Automatic extraction of roads from aerial images based on scale space and snakes

I. Laptev[1], H. Mayer[2], T. Lindeberg[1], W. Eckstein[3], C. Steger[3], A. Baumgartner[4]

[1] Computational Vision and Active Perception Laboratory (CVAP), NADA, KTH, 100 44 Stockholm, Sweden
[2] Institute for Photogrammetry and Cartography, University of the Federal Armed Forces Munich, 85577 Neubiberg, Germany
[3] MVTec Software GmbH, Neherstrasse 1, 81675 Munich, Germany
[4] Chair for Photogrammetry and Remote Sensing, Technische Universität München, Arcisstrasse 21, 80290 Munich, Germany

Abstract. We propose a new approach for automatic road extraction from aerial imagery with a model and a strategy mainly based on the multi-scale detection of roads in combination with geometry-constrained edge extraction using snakes. A main advantage of our approach is, that it allows for the first time a bridging of shadows and partially occluded areas using the heavily disturbed evidence in the image. Additionally, it has only few parameters to be adjusted. The road network is constructed after extracting crossings with varying shape and topology. We show the feasibility of the approach not only by presenting reasonable results but also by evaluating them quantitatively based on ground truth.

Key words: Automatic road extraction – Aerial imagery – Snakes – Multi-scale – Evaluation

1 Introduction

Aerial imagery is one of the standard data sources for the acquisition of topographic objects, like roads or buildings for geographic information systems (GIS). Road data in GIS are of major importance for applications such as car navigation or guidance systems for police, fire services or forwarding agencies. Since the manual extraction of road data is time consuming, there is a need for automation.

In practical applications, human interaction will, at least for some time, be needed as a complement, resulting in semi-automatic methods. Vosselman and Knecht [36] and McKeown and Denlinger [24] present approaches based on road tracking which rely strongly on this interaction. They start from a given point and a given direction after extracting parallel edges or by extrapolating and matching profiles in high-resolution images. Other semi-automatic approaches search for an optimal path between a few given points. Grün and Li [12] and Merlet and Zerubia [25] connect points using dynamic programming. Fischler et al. [9] apply the F* algorithm based on line extraction from low-resolution images.

Neuenschwander et al. [26] use so-called "ziplock snakes" which provide another means to connect given points in the presence of obstacles. When more than one image is available, it is possible to track the line in 3D which constrains the path of the road and renders it possible to handle occlusions using robust optimization. In [10], Fua presents the model-based optimization of ribbon snake networks to improve coarsely digitized road networks.

Another way to tackle the problem is to start with fully automatic extraction and manually edit the result afterwards. This is the approach taken by us and many others. A survey on this topic can be found in Mayer et al. [22]. Zlotnick and Carnine [39] extend [24] to fully automatic extraction by finding starting points. De Gunst and Vosselman [7] and Bordes et al. [4] extract roads using a priori information in the from of GIS data. One of the recent approaches that are similar to the one proposed here is by Barzohar et al. [1]. It complements a low-level Markov-Random-Field model for the extraction of road seeds and the tracking of roads with a simple clutter and occlusion model and a Kalman filter. Another recent approach by Ruskoné et al. [28, 29] improves road extraction by modeling the context, i.e., other objects like shadows, cars or trees hindering or supporting the extraction of the road. A division of the context into spatially more global and more local parts was proposed in Baumgartner et al. [2] and Steger et al. [33]. In [28, 29], roads in urban areas are extracted by detecting and grouping of cars. Baumgartner et al. [2] and Steger et al. [33] utilize the scale-space behavior of roads; the road network is globally optimized, crossings are modeled, and markings are used to verify the existence of the road. Trinder and Wang [35] are on the same line, exploiting the scale-space behavior of roads as well as grouping of parallel segments. Related works on detecting ridge-like descriptors using multi-scale methods have been presented by Pizer et al. [27] and Lindeberg [20, 21]. Sing and Sowmya illustrate in [31] how grouping parameters might be learned. Boichis et al. present in [3] a more conceptual work for the extraction of crossings based on the Hough transform. The tracking of roads based on the A* algorithm is shown by Bückner [5]. In [34], Tönjes and Growe give an idea how information from dif-

ferent sensors, e.g., optical and radar, can be fused and help to improve the results.

In this paper, we present an extension of [17]. Our main idea is to take advantage of the scale-space behavior of roads in combination with geometrically constrained edge extraction by means of snakes. From the scale-space behavior of roads, we inferred to start by extracting lines at a coarse scale. Such lines are less precise but also less disturbed by cars, shadows, etc., than features at fine scales. The lines initialize ribbon snakes at fine scales, where roads often appear as bright, more or less homogeneous elongated areas. Optimized ribbons of constant width are accepted as *salient* roads (defined below). The connections between adjacent ends of *salient* roads are checked if they correspond to *non-salient* roads (defined below). As these roads are disturbed, the evidence for the road in the image can only be exploited when additional constraints and a special strategy focus the extraction. The constraints are low curvature and constant width of roads as well as the connectivity of the road network, i.e., the start and the end points are given. The strategy is to optimize the center of the ribbon snake using the "ziplock" method.

