Model-Based Road Extraction from Images

Carsten Steger, Clemens Glock, Wolfgang Eckstein, Helmut Mayer, Bernd Radig

Lehrstuhl für Bildanalyse und wissensbasierte Systeme
Technische Universität München
Orleansstraße 34, 81667 München, Germany
Ph.: +49-89-48095-122, Fax: +49-89-48095-203
e-mail: {stegerc|eckstein|radig}@informatik.tu-muenchen.de

Lehrstuhl für Photogrammetrie und Fernerkundung
Technische Universität München
Arcisstraße 21, 80290 München, Germany
Ph.: +49-89-2105-2671, Fax: +49-89-280-9573
e-mail: {clemens|helmut}@photo.verm.tu-muenchen.de

Abstract

In this paper we present an approach to the automatic extraction of roads from aerial images. We argue that a model for road extraction is needed in every step of the image interpretation process. The model needs to include knowledge about different aspects of roads, like geometry, radiometry, topology, and context. The main part of this paper discusses the parts of that knowledge that we have implemented so far. It is shown that roads can be successfully detected at various resolution levels of the same image. Furthermore, we show that combining the results obtained in each level helps to eliminate false hypotheses typical for each level. The approach has been successfully applied to a variety of images.

1 Introduction

The extraction of roads from images has received considerable attention in the past. Several schemes have been proposed to solve this problem at resolutions that range from satellite images to low altitude aerial images. The strategies proposed fall into two broad categories. The work described in Gruen et al. (1994), Heipke et al. (1994), and McKeown et al. (1988) deals with the semi-automatic extraction of roads. The human operator has to select a certain number of points of the road which is then extracted. On the other hand, the work presented in Barzohar et al. (1993), Ruskoné et al. (1994), and Zerubia et al. (1993) is concerned with the automatic extraction of roads.

The aim of the work presented here is the automatic extraction of roads from aerial images for input into geographic information systems (GIS), in particular the German Authoritative Cartographic Topographic Information System ATKIS DLM25, which uses a scale of 1 : 25 000. Orthophotos with a ground resolution of 25 cm per pixel are used because they contain the necessary detail for a successful interpretation and are sufficient to obtain the geometric precision required.

Section 2 gives a short outline of the kinds of knowledge necessary for the extraction process. In section 3 the implemented parts of the model are presented in detail. Section 4 shows results. Finally, section 5 concludes the paper.

2 A Model for the Extraction of Roads

In order to be able to extract roads from aerial images a vision system must incorporate various kinds of knowledge. The proposed model groups knowledge into the categories geometry, radiometry, topology, and context. For example, the knowledge that roads are linear objects that have parallel roadsides, and that their width is more or less constant belongs to the geometric part of the model. The fact that roads have a surface that is relatively homogeneous in the direction of the road, and that they can have bright markings is radiometric knowledge. Furthermore, roads build up a hierarchical network which is used to connect towns on a coarse level and buildings on a fine level. This means, for example, that usually roads do not end without there being an object to connect to nearby. This aspect belongs to the knowledge about the topology of roads. It can be used to reject road hypotheses that cannot be connected with other roads or to start a process that looks for a reason why no connection can be made, e.g., shadows cast by trees. Finally, the context in which roads can appear has to be modeled as well in order to be able to explain failures to extract roads and correct them, and to discriminate roads from objects that look like roads, e.g., drive ways or roofs.

The extraction process has to receive attention as well. We argue to use knowledge as early as possible. This leads us to use a process that uses the parts of the road network which can be extracted quickly and with high reliability as the starting point for all further computations. Furthermore, an automatically generated digital elevation model (DEM) can be used to exclude objects that are higher than their surroundings, e.g., buildings or trees, from processing in later steps. Additionaly, color can be used to exclude regions that cannot be roads, e.g., lawns or parks. Utilisation of this kind of knowledge in very early stages helps to avoid generating false road hypotheses in later stages and speeds up processing.

