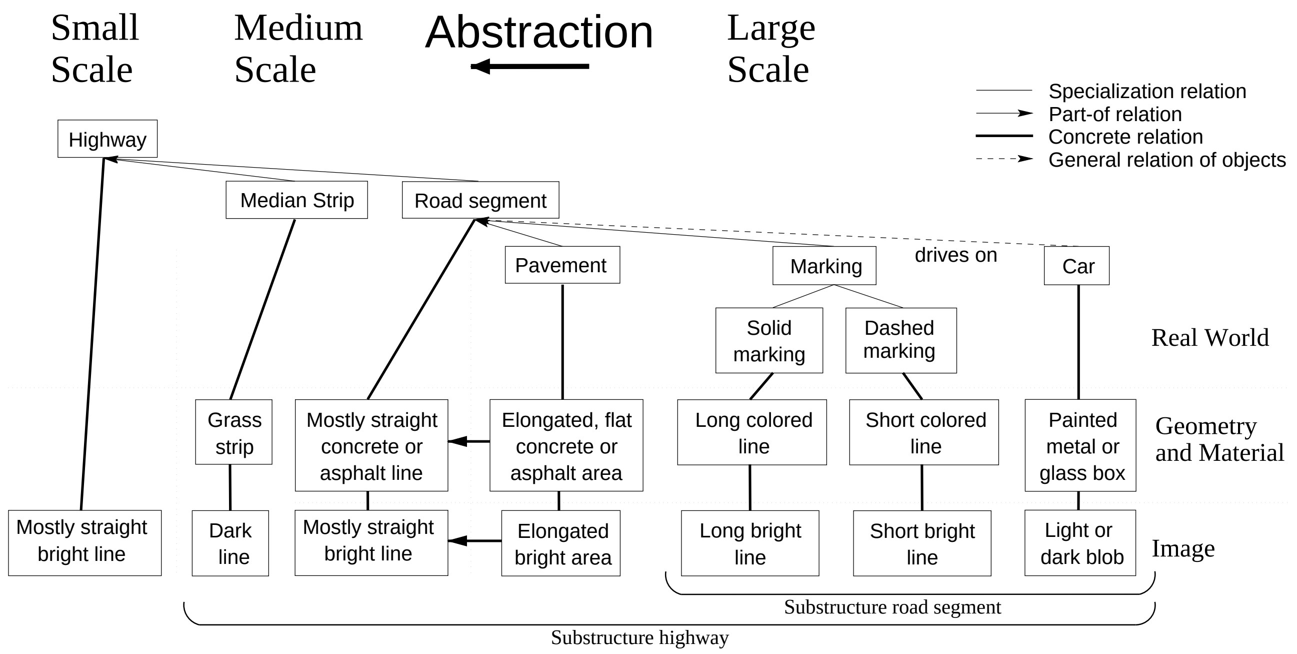


Figure 8: Model for highway and road segment (concepts are depicted in boxes)

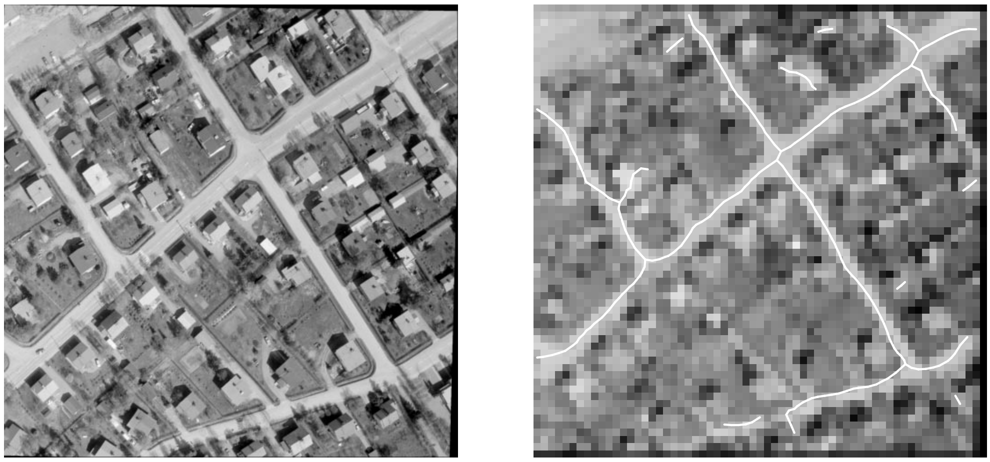
line. The elongated area has been condensed into a line.