Our paper is organized as follows: In Sect. 2, we introduce model and inherent strategy essentially based on the scale-space behavior of roads and on ribbon snakes. Then, we define the terms *salient* road, *non-salient* road, and crossing. While Sect. 3 gives basic theory for ribbon snakes, we present the main part of the paper in Sect. 4. The extraction of *salient* roads supplies the information needed to extract the *non-salient* roads. We link *salient* as well as *non-salient* roads by crossings with varying shape and topology. Section 5 shows that the approach gives reasonable results for which we have evaluated the performance based on ground truth. We conclude the paper with a summary in Sect. 6.

2 Model and strategy

The appearance of roads in digital imagery depends on the spectral sensitivity [34] as well as on the resolution of the sensor. In this paper, we restrict ourselves to grey-scale images, and only scale dependencies are considered. Notably, a road can appear in different ways depending on the scale of observation. In images with coarse resolution (more than 2 m per pixel), roads appear mainly as lines, which establish a more or less dense network. In contrast, in high-resolution images (less than 0.50 m per pixel) roads are depicted as bright, more or less homogeneous elongated areas with almost constant width and bounded curvature.

In this paper, we take advantage of results from Mayer and Steger [23] and Lindeberg [18]. According to them, coarser scale lines representing road axes can be extracted in a stable manner from smoothed images[1] even in the presence of background objects such as trees or cars. Since our goal is to extract objects in the real world that have specific sizes, the "scale" in object space is of utmost importance. This scale depends on the resolution of the image, which we assume to be given, the type of scale space and the scale parameter. A typical order of magnitude for a descriptor at

a coarse scale is 2 m. This is equivalent to a σ of the Gaussian of 4 pixels at a ground resolution of 0.5 m. Seen from a symbolic point of view, finer scale substructures of the road, i.e., cars on the road or markings, as well as disturbances, like shadows or partial occlusions, are eliminated by considering representations at coarse scales. This can be interpreted as an *abstraction*, i.e., an increase of the level of simplification of the road. In contrast to the approach by Lindeberg [21] involving automatic scale selection, we assume that the width of the road is known, and that we can thereby determine the scale levels from given a priori information. Whereas coarse scales give global information which is especially suited for initial detection of the road, fine scales add detailed information which can be used to verify and complete the road network. If the information at both levels is fused, wrong hypotheses for roads can be eliminated by using the abstract coarse-scale information, while details can simultaneously be integrated from finer scales (for example, the correct width and precise position of the roads). In this way, relative advantages of using both scales can be combined.

For line extraction at coarse scales, we use the method of Steger [32], which builds upon previous works on multi-scale ridge detection [8, 15, 16, 20, 27], and is specifically adapted to road extraction, including an analysis of the scale-space behavior of roads. Baumgartner et al. [2] combine this framework with the following idea: the elongated areas of constant width describing roads at fine scales are extracted as parallel edges in the image using a local edge detector, such as the Canny operator [6], in combination with grouping (Sarkar and Boyer [30]). We found the following problems: the quality of edges varies due to noise, changes in radiometry of the road surface and its background, occlusions, etc. [11]. Therefore, the extraction of edges based on purely local photometric criteria often results in an incomplete detection of few significant edges or in the detection of many irrelevant edges. When these edges are grouped into parallelograms, they tend to be fragmented, even when lines from coarse scales are used in the grouping process and help to eliminate many wrong hypotheses. When parallel edges are linked based on purely geometric criteria, the precision is poor and wrong hypotheses are common [2].

Contrary to this, the framework of snakes introduced in Kass et al. [13] gives us the possibility to focus the extraction of edges by their known geometric properties. We do this by linking two edges into a "ribbon" defined by its center and its width. The ribbon is optimized using the "snake" concept [13] which leads to the so-called "ribbon snakes". Based on the scale-space behavior and the ribbon snakes, our model and the inherent strategy for the extraction of the road network can be summarized as follows:

– **Salient roads** have a distinct appearance in the image. A line at a coarse scale initializes the center of a ribbon snake at a finer scale. Criteria for the verification of a *salient* road are the constancy of the width and the homogeneity of the corresponding image region. Since man-made objects such as buildings have similar properties, the extraction of *salient* roads is restricted to homogeneous rural areas, given by a GIS or segmented by means of texture features. This is similar to [2], where the

---

[1] Here a Gaussian scale space [14, 19, 38] was used.
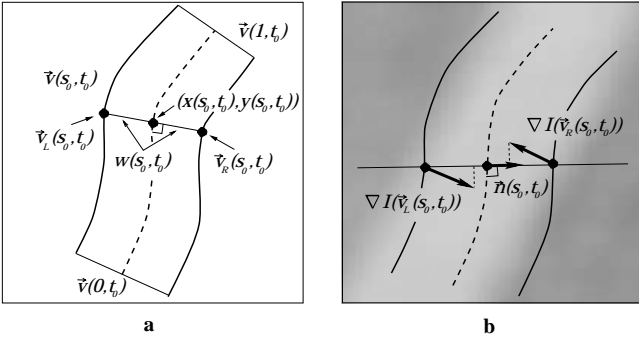
Fig. 1. a Parametric representation of the ribbon snake. Each slice of the ribbon $\mathbf{v}(s_0, t_0)$ is characterized by its center $(x(s_0, t_0), y(s_0, t_0))$ and width $w(s_0, t_0)$. Center and width define points $\mathbf{v}_L(s_0, t_0)$ and $\mathbf{v}_R(s_0, t_0)$ corresponding to the left and right side of the ribbon. **b** Image gradients for the left and the right side and their projection to the unit normal vector of the ribbon $\mathbf{n}(s_0, t_0)$

"context regions" *suburb_urban*, *forest*, and *open_rural* are introduced and the extraction of roads is restricted to the latter ones.