Furthermore, knowledge about the areas of application of operators and knowledge about their strengths and weaknesses have to be incorporated in to the system. This can be used to select the appropriate extraction strategy in different contexts. For example, within cities a DEM can give most of the relevant hints for finding roads, if “ravines” between buildings are extracted. On the other hand, in rural or suburban areas the DEM serves mainly to eliminate false hypotheses, while the main clues are provided by radiomety and geometry.

3 Implemented Parts of the Model

3.1 Detection of Lines at Low Resolutions

In this section we will describe how roads can be detected in aerial images of reduced resolution. Because of the implicit smoothing performed in the scaling process, several kinds of problems that make the detection of roads at high resolutions more difficult are alleviated, e.g., cars on the road or shadows cast onto the road by adjacent trees.

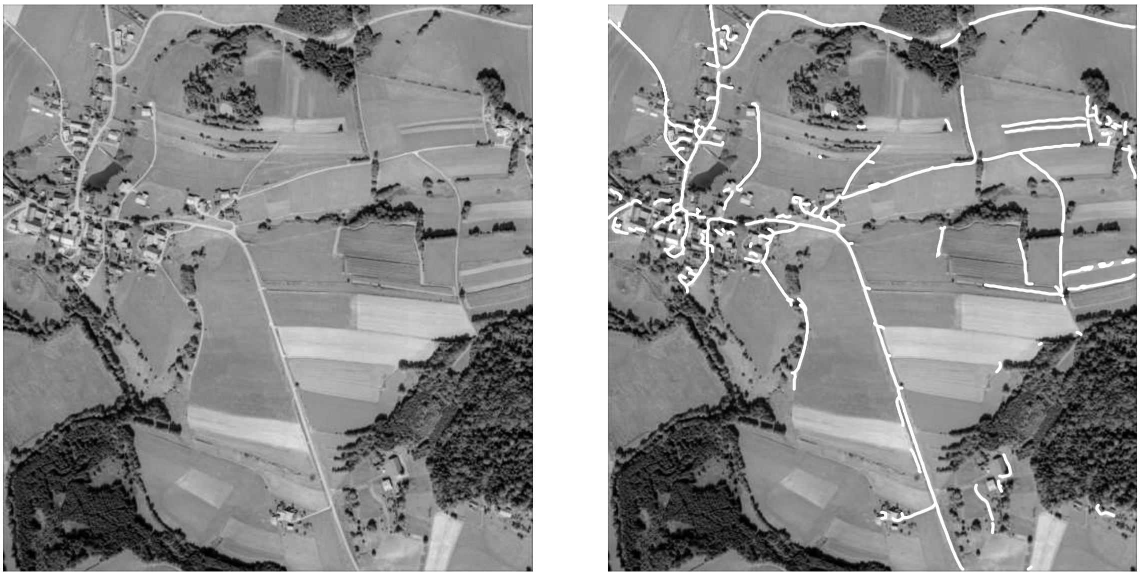
To start the extraction process the original orthophoto is reduced by a factor chosen such that roads in the reduced image are at most five to six pixels wide. At this level of detail roads can be modeled as lines that usually are brighter than their surroundings. To extract the lines from this image each pixel is compared to a locally computed Gaussian mean. Pixels that are brighter than the local mean by a certain threshold are selected for further processing (R_1). In a second step the image is thresholded such that only regions which have an intensity that lies within a certain range are selected (R_2). For the image in figure 1 only relatively bright regions were selected. Then the intersection of these two regions is computed, resulting in $R_3 = R_1 \cap R_2$. This result corresponds to the regions in the image that are bright and brighter than their surroundings. After this step the skeleton S of R_3 is computed since we are only interested in the center lines of the roads at this resolution.

Unfortunately, the process so far selects lines as well as certain kinds of edges in the image. To select the lines, contours are computed from S. A contour is defined as a sequence of 8-connected points that starts and ends either in a junction point or a point that has exactly one neighbour in S. The contour-points are then examined whether they are local maxima on a line perpendicular to the direction of the contour. Only contours that have more than a certain percentage of points that are local maxima are selected. The output of this step are the road hypotheses at this resolution level. Figure 1b) shows the result of this process.

3.2 Edge-Extraction and Polygonal Approximation

The next three sections will give details about the road extraction process at the highest resolution level. At this level we assume that roads are relatively homogeneous regions in the image that have a significantly different brightness than their surrounding areas. Hence it follows that roadsides can be detected by an edge extraction algorithm. For this task we use a modified version of a Deriche edge detector that has been described in Lanser et al. (1992). The advantages of this operator are very good detection quality, accurate edge localisation, few multiple responses, and isotropic response. The resulting edges are then thinned by a non-maximum-suppression algorithm, yielding one pixel wide edges. From these edges contours are computed.

To reduce the amount of data which has to be handled, and to faciliate the perceptual grouping of parallel lines, a polygonal approximation of each contour is computed by the algorithm given in Ramer (1972). This algorithm splits contours into polygon segments which have a limited distance to the approximated contour. The advantage of this algorithm is that it yields rather long line segments with an acceptably small approximation error.



a) b)

Fig. 1: a) Reduced version of the original orthophoto (reduced by a factor of 8) b) Result of the line detection algorithm

Figure 2 shows the result of the edge detection and polygonal approximation step with the maximum deviation set to 2 pixels.

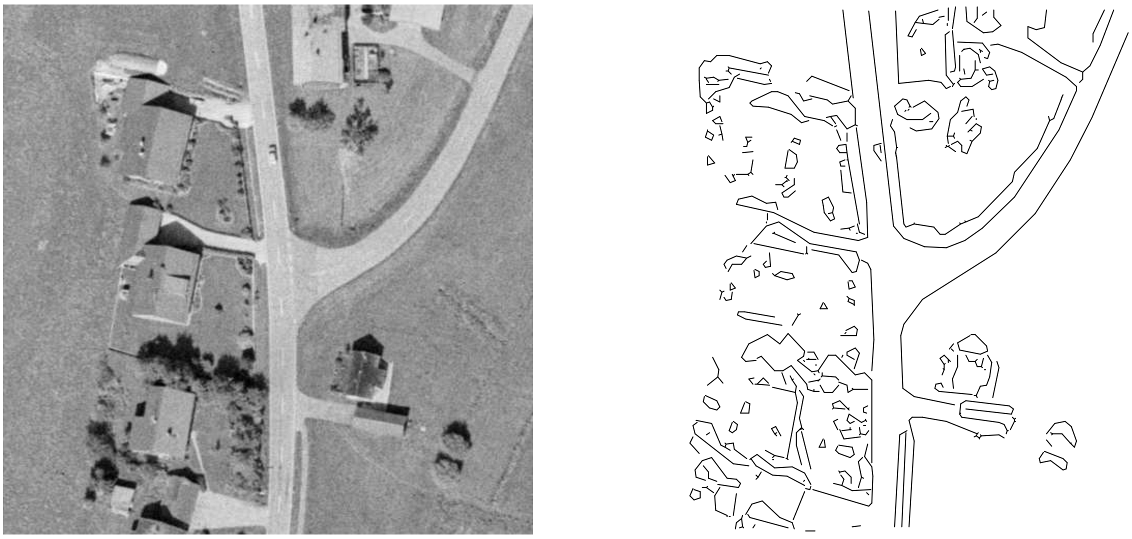
3.3 Perceptual Grouping of Parallel Edges

One feature which characterises roads is the parallelism of opposite roadsides. Therefore the next step in the road-extraction process is to construct relations of parallel polygons.

To be included into this relation, line segments have to fulfill several criteria. First of all, the line segments should be parallel. Because of the discrete nature of the edge extraction process and because the edges are approximated by polygons, opposite roadsides are never perfectly parallel. Furthermore, the direction of a long line segment is determined more accurately than that of a short segment. Therefore the criterion to determine whether two line segments are parallel has to allow for a certain angular difference that depends on the length of the lines involved. The following formula is used as a threshold for the angle α that two lines are allowed to enclose:

$$\alpha \leq \tau + \frac{w\tau}{\max(l_1, l_2)} \tag{1}$$

Here, l_1 and l_2 are the lengths of the line segments involved, and τ is a user-selectable threshold. This formula allows short lines to enclose a greater angle than long lines. The weight w can be used to control how far the threshold grows for short lines. In the current implementation it is set to 1.



a) b)

Fig. 2: a) Input image for the edge detection process b) Detected edges approximated by polygons

The second criterion is that parallel line segments have to overlap. This is determined by projecting the end points of both lines onto a line that has a direction determined by the bisection of the angle that the lines enclose. If the projected lines overlap, then the original lines overlap as well. The last criterion is that the lines involved have to be closer than a certain threshold. This is motivated by the fact that roads cannot be wider than a certain distance. Figure 3 shows the results of this step with $\tau = 10^\circ$ and a maximum distance of 7.5 m (30 pixels).

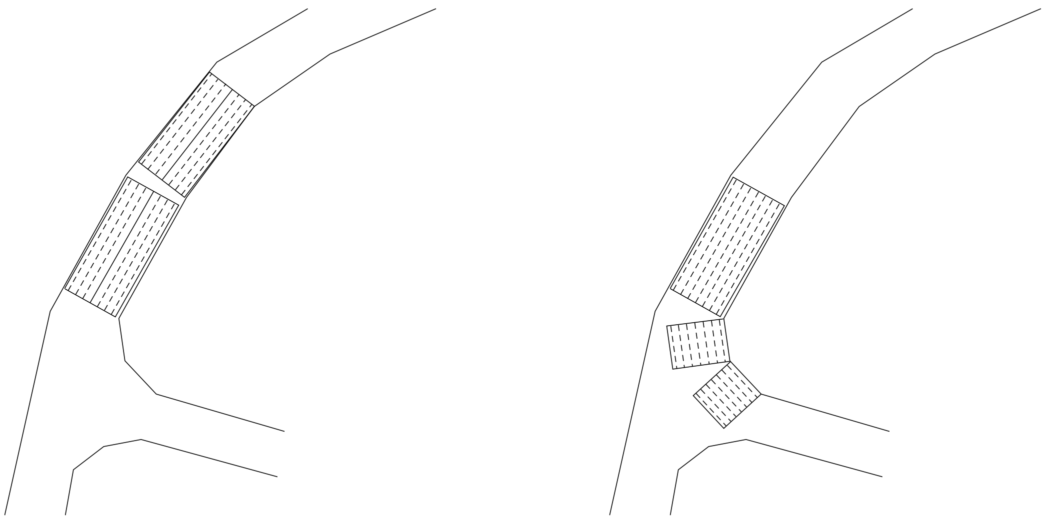
3.4 Selection of Homogeneous Areas between Parallels

The next step in our approach is concerned with the radiometric part of the road model. In this step we examine whether the area between the parallel line segments found in section 3.3 is homogeneous in the direction of the center line between them. This corresponds to the assumption that the surface intensity of a road is relatively constant in the direction of the road, whereas it can vary considerably across the road due to road markings or tire tracks.

To determine whether the region between two parallel line segments is homogeneous, slices parallel to the center line of the two lines are generated. These slices are 1 pixel apart, and the intensities within each slice are computed by bilinear interpolation. Figure 4a) shows how these slices are placed between two parallels. The solid line indicates the center line, while the dashed lines show the parallel slices. Note that only the minimal area between the parallels is used for computation.

The procedure then computes the mean intensity within each slice. If the mean falls outside a certain user-settable range, the region is immediately rejected. This is due to the

Fig. 3: Parallel line segments detected



a) b)

Fig. 4: a) Slices computed between parallel line segments b) Slices computed while following the original edges

model assumption that a road is normally a relatively bright object, but it also allows one to select dark roads if present. After this the intensity variance in each slice is determined. This is the main criterion for homogeneity of the region in question. If the variance in each slice is smaller than a threshold, the parallel lines that enclose the region are accepted as hypothetical roadsides. The results of this step are shown in figure 5a).

As is evident from figure 5a), the algorithm is quite successful in finding the parts of the road where parallel roadsides exist. It fails in regions of the road where no parallelism can be detected, e.g., at intersections. Therefore all edges which neighbour the last parallel edge selected are examined to see if they still border homogeneous regions. To this end, rectangles are constructed that have the width of the last rectangle processed in the previous step. These rectangles are then sliced as above and examined for homogeneity. Figure 4b) shows the last segment processed and the two following segments that are examined. The result of this extension process can be seen in figure 5b).

a) b)

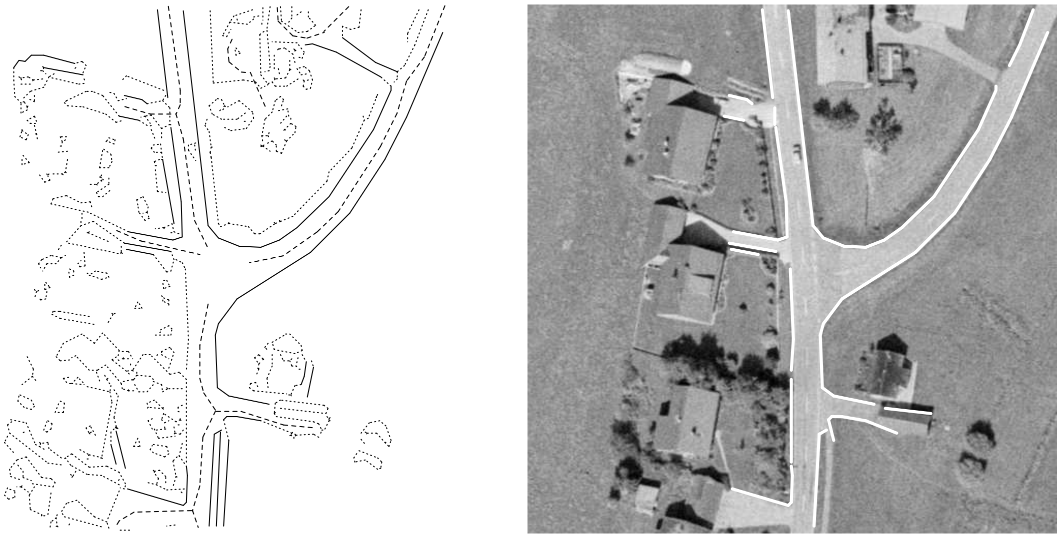
Fig. 5: a) Selected parallels that enclose a homogeneous region b) Result of the parallel line extension process

3.5 Combining the Different Resolution Levels

As can be seen from the examples given so far, the extraction process in the two different resolution levels have their advantages and deficiencies. Therefore in this section we show how to combine the results of both levels to eliminate incorrectly detected road-segments and to extend results to regions where the extraction process so far has failed for some reason. The basic strategy in this step is to take the results of both levels which support each other. To start the process, parallel lines that enclose a homogeneous area of the high resolution level are selected if a center line of the reduced resolution level is found between the two parallels. Starting from these strong hypotheses, gradually all the roadsides are extracted.

This is done using a number of rules. One rule is, e.g.: Take roadside R_1 and its parallel roadside R_2. If there exists an edge R_3 that is colinear to R_1, and R_2 overlaps both R_1 and R_3, and there is a center line C that lies between R_2 and R_3, then accept R_3 as a new roadside. As can be seen from this example, the rules are rather conservative, selecting only edges as roadsides that have support in all resolution levels. Figure 6a) depicts the input to the combination process. The results obtained in section 3.4 are shown as solid lines, the original edges as dotted lines, and the center lines obtained in section 3.1 as dashed lines.

Figure 6b) shows the final result of the road extraction process. It can be seen that the algorithm was able to bridge the gap on the left hand side of the road at the intersection that resulted from lacking parallelism. Furthermore, some erroneously detected road hypotheses next to the houses on the left hand side of the image and a drive way on the upper right hand side were eliminated. The only errors that need to be eliminated are the drive ways on the left hand side of the image. The ATKIS model states that these roads must not be included in the GIS dataset.



a) b)

Fig. 6: a) Input to the combination step b) Final result of the road extraction process

4 Results

In this section we will present two more examples of the results obtained with the approach we have described. The first example is a different part of the image shown in figure 1a). Figure 7 shows that most parts of the road network are extracted correctly. However, due to shadows cast onto the road by trees, some parts of the road in the upper part of the image could not be detected. Also, adjacent to that road, an erroneous road segment has been detected because the shadow of a tree has pushed the center line outside the road, thereby making it support that segment. Finally, not all parts of the track that leads into the field on the right hand side of the image have been eliminated because a short center

a) b)

Fig. 7: a) Input image to the algorithm b) Final result of the road extraction process

line was detected that gave support to the initial part of the track.

Figure 8 shows the result obtained with the image 5987 of the Ascona testset. Since this image has a ground resolution of 7.5 cm, it was reduced by a factor of 3 to fit our model. Again, most of the road network is detected correctly. However, several buildings are extracted as well. These errors could be eliminated very easily through the use of a DEM by excluding areas that are higher than their surroundings from being input to the algorithm.

5 Conclusions

The results given in sections 3.5 and 4 indicate that the strategy we have chosen is quite general and leads to good results. However, more parts of the knowledge mentioned in section 2 have to be implemented for a successful automatic extraction of roads. Especially the topology and the context are crucial, as can be seen from the examples.

Therefore, we will focus on implementing knowledge about the topology of roads and the context in which roads appear in the next phase of the project. To faciliate this we plan to use color images and automatically computed digital elevation models as additional input data.

Acknowledgements

We would like to acknowledge the vital contributions of Heinrich Ebner to this project. This work was funded by the Deutsche Forschungsgemeinschaft under grant no. Eb 74-8/1.

a) b)

Fig. 8: a) Image 5897 of the Ascona testset b) Final result of the road extraction process

References

Barzohar M., D.B. Cooper (1993) *Automatic Finding Of Main Roads In Aerial Images By Using Geometric - Stochastic Models and Estimation*, CVPR 93, pp. 459–464.

Gruen A., H. Li (1994) *Semi-Automatic Road Extraction by Dynamic Programming*, IntArchPhRS, Vol. 30, Part 3/1, pp. 324–332.

Heipke C., A. Englisch, T. Speer, S. Stier, R. Kutka (1994) *Semi-Automatic Extraction of Roads from Aerial Images*, IntArchPhRS, Vol. 30, Part 3/1, pp. 353–360.

Lanser S., W. Eckstein (1992) *A Modification of Deriche's Approach to Edge Detection*, ICPR 92, Vol. III, pp. 633–637.

McKeown D.M., J.L. Denlinger (1988) *Cooperative Methods For Road Tracking In Aerial Imagery*, CVPR 88, pp. 662–672.

Ramer U. (1972) *An Iterative Procedure for the Polygonal Approximation of Plane Curves*, Computer Graphics and Image Processing, Vol. 1, pp. 244–256.

Ruskoné R., S. Airault, O. Jamet (1994) *Road Network Interpretation: A Topological Hypothesis Driven System*, IntArchPhRS, Vol. 30, Part 3/2, pp. 711–717.

Zerubia J., N. Merlet (1993) *Classical Mechanics and Roads Detection in SPOT Images*, INRIA Rapport de Recherche No. 1889, INRIA Sophia Antipolis, France.