- But more importantly, the substructure of objects in the larger scale (the markings, or the car on the road for the large scale, or the median strip in the small scale), has been eliminated.

This means that the complex object highway composed of median strip and road segment, which itself is made up from region-like pavements, and which has markings and cars on it, is changed into a more abstract linear highway. Whereas the large scale gives the detail, the medium and small scale add global information. If the information of both levels is fused, false hypotheses for roads are eliminated by using the abstract small scale information. Furthermore, details from the large scale (e.g., the correct width of the roads) are integrated in the result. With this, the advantages of both scales are merged.

6. Results of the Road Extraction System

In Fig. 9 the image “suburb” from the ISPRS Working Group III/2 dataset [7] is shown together with lines extracted from it in small scale. As was theoretically motivated in the last sections, this gives global information. Figure 10(a) presents the result for extracting parallel edges having bright regions in between which are homogeneous in the direction of the road (for details cf. [25]). In Figure 10(b) the result of fusing the two levels of scale/abstraction is shown. The global information in the small scale was used to eliminate some errors which have arisen in the large scale because of its too local context. In [3] it is shown that markings can be extracted from the image. This gives an additional



(a) *(b)*

Figure 9: Original image (a) and extracted lines (b) in small scale (reduced image)

(a) *(b)*

Figure 10: Bright elongated regions in large scale (a) and result of fusing lines in small scale and bright elongated regions in large scale (b)

level of scale/abstraction.

7. Summary and Conclusions

This paper shows how the abstraction of concepts can be linked to scale-space events. The notion of a scale-space event was studied on the example of line extraction and quantitative results were given and verified on practical examples. The results were then transferred to road extraction and a model for different scales was elaborated. Practical results for a system built on top of this theory were shown.

By and large, the outcome of this paper strongly recommends to use more than one scale for the recognition of objects. This not only has advantages in the performance but, as was shown, the

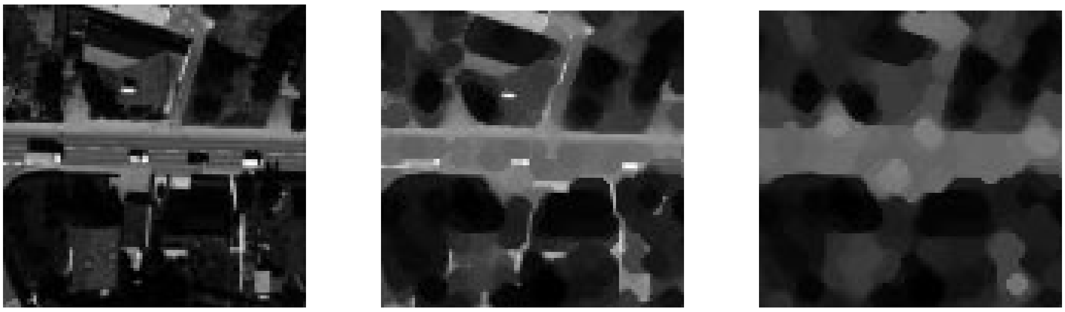


Figure 11: Car on road in morphological scale-space — Images (from left to right: original image, result of closing alone and of closing as well as opening - radius of disk $= 1.5\,\mathrm{m}$**)**

emphasis put on an object by smoothing away its substructure can be an important prerequisite for the recognition of objects ("only from a distance you can see clear"). A question which arises is the optimal scale to detect clues for an object. Small scales are especially suited for the detection of global structure and the formation of context while large scales add detailed/specific information for a detailed classification or verification of objects. Analytical means to answer the question were given for the extraction of lines.

In conclusion, a detailed analysis of the link of scale and abstraction of objects of an application (e.g., road or building) could be, besides the question of context, one of the central issues of modeling in image understanding in the future.

Problems which have to be answered in future are for instance the appropriateness of different types of scale-spaces. In Fig. 11 the image shown in Fig. 1 was put into a morphological scale-space (cf. [15]. When the image is closed and opened with a structuring element of suitable size the cars on the road can be totally eliminated.

References

[1] S. Airault, R. Ruskoné, and O. Jamet. Road Detection from Aerial Images: A Cooperation between Local and Global Methods. In *Image and Signal Processing for Remote Sensing*, volume 2315, pages 508–518. SPIE, 1994.

[2] H.-P. Bähr and A. Schwender. Linguistic Confusion in Semantic Modelling. In *International Archives of Photogrammetry and Remote Sensing*, volume (31) B6/VI, pages 7–12, 1996.