– **Non-salient roads** correspond to parts of the road network that are more or less disturbed in the image by shadows or partial occlusions. They can only be extracted in a top-down manner, when the start and end points are available. Thereby, we exploit the fact that all roads are connected into a global network. Taking adjacent ends of *salient* roads as start and end points, we connect them by ribbon snakes and optimize only the position of the center with the "ziplock" method. Finally, the width of the ribbon is optimized for the purpose of verification.

– **Crossings** link the road network together. At coarse scales, they correspond to junctions from the line extraction step [32]. We verify these hypotheses at a fine scale by expanding a closed snake around each junction and checking the connections between the outline of the crossing and its adjacent roads.

3 Ribbon snakes

The original snake concept introduced by Kass et al. [13] consists of curves with a parametric representation, where the position of the snake is optimized under a number of constraints. On one hand, the photometric constraints evoke the *image* forces that "pull" the snake to features in the image. On the other hand, the geometric constraints give rise to *internal* forces that control the shape of the snake and guarantee its piecewise smoothness. During optimization, the snake evolves from its initial position to a position where the forces compensate each other and the energy of the snake is minimized. This state implies that the snake is located in a way that best satisfies the desired properties.

For road extraction, we extended the original snakes with a width component analogously to Fua [10], leading to a ribbon snake defined as

$$\mathbf{v}(s, t) = (x(s, t), y(s, t), w(s, t)), \qquad (0 \leq s \leq 1), \qquad (1)$$

where s is proportional to the length of the ribbon, t is the current time, x and y are the coordinates of the centerline of the ribbon, and w is the half width of the ribbon measured perpendicular to the centerline. As shown in Fig. 1a, the centerline $(x(s, t), y(s, t))$ and the width $w(s, t)$ define the sides of the ribbon $\mathbf{v}_L(s, t)$ and $\mathbf{v}_R(s, t)$. Using this representation, the original expression for the *internal* energy of the snake still holds and the width is constrained by the same "tension" and "rigidity" forces as the two coordinate components. Differently from the original snake concept, however, the image forces for ribbon snakes are applied along the sides. When optimizing a ribbon to a bright road on a dark background, the image function P can be re-defined as the sum of the magnitudes of the image gradient along the curves $\mathbf{v}_L(s, t)$ and $\mathbf{v}_R(s, t)$. An even better way is to project image gradient onto the ribbon's normal $\mathbf{n}(s, t)$ and to constrain the projection to be positive at the left side of the ribbon and negative at its right side (cf. Fig. 1b). In this way, a correspondence between road sides and the sides of the ribbon is obtained. Using

$$P(\mathbf{v}(s, t)) = (\nabla I(\mathbf{v}_L(s, t)) - \nabla I(\mathbf{v}_R(s, t))) \cdot \mathbf{n}(s, t) \qquad (2)$$

for the image function, the expression for the total energy of the ribbon snake remains identical with the corresponding formula for the energy of the original snake:

$$E(\mathbf{v}) = - \int_0^1 P(\mathbf{v}(s, t)) ds +$$

$$+ \frac{1}{2} \int_0^1 \alpha(s) \left| \frac{\partial \mathbf{v}(s, t)}{\partial s} \right|^2 + \beta(s) \left| \frac{\partial^2 \mathbf{v}(s, t)}{\partial s^2} \right|^2 ds. \qquad (3)$$

The first term in Eq. 3 represents the *image* energy and the second corresponds to the *internal* energy. $\alpha(s)$ and $\beta(s)$ are arbitrary functions which determine the influence of the geometric constraints on the optimization.
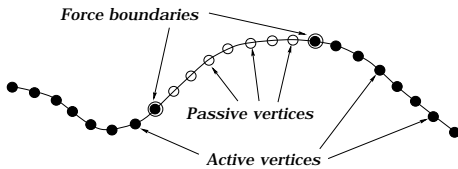
The application of ribbon snakes to fully automatic extraction requires that the balance between their *image* and *internal* energies has to be achieved automatically. When we assume that the initial estimate of the ribbon is close to the final solution, this can be enforced by substituting the functions $\alpha(s)$ and $\beta(s)$ in Eq. 3 with a factor λ [11]:

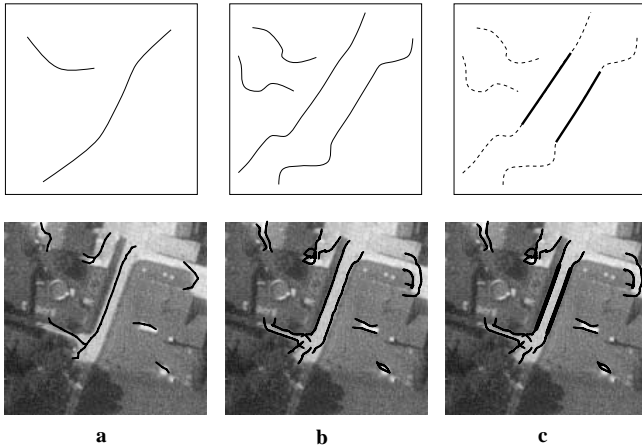$$\lambda = \frac{|\delta E_{img}(\mathbf{v})|}{|\delta E_{int}(\mathbf{v})|}, \qquad (4)$$

where δ is the variational operator. This simplification avoids the manual adjustment of $\alpha(s)$ and $\beta(s)$.

What is also important for the extraction is the ability to restrict and control the motion of the ribbon during optimization. For this reason, we "embed" the ribbon in a viscous medium and obtain the solution by minimizing the term $\int E(\mathbf{v}) + D(\mathbf{v}) dt$, where D is the dissipation functional $D(\mathbf{v}) = \frac{1}{2} \int_0^1 \gamma(s) |\mathbf{v}_t|^2 ds$ with the damping coefficient γ. For the discretized version of a ribbon with n vertices, γ can be derived from

$$\gamma = \frac{\sqrt{2n}}{\Delta} \left| \frac{\partial E(\mathbf{v})}{\partial \mathbf{v}} \right|. \qquad (5)$$
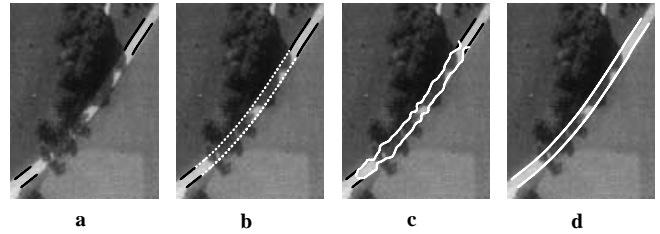
Fig. 2. Force boundaries divide the ziplock snake into two active parts and one passive part. During optimization the force boundaries are gradually propagated from the ends towards the center of the snake. Each time the force boundaries move, the snake is re-optimized
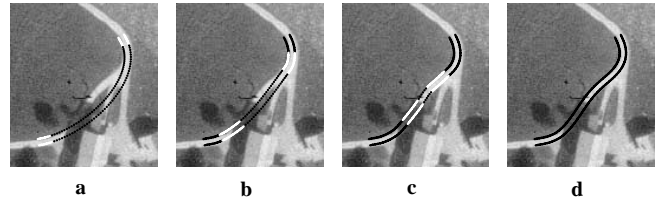

Fig. 4a–d. Extraction of *non-salient* roads. **a** Ends of extracted *salient* roads. **b** Extraction of optimal path. **c** Verification by optimization of width. **d** Acception of a hypothesis with low variation of width

Fig. 5a–d. Steps of the optimization of a "ziplock" ribbon. **a–c** *Black lines* in the middle of the ribbon indicate its passive part. *White parts* are currently optimized. *Black ends* indicate the result of the optimization so far. **d** Final result

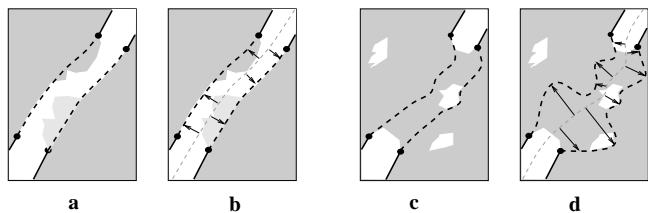## 4 Road extraction

4.1 Extraction of salient roads

According to the model and the strategy presented in Sect. 2 we start road extraction by detecting lines at a coarse scale [32]. As shown in Fig. 3a, line extraction often gives response to irrelevant image features. In order to separate them from the roads, i.e., to verify the roads, more evidence is required. At a fine scale, it is possible to determine the precise width of the structure corresponding to the line at a coarse scale. By evaluating the variation of the width along each hypothesis, we can mostly discriminate irrelevant structures, since their width is much more unstable than the width of roads.

Following this idea, we determine the width of image structures at fine scales by optimization of ribbon snakes. For each detected line, a ribbon is initialized such that its centerline coincides with the line and its width is zero. In order to obtain a rough approximation of the sought boundaries, we first optimize ribbons on images at coarse scales. We then re-optimize ribbons at fine scales in order to delineate the details of boundaries. During optimization, the width of ribbons is expanded to ensure the correct delineation of wide structures.

As shown in Fig. 3b, the optimized ribbons that correspond to wrong road hypotheses have a much stronger variation of width than the ribbon optimized at the *salient* road. Thus, thresholding the variance of width for short pieces of ribbons enables the selection of ribbon parts that correspond to roads with high probability (cf. Fig. 3c).

4.2 Extraction of non-salient roads

Typical reasons for gaps between *salient* roads are shadows or partial occlusions (cf. Sect. 2). To bridge the gaps, we

Fig. 3a–c. Extraction of *salient* roads. **a** Extraction of lines. **b** Optimization of ribbons' boundaries. **c** Selection of ribbon parts with constant width. Only the ribbon drawn with *thick lines* is accepted as a correct hypothesis for a *salient* road

As shown in [11] this ensures that the displacement of each vertex of the ribbon during one optimization step is on average of magnitude Δ. We choose a large value for Δ (20–30 pixels) at the beginning of optimization and then gradually decrease it toward the end of optimization in order to reduce the search space of the ribbon.

One well-known limitation of snakes is their sensitivity to initialization. Obstacles between the initial and the desired position of a snake often attract it and hinder a correct extraction. Neuenschwander et al. [26] address this problem and present a complementary optimization strategy called "ziplock snakes". Its idea is to divide the snake (or the ribbon snake) into two active parts and one passive part, as shown in Fig. 2. During optimization the force boundaries are gradually propagated from the ends toward the center of the snake. Image forces are only applied at active parts while the passive part is optimized with respect to purely geometric constraints. Given a correct initialization of the end points, the strategy ensures that the active parts stay always close to their desired positions, while irrelevant image structures around the passive part do not influence the extraction. We use the "ziplock" strategy for the optimization of the ribbon snakes while extracting *non-salient* roads in the next section.

Fig. 6a–d. Verification of hypotheses by expanding the width. **a**, **b** Verification of correct hypothesis. Since the centerline of the ribbon in **b** is fixed, the contrast at only one side is enough to stop the expansion. **c**, **d** Verification of wrong hypothesis. Random features result in a big variation of the width **d**
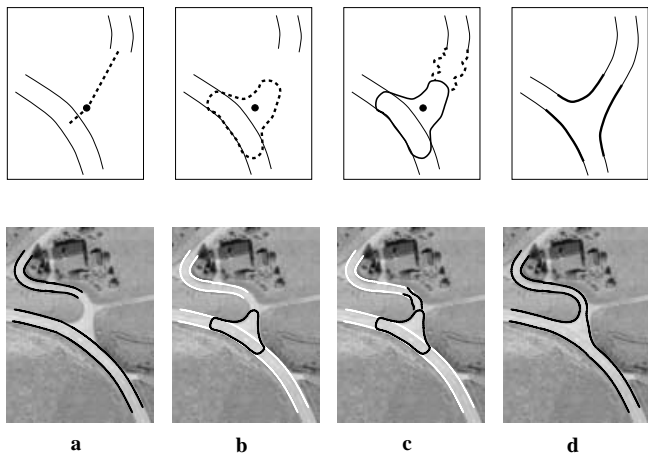

Fig. 7a–d. Extraction of crossings. **a** Selection of initial hypothesis. **b** Approximation of the outline of the crossing. **c** Verification of connections to adjacent roads. **d** Construction and connection of the crossing

connect two adjacent ends of *salient* roads with a road hypothesis, which is then verified based on homogeneity and the constancy of width (cf. Fig. 4). Regions corresponding to gaps show in many cases at least one traceable road side, although they may contain many other irrelevant edges which potentially disturb the road extraction. However, as can be seen from Fig. 4a, the curvature of these edges is mostly much higher than the curvature of road sides. We use this observation and apply low curvature constraint to discriminate irrelevant edges and to find correct sides of the road.

Each hypothesis for a *non-salient* road is initialized with the ends of previously extracted *salient* roads (cf. Figs. 4a, 5a). Note that the initial position of the ribbon can be far off from the correct position of a road. As mentioned in Sect. 3, this may cause the ribbon to lock to irrelevant image structures. To avoid this effect, we apply "ziplock" strategy and propagate the correct road extraction from the end to the middle of of the ribbon (cf. Fig. 5). At this point, we optimize
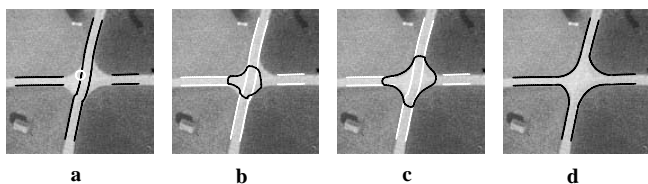
Fig. 8a–d. Steps of the optimization of the outline of a crossing. **a** Initial state of the closed snake. **b** and **c** 10th and 20th step. **d** Final result
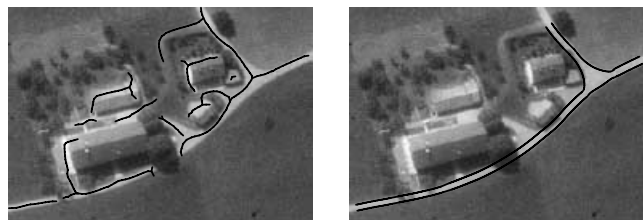
Fig. 9a,b. Road extraction for a complicated scene. **a** Line extraction. **b** Final result
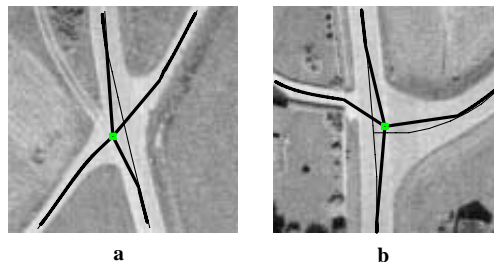
Fig. 10a,b. Centerlines of roads corresponding to the result of automatic road extraction (*thick lines*) compared to manually extracted reference data (*thin lines*). Manually and automatically extracted centerlines coincide outside the crossings except the left upper part of **b** where an approach road was not included to the reference data

only the centerline of the ribbon. The width of the ribbon is fixed and equals the width of the adjacent *salient* roads. Note, that due to the choice of the image function P in Eq. 2, the ribbon is insensitive to edges that cross perpendicular to its direction. This property is important because shadows and occlusions often result in this type of edges. At the same time, even a weak contrast of road sides attracts the ribbon and supports correct road extraction.

The procedure described so far connects *salient* roads, but does not really verify whether the detected connections correspond to roads or not. In order to verify hypothesized connections, we consider the following two steps. First, we evaluate the homogeneity of the region in the image which corresponds to the ribbon and accept ribbons with low variation of the image intensity as *non-salient* roads. This, however, will reject most of the shadowed and occluded roads and therefore we take a second step. The centerlines of the ribbons are fixed and only the width is expanded and optimized. As shown in Fig. 6a,b, given a correctly located road hypothesis, the fixed centerline together with the contrast on at least one road side stops the expansion of the width at the right place. On the contrary, random features localized close to wrong hypotheses will in most cases result in a large variation of a ribbon's width (cf. Fig. 6c,d). This observation implies that many wrong hypotheses can be rejected based on their large variation of width.

4.3 Extraction of crossings

The extraction of crossings is complicated since the variability of their shape makes it difficult to distinguish them from other objects. However, as shown below, the *salient* and *non-salient* roads greatly help to reduce the search space.

Fig. 11. Road extraction in larger image

Fig. 13. Road extraction in image "Marchetsreut" (centerlines only)

extracted roads (cf. Fig. 7c) and verify such connections using the method described in Sect. 4.2. Finally, we accept the crossing if at least one of the considered connections has been verified (cf. Figs. 7d, 8d).

The present strategy does not make any restricting assumptions on the shape and the topology of crossings. Therefore, crossings with different sizes and geometry and with varying numbers of adjacent roads can be extracted in a uniform way.

5 Results and their evaluation

Although the approach was developed for rural areas, in some cases, a successful recognition is also possible in built-up areas (cf. Fig. 9). As shown in Fig. 9a, line extraction alone does not detect some parts of roads and at the same time may result in more erroneous hypotheses than correct ones. Using the proposed approach, all wrong hypotheses are eliminated and additionally gaps caused by shadows are bridged (cf. Fig. 9b). Figures 11, 12 and 13 show the results for larger scenes. As can be seen, most of the roads are correctly delineated and connected into a network by the crossings. Typical failures occurred at roads that violate the assumption of constant width. This is common in urban areas, e.g., in the upper part of Fig. 11. However, the extraction fails even for some roads in rural areas. For instance, the main road running at the lower right corner of Fig. 12 could not be extracted since its width appears to be very unstable when examining it at a fine scale.

An evaluation of the results according to [37] is shown in Table 1. The evaluation is based on the comparison of the extracted road centerlines to ground truth, i.e., manually plotted road axes used as reference (cf. Fig. 10). Figure 11 is a part of the image "Erquy", i.e., Fig. 12, which can be described as "flat, agricultural, difficult", whereas the image "Marchetsreut" (Fig. 13) is "flat, agricultural, easy".

Fig. 12. Road extraction in image "Erquy" (centerlines only)

In order to generate initial hypotheses for crossings, we take advantage of two observations:

– During the extraction of *non-salient* roads most of the gaps in *salient* roads caused by crossings will be bridged (cf. Figs. 5d and 7a).
– As exemplified in Figs. 3a and 9a, line extraction often results in junctions inside crossings. However, a large number of these junctions correspond to irrelevant features.

Combining these two observations, we hypothesize crossings at junctions that are adjacent with previously extracted roads (cf. Fig. 7a). Next, we approximate the outline of crossings by blowing up and optimizing closed snakes around the junctions (cf. Figs. 8a–c, 7b). We than establish connections between crossings and adjacent ends of previously

Table 1. Evaluation of results obtained by snake-based road extraction in images with a resolution of 0.5 m

	Fig. 11	Erquy (Fig. 12)	Marchetsreut (Fig. 13)
Correctness	0.97	0.95	0.99
Completeness	0.83	0.72	0.84
RMS [m]	0.46	0.46	0.37
Image size [pixel]	1800×1600	4500^2	2000^2
Time [min]	18	80	15

The "correctness" in Table 1 represents the ratio of the length of correctly extracted roads to the length of all extracted roads. The values prove that only a small number of false roads is extracted. "Completeness" corresponds to the ratio of the length of the correctly extracted roads and the length of the reference roads. Although the values for completeness are quite high for an automatic approach, they could have been even higher if forest and urban areas, where contrast on both road sides is weak, had been excluded from evaluation. The RMS values show that the precision is high and, more importantly, much better than demanded in most standards for topographic objects in GIS. For all three examples, they are better than one pixel and close to the value that arises from the fuzzy definition of the road sides. The time for the extraction (Sun Sparc 20) is reasonable. It is mostly proportional to the size of the image. However, it also depends on the scene and the number of roads, or more specifically, lines in it. The more gaps there are due to shadows and other disturbances in the *salient* roads, the more time it takes to verify all possible connections during the extraction of the *non-salient* roads.

6 Conclusions

In this paper, we have shown how it is possible to use ribbon snakes to take advantage of the little evidence that is available in shadowed parts of the road or when one side of the road is occluded. Thereby, we overcome some of the problems with other recent approaches for road extraction. Nevertheless, our approach is mostly intended for rural areas. Therefore, the best idea would be to combine it with the assets of other approaches. These are particularly the modeling of the context [2,28], the extraction of groups of cars to extract roads in urban areas [29], the use of GIS information [4, 5], and the exploitation of images from more than one sensor [34].

References

1. Barzohar M, Cohen M, Ziskind I, Cooper DB (1997) Fast Robust Tracking of Curvy Partially Occluded Roads in Clutter in Aerial Images. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Basel, Switzerland, 1997. Birkhäuser Verlag, Basel, p 277–286

2. Baumgartner A, Steger C, Mayer H, Eckstein W (1997) Multi-Resolution, Semantic Objects, and Context for Road Extraction. In: Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, Basel, Switzerland, 1997. Birkhäuser Verlag, Basel, p 140–156

3. Boichis N, Cocquerez JP, Airault S (1998) A Top Down Strategy for Simple Crossroads Extraction. Int Arch Photogramm Remote Sensing 32(2/1): 19–26

4. Bordes G, Giraudon G, Jamet O (1997) Road Modeling Based on a Cartographic Database for Aerial Image Interpretation. In: Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, Basel, Switzerland, 1997. Birkhäuser Verlag, Basel, p 123–139

5. Bückner J (1998) Model Based Road Extraction for the Registration and Interpretation of Remote Sensing Data. Int Arch Photogramm Remote Sensing 32(4/1): 85–90

6. Canny J (1986) A Computational Approach to Edge Detection. IEEE Trans Pattern Anal Mach Intell 8(6): 679–698

7. de Gunst M, Vosselman G (1997) A Semantic Road Model for Aerial Image Interpretation. In: Semantic Modeling for the Acquisition of Topographic Information from Images and Maps, Basel, Switzerland, 1997. Birkhäuser Verlag, Basel, p 107–122

8. Eberly D, Gardner R, Morse B, Pizer S, Scharlach C (1994) Ridges for image analysis. J Math Imaging Vision 4(4): 353–373

9. Fischler MA, Tenenbaum JM, Wolf HC (1981) Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique. Comput Graphics Image Process 15: 201–223

10. Fua P (1996) Model-Based Optimization: Accurate and Consistent Site Modeling. Int Arch Photogramm Remote Sensing (31) B3/III: 222–233

11. Fua P, Leclerc YG (1990) Model Driven Edge Detection. Mach Vision Appl 3: 45–56

12. Grün A, Li H (1997) Linear Feature Extraction with 3-D LSB-Snakes. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Basel, Switzerland, 1997. Birkhäuser Verlag, Basel, p 287–298

13. Kass M, Witkin A, Terzopoulos D (1987) Snakes: Active Contour Models. Int J Comput Vision 1(4): 321–331

14. Koenderink JJ (1984) The structure of images. Biol Cybern 50: 363–370

15. Koenderink JJ, van Doorn AJ (1994) Two-plus-one-dimensional differential geometry. Pattern Recognition Lett 15(5): 439–444

16. Koller TM, Gerig G, Szèkely G, Dettwiler D (1995) Multiscale detection of curvilinear structures in 2-D and 3-D image data. In: Fifth International Conference on Computer Vision, p 864–869, 1995

17. Laptev I (1997) Road Extraction Based on Snakes and Sophisticated Line Extraction. Master Thesis, Computational Vision and Active Perception Lab (CVAP), Royal Institute of Technology, Stockholm, Sweden

18. Lindeberg T (1994) Scale-space theory: A basic tool for analysing structures at different scales. J Appl. Stat, 21(2): 225–270. Supplement Adv Appl Stat: Statistics and Images: 2

19. Lindeberg T (1994) Scale-Space Theory in Computer Vision. Kluwer Academic Publishers, Boston, USA

20. Lindeberg T (1996) Edge detection and ridge detection with automatic scale selection. In: Computer Vision and Pattern Recognition, p 465–470

21. Lindeberg T (1998) Edge detection and ridge detection with automatic scale selection. Int J Comput Vision 30(2): 117–154

22. Mayer H, Baumgartner A, Steger C (1998) Road Extraction from Aerial Imagery, http://www.photo.verm.tu-muenchen.de/cvonline_roads/. In: CVonline: On-line Compendium of Computer Vision. http://www.dai.ed.ac.uk/CVonline/

23. Mayer H, Steger C (1998) Scale-Space Events and Their Link to Abstraction for Road Extraction. ISPRS J Photogramm Remote Sensing 53: 62–75

24. McKeown DM, Denlinger JL (1988) Cooperative Methods For Road Tracking In Aerial Imagery. In: Computer Vision and Pattern Recognition, p 662–672

25. Merlet N, Zerubia J (1996) New Prospects in Line Detection by Dynamic Programmming. IEEE Trans Pattern Anal Mach Intell 18(4):

426–431

26. Neuenschwander W, Fua P, Székely G, Kübler O (1995) From Zi-plock Snakes to Velcro^{TM} Surfaces. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images, Basel, Switzerland, 1995. Birkhäuser Verlag, Basel, p 105–114

27. Pizer SM, Burbeck CA, Coggins JM, Fritsch DS, Morse BS (1994) Object shape before boundary shape: Scale-space medial axis. J Math Imaging Vision 4: 303–313

28. Ruskoné R, Airault S (1997) Toward an Automatic Extraction of the Road Network by Local Interpretation of the Scene. In: Photogrammetric Week, Heidelberg, Germany, 1997. Wichmann Verlag, Heidelberg, p 147–157

29. Ruskoné R, Guigues L, Airault S, Jamet O (1996) Vehicle Detection on Aerial Images: A Structural Approach. In: 13th International Conference on Pattern Recognition, volume III, p 900–903

30. Sarkar S, Boyer KL (1993) Integration, Inference, and Management of Spatial Information Using Bayesian Networks: Perceptual Organization. IEEE Trans Pattern Anal Mach Intell 15(3): 256–274

31. Sing S, Sowmya A (1998) RAIL: Road Recognition from Aerial Images Using Inductive Learning. Int Arch Photogramm Remote Sensing 32 (3/1): 367–378

32. Steger C (1998) An Unbiased Extractor of Curvilinear Structures. IEEE Trans Pattern Anal Mach Intell 20: 113–125

33. Steger C, Mayer H, Radig B (1997) The Role of Grouping for Road Extraction. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), Basel, Switzerland, 1997. Birkhäuser Verlag, Basel, p 245–256

34. Tönjes R, Growe S (1998) Knowledge-Based Road Road Extraction from Multisensor Imagery. Int Arch Photogramm Remote Sensing 32 (3/1): 387–393

35. Trinder JC, Wang Y (1998) Knowledge-Based Road Interpretation in Aerial Images. Int Arch Photogramm Remote Sensing 32 (4/1): 635–640

36. Vosselman G, de Knecht J (1995) Road Tracing by Profile Matching and Kalman Filtering. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images, Basel, Switzerland, 1995. Birkhäuser Verlag, Basel, p 265–274

37. Wiedemann C, Heipke C, Mayer H, Jamet O (1998) Empirical Evaluation of Automatically Extracted Road Axes. In: CVPR Workshop on Empirical Evaluation Methods in Computer Vision, p 172–187

38. Witkin AP (1983) Scale-space filtering. In: Eighth International Joint Conference on Artificial Intelligence, August 1983, p 1019–1022

39. Zlotnick A, Carnine PD (1993) Finding Road Seeds in Aerial Images. Comput Vision Graphics Image Process: Image Understanding, 57: 243–260

Helmut Mayer is an Associate Professor at the Chair for Photogrammetry and Remote Sensing, Technische Universität München. He received his Dr.-Ing. degree in 1993 on the automatic interpretation of cadastral maps and the Dr.-Ing. habil. degree in 1997 with a survey about the automatic extraction of objects from aerial imagery. Dr. Mayer is chairing the working group III/3 on "Feature Extraction and Grouping" of the International Society for Photogrammetry and Remote Sensing (ISPRS). His interests include various aspects of computer vision and image analysis.

Tony Lindeberg received his M.Sc. degree in engineering physics and applied mathematics from KTH (Royal Institute of Technology), Stockholm, Sweden in 1987, and his Ph.D. degree in computing science in 1991 with the dissertation "Discrete Scale-Space Theory and the Scale-Space Primal Sketch". He is currently an Associate Professor at the Department of Numerical Analysis and Computing Science at KTH. His main research interests are in computer vision and relate to multi-scale representations, focus-of-attention and shape. He has contributed to the foundations of continuous and discrete scale-space theory, as well as to the application of these theories to computer vision problems. Specifically, he has developed principles for automatic scale selection, methodologies for extracting salient image structures, and theories for multi-scale shape estimation. He is author of the book "Scale-Space Theory in Computer Vision".

Ivan Laptev received his M.Sc. degree in computer science and technology from KTH (Royal Institute of Technology), Stockholm, Sweden in 1997. He is currently a Ph.D. student at the Department of Numerical Analysis and Computing Science at KTH. His research interests are in computer vision and include shape recognition, multi-scale image representation, temporal image processing and tracking.


Wolfgang Eckstein received his master's degree in computer science in 1988 and his doctorate in 1993, both from the Technische Universität München. Then he was Assistant Professor at the Forschungsgruppe Bildverstehen FG BV (computer vision research group) at the same institution. Now he is Managing Directory of the machine vision company MVTec. He also serves as chairman of Working Group III/4 "Image Understanding / Object Recognition" of the International Society for Photogrammetry and Remote Sensing (ISPRS). The main fields of his research are tools for image analysis, aerial image interpretation, morphology, robotics, image sequence analysis, and knowledge-based graphical user interfaces for computer vision.

Carsten Steger received his M.Sc. and Ph.D. degree in computer science from the Technische Universität München in 1993 and 1998, respectively. From 1998 to 1999 he worked as a post-doctoral researcher at the Forschungsgruppe Bildverstehen FG BV (computer vision research group) at the same institution. He is one of the founders of the machine vision company MVTec, where he currently works. He also serves as secretary of Working Group III/4 "Image Understanding/Object Recognition" of the International Society for Photogrammetry and Remote Sensing (ISPRS). His main research interests are low-level vision, machine vision, perceptual grouping, and semantic models for the interpretation of aerial images.
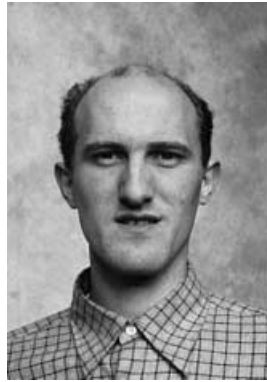
Albert Baumgartner is a scientific collaborator at the Chair for Photogrammetry and Remote Sensing, Technische Universität München (TUM). He studied geodesy at TUM and graduated in 1995. Since then the main focus of his work was on automatic extraction of roads from aerial imagery.