Automatic Road Extraction Based on Multi-Scale, Grouping, and Context*

Albert Baumgartner, Carsten Steger, Helmut Mayer, Wolfgang Eckstein, and Heinrich Ebner

Abstract

An approach for the automatic extraction of roads from digital aerial imagery is proposed. It makes use of several versions of the same aerial image with different resolutions. Roads are modeled as a network of intersections and links between these intersections, and are found by a grouping process. The context of roads is hierarchically structured into a global and a local level. The automatic segmentation of the aerial image into different global contexts, i.e., rural, forest, and urban area, is used to focus the extraction to the most promising regions. For the actual extraction of the roads, edges are extracted in the original high resolution image (0.2 to 0.5 m) and lines are extracted in an image of reduced resolution. Using both resolution levels and explicit knowledge about roads, hypotheses for road segments are generated. They are grouped iteratively into larger segments. In addition to the grouping algorithms, knowledge about the local context, e.g., shadows cast by a tree onto a road segment, is used to bridge gaps. To construct the road network, finally intersections are extracted. Examples and results of an evaluation based on manually plotted reference data are given, indicating the potential of the approach.

Introduction

In digital photogrammetry, operational automatic solutions exist for geometric tasks such as the measurement of fiducials and matching of homologous points. The latter is used for the reconstruction of the relative orientation, for the generation of digital surface models, or for automatic aerotriangulation. For data acquisition and update of geographic information systems (GIS), the determination of the meaning of individual topographic objects, e.g., buildings and roads, is necessary. This semantic task still has to be done manually. Because this is time consuming and expensive, automatic solutions are highly welcome.

Research on the automatic extraction of topographic objects from aerial and space imagery goes back to the seventies. Nowadays the goal is the update of GIS data. Using existing, albeit outdated, data can facilitate object extraction. However, the approach proposed in this paper is restricted to the extraction of objects (here: roads) without GIS data. There are several reasons for this. First, automatic extraction without *a priori* information shows the potential and deficits of an extraction scheme much better than a GIS-based extraction, because it only relies on the given model and strategy, and therefore can deepen the understanding of the problem. Second, the extraction of new objects is possible only in this manner and is needed for GIS update in any case. Third, to make the system reliable, it is wise to base the decision about an object on new imagery and not on old GIS data. Nevertheless, work on GIS-based extraction of roads (de Gunst and Vosselman, 1997; Bordes *et al.*, 1997; Plietker, 1997) is useful, and has been carried out also within our approach (Wiedemann and Mayer, 1996).

The most common techniques for road extraction in images with low resolution are the detection and following of lines. In high resolution, matching of profiles and detection of roadsides, i.e., (anti-)parallel edges, are used. The different approaches apply specialized algorithms and additional knowledge, e.g., geometric constraints. The main criterion to classify the extraction schemes is human interaction. In semi-automatic approaches, an operator provides, for example, starting points and starting directions for road following (McKeown and Denlinger, 1988; Vosselman and de Knecht, 1995). In Merlet and Zerubia (1996), points along a road are measured and the algorithm finds the road, i.e., a line which connects these points. If more than one image is used, this can also be done in 3D (Grün and Li 1997). The advantage of the approaches with multiple points is that the path of the road is more constrained, which results in a more reliable handling of critical areas. A similar approach based on so-called "ziplock" snakes is presented in Neuenschwander *et al.* (1995).

By automatic detection of the seed points, semi-automatic schemes can be extended to (fully) automatic ones. An automatic approach is described in Barzohar *et al.* (1997). The selection of starting points is based on gray-value histograms. Further assumptions about geometry and radiometry are modeled by a Markov stochastic process. Road extraction is performed by dynamic programming. In Ruskoné *et al.* (1994), the centers of elongated regions are detected using a watershed transform of the gradient image. Starting from these points, the homogeneity of the road surface in the images is used to extract road segments. Using geometrical constraints, connection hypotheses between road segments are checked, and a road network is constructed. Similarly to the approach proposed in this article, Trinder and Wang (1998) extract roads using different resolutions and grouping.

When relations between roads and other objects, e.g., vehicles, buildings, or trees, are neglected, a reliable extraction is

*This is a modified version of an article published in German in *Photogrammetrie-Fernerkundung-Geoinformation (PFG)*, No. 1/99, pp. 5–17, 1999.

A. Baumgartner, H. Mayer, and H. Ebner are with the Lehrstuhl für Photogrammetrie und Fernerkundung, Technische Universität München, D-80290 Munich, Germany ({albert}{helmut}{ebn}@photo.verm.tu-muenchen.de).

C. Steger and W. Eckstein are with the Forschungsgruppe Bildverstehen (FG BV), Informatik IX, Technische Universität München, D-80290 Munich, Germany ({stegerc}{eckstein}@informatik.tu-muenchen.de).

Photogrammetric Engineering & Remote Sensing
Vol. 65, No. 7, July 1999, pp. 777–785.

0099-1112/99/6507–777$3.00/0

not possible in many cases. These background objects often have a strong influence on the appearance of roads in aerial imagery. For example, high objects may cause occlusions or cast shadows on the road. Therefore, Strat (1995) strongly demands the use of context information to guide image understanding algorithms in complex scenes. In Bordes *et al.* (1997), the influence of neighboring objects on the reliability of road extraction is investigated and employed to improve the verification of roads from a GIS database. In urban areas tracking of profiles or extraction of roadsides is often difficult. For these areas, Ruskoné *et al.* (1996) use the extraction and grouping of vehicles for road extraction.

One major finding in recent years is that the characteristic properties of roads described by a model are not the same at different image resolution levels (Mayer and Steger, 1998) and in different contexts (Baumgartner *et al.*, 1997). To describe the appearance of roads, three so-called "global contexts" are distinguished in this work: rural, forest, and urban area. For every global context, specific relations between road objects and background objects are described in so-called "local contexts." The proposed approach starts with the generation of hypotheses for roadsides using two resolution levels. Edges are extracted in the original high resolution image and lines in an image with reduced resolution. Using both resolution levels and explicit knowledge about geometric and radiometric properties of roads, road segments are constructed from the hypotheses for roadsides. From these road segments, road links, i.e., roads connecting two intersections, are built. Road segments and links are semantic objects with attributes and methods. Because they explicitly represent a specific area in the image, a search for additional evidence, e.g., road markings, is facilitated. Attributes and methods of the road objects vary according to the context.

The road model and the different contexts are described in the next section. The basic strategy for road extraction as well as the basic steps and characteristic elements of this approach are then explained. Finally, after an evaluation of the results, a short outlook concludes the paper.

Model

In order to extract roads from an aerial image, it is necessary to have a clear idea about the concept "road." For the proposed approach, the model comprises explicit knowledge about geometry (road width, parallelism of roadsides, etc.), radiometry (reflectance properties), topology (network structure), and context (relations with other objects, e.g., buildings or trees). The model described below consists of two parts. The first part describes characteristic properties of roads in the real world and in aerial imagery, and presents a road model derived from these properties. The second part defines different local contexts and assigns those to the global contexts. In this way, the complex model for the object road is split into sub-models that are adapted to specific contexts. Because the sub-models emphasize certain characteristics of the objects, they can also be regarded as specialized models.

Roads

Roads in the Real World.

A description of roads in the real world can be derived from their function for human beings. Roads access space and are therefore organized as a network connecting all areas inhabited and exploited by human beings. The denser an area is inhabited and the more intensively it is used, the denser the road network is. According to their importance, network components are classified into a hierarchy of different categories with different attributes. Roads of national importance are much wider than roads accessing rural areas. Field paths and less important roads follow the natural terrain surface more closely than highways, which serve as fast connections between conurbations. According to the different categories, roads differ with respect to minimum curvature radius and maximum allowed slope. As a consequence of greater curvature radii, more gentle slopes, and multi-layered interchanges, there is a strong tendency for roads of greater importance to run along embankments, over bridges, or in tunnels. Some important attributes for parts of the road net are the type and state of the road surface material; existence of road markings, sidewalks, and cycle-tracks; or legal instructions, such as traffic regulations.

Roads in Aerial Imagery

The appearance of roads in digital aerial imagery strongly depends on the sensor's spectral sensitivity and its resolution in object space. The proposed approach is restricted to gray-scale images, and only scale dependencies are considered. In images with low resolution, i.e., more than 2 m per pixel, roads mainly appear as lines that form a more or less dense network. Contrary to this, in images with a higher resolution, i.e., less than 0.5 m, roads are projected as elongated homogeneous regions with almost constant width. Here the attainable geometric accuracy is better, but background objects such as cars, trees, or buildings disturb the road extraction more severely.

In a smoothed image-which corresponds to a reduced resolution image-lines representing road centerlines can be extracted in a stable manner even in the presence of these background objects. The smoothing eliminates substructure of the road, e.g., vehicles or markings. This can be interpreted as abstraction, i.e., the object road is simplified and its fundamental characteristics are emphasized, as shown in Mayer and Steger (1998).

Road Model

From the last paragraph, it follows that the fusion of low and high resolution imagery can contribute to improving the reliability of road extraction. Additionally, details such as road markings, which can be recognized at a resolution of less than 0.2 m, can be used as further evidence to validate the detected road hypotheses. On the one hand, using multiple resolution levels improves the robustness of the road extraction. On the other hand, it results in different features at each resolution level, and the necessity to combine all features of all resolution levels into one road model. The road model condensed from the findings above is illustrated in Figure 1.

This road model describes objects by means of "concepts," and is split into three levels defining different points of view. The *real world* level comprises the objects to be extracted and their relations. On this level, the road network consists of intersections and road links that connect intersections. Road links are constructed from road segments. In fine scale, road segments consist of pavement and markings. The concepts of the real world are connected to the concepts of the *geometry and material* level by means of *concrete* relations (Tönjes, 1997), which connect concepts representing the same object on different levels. The geometry and material level is an intermediate level which represents the 3D shape of an object as well as its material (Clément *et al.*, 1993). The idea behind this level is that it describes objects independently from sensor characteristics and viewpoint, which is in contrast to the *image* level. Road segments are linked to the "mostly straight bright lines" of the image level in coarse scale. In contrast to this, the pavement as a part of a road segment in fine scale is linked to the "elongated bright region" of the image level by using the "elongated, flat concrete or asphalt region."

Whereas the fine scale gives detailed information, the coarse scale adds global information. Because of the abstraction in coarse scale, additional correct hypotheses for roads can be found and sometimes also false ones can be eliminated,

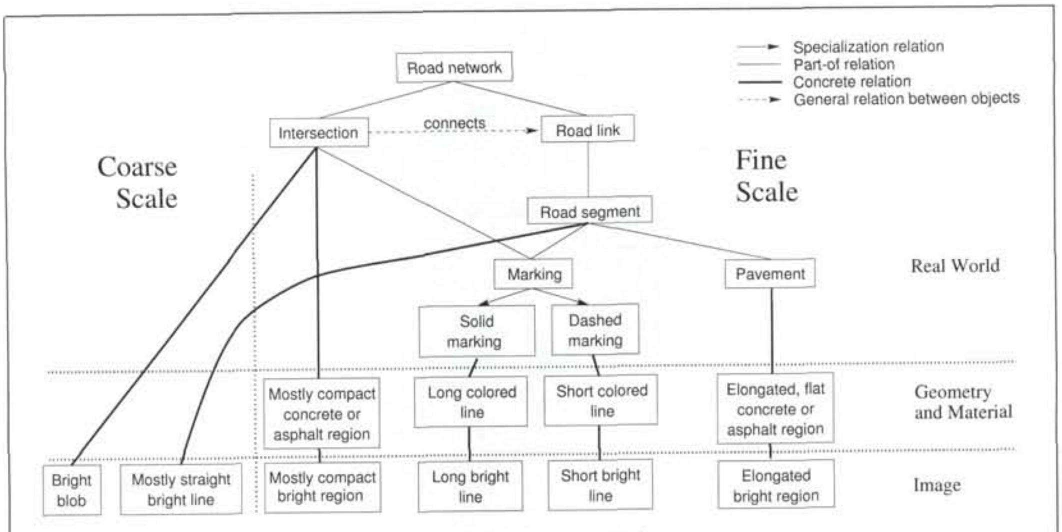


Figure 1. Road model.

while details, such as exact width and position, or markings, from fine scale are integrated. In this way the extraction benefits from both scales.

Context

The road model presented above comprises knowledge about radiometric, geometric, and topological characteristics of roads. This model is extended by knowledge about context. Background objects, such as buildings, trees, or vehicles, can support road extraction (e.g., usually there is a road to every building), but also interfere (e.g., a building occludes a part of a road, roofs might look similar to roads). This interaction between road objects and background objects is modeled *locally* and *globally*.

Local Context

With local context, typical relations between a small number of road and background objects are modeled. Situations in which background objects make road extraction locally difficult are in an open rural area, for example, paths to agricultural fields or individual cars. Driveways to buildings cause problems in urban areas. Buildings are mostly parallel to roads. In urban areas sidewalks and cycle tracks run parallel to roads, potentially hindering or supporting road extraction. For the local context, these situations are described in sketches. The local context *occlusion_shadow* (Figure 2) illustrates a situation where a high object occludes a part of a road or casts a shadow on a road, thereby causing the detection of two disconnected road segments. Other local contexts are, e.g., *rural_driveway, building_driveway_road*, or *sidewalk/cycle-track_parallel_to_road*. These basic local contexts can be aggregated into more complex local contexts, in which, for example, *occlusion_shadow* and *building_driveway_road segment* interact.

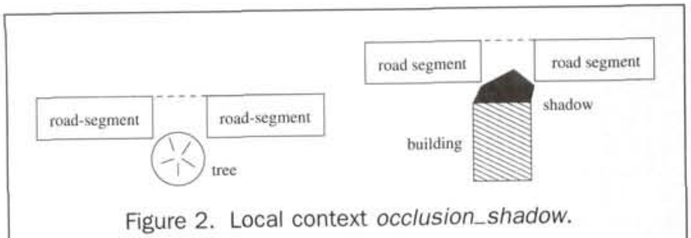


Figure 2. Local context *occlusion_shadow*.

Global Context

It is not necessary to always take into account every local context. Relations to background objects and their relevance for road extraction depend also on the region where they occur. For instance, roads in urban or suburban areas have a quite different appearance from roads in forest areas or in open rural areas. The differences in appearance are partly consequences of different relations between roads and buildings. In downtown areas, buildings typically are closer and more parallel to roads. Therefore, this paper proposes to use different local contexts for different areas, i.e., different global contexts. Here, *urban, forest*, and *rural* contexts are distinguished. The global context is not only relevant for the importance of the local contexts, but also for the extraction of objects. Experience shows that approaches that are suitable for road extraction in rural areas usually cannot be applied in other global contexts without modifications. In forest or urban areas, other parameter settings might be necessary or, more likely, even a completely different approach is required. From this, it is clear that the global context enables a more efficient use of the knowledge about roads. In Figure 3, some frequently occurring local contexts are assigned to the global contexts.

Tightly linked to the use of knowledge about context are the questions of how to get information about background objects and what their influence is on roads in the image. Because road extraction is the main goal, the information about background objects and different global contexts is not required with a high level of detail and accuracy. This information can be provided by an existing GIS or derived from the image itself, e.g., for the global context, by analyzing the texture in the image. The latter strategy has the advantage that it avoids problems caused by inconsistencies between image and outdated GIS data. A digital surface model (DSM), which can, for example, be automatically derived from two or more images, can be used to explain many situations in which road extraction is difficult.

Strategy

In addition to the road and context models, the strategy, i.e., the knowledge about how and when to use which part of the model, is very important for the performance of the approach. First, in this section a general description of the strategy for road extraction is given. Then, the steps of the extraction process are explained in more detail.

The basic idea of the strategy is to focus the extraction process on those parts of the road network that can be detected most