[3] A. Baumgartner, C. Steger, C. Wiedemann, H. Mayer, W. Eckstein, and H. Ebner. Update of Roads in GIS from Aerial Imagery: Verification and Multi-Resolution Extraction. In *International Archives of Photogrammetry and Remote Sensing*, volume (31) B3/III, pages 53–58, 1996.

[4] A.F. Bobick and R.C. Bolles. Representation Space: An Approach to the Integration of Visual Information. In *Computer Vision and Pattern Recognition*, pages 492–499, 1989.

[5] R.J. Brachman. On the Epistomological Status of Semantic Networks. In *Associative Networks*, pages 191–215, New York, USA, 1979. Academic Press, Inc.

[6] R.J. Brachman. What IS-A Is and Isn't: An Analysis of Taxonomic Links in Semantic Networks. *IEEE Computer*, 16(10):30–36, 1983.

[7] D. Fritsch and M. Sester. Test on Image Understanding. In *International Archives of Photogrammetry and Remote Sensing*, volume (30) 3/1, pages 243–248, 1994.

[8] L.W. Fritz. The Era of Commercial Earth Observation Satellites. *Photogrammetric Engineering & Remote Sensing*, 62(1):39–45, 1996.

[9] A. Grün, P. Agouris, and H. Li. Linear Feature Extraction with Dynamic Programming and Globally Enforced Least Squares Matching. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 83–94, Basel, Schweiz, 1995. Birkhäuser Verlag.

[10] P.T. Jackway and M. Deriche. Scale-Space Properties of the Multiscale Morphological Dilation-Erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):38–51, 1996.

[11] B. Kimia, A.R. Tannenbaum, and S.W. Zucker. Shapes, Shocks, and Deformations I: The Components of Two-Dimensional Shape and the Reaction-Diffussion Space. *International Journal of Computer Vision*, 15(3):189–224, 1995.

[12] J.J. Koenderink. The Structure of Images. *Biological Cybernetics*, 50:363–370, 1984.

[13] T.M. Koller, G. Gerig, G. Székely, and D. Dettwiler. Multiscale Detection of Curvilinear Structures in 2-D and 3-D Image Data. In *Fifth International Conference on Computer Vision*, pages 864–869, 1995.

[14] S. Kosslyn. *Image and Brain*. MIT Press, Cambridge, USA, 1994.

[15] U. Köthe. Local Appropriate Scale in Morphological Scale-Space. In *Fourth European Conference on Computer Vision*, volume I, pages 219–228, 1996.

[16] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, Boston, USA, 1994.

[17] D. Marr. *Vision*. Freeman and Company, New York, USA, 1982.

[18] H. Mayer. Abstraction and Scale-Space Events in Image Understanding. In *International Archives of Photogrammetry and Remote Sensing*, volume (31) B3/III, pages 523–528, 1996.

[19] H. Mayer. Using Real World Knowledge for the Automatic Acqusition of GIS-Objects from Scanned Maps. *GIS*, 9(2):14–20, 1996.

[20] H. Niemann, G.F. Sagerer, S. Schröder, and F. Kummert. ERNEST: A Semantic Network System for Pattern Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(9):883–905, 1990.

[21] M. Sester. Multiscale Representation for Knowledge Based Recognition and Tracking of Objects in Image Sequences. In *International Archives of Photogrammetry and Remote Sensing*, volume (28) 3/2, pages 868–888, 1990.

[22] J.F. Sowa, editor. *Principles of Semantic Networks – Explorations in the Representation of Knowledge*. Morgan Kaufmann Publishers, San Mateo, USA, 1995.

[23] C. Steger. Extracting Curvilinear Structures: A Differential Geometric Approach. In *Fourth European Conference on Computer Vision*, volume I, pages 630–641, 1996.

[24] C. Steger. Extraction of Curved Lines from Images. In *13th International Conference on Pattern Recognition*, 1996.

[25] C. Steger, C Glock, W. Eckstein, H. Mayer, and B. Radig. Model-Based Road Extraction from Images. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 275–284, Basel, Schweiz, 1995. Birkhäuser Verlag.

[26] T.M. Strat. Using Context to Control Computer Vision Algorithms. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 3–12, Basel, Schweiz, 1995. Birkhäuser Verlag.

[27] R. Tönjes. Knowledge Based Modelling of Landscapes. In *International Archives of Photogrammetry and Remote Sensing*, volume (31) B3/III, pages 868–873, 1996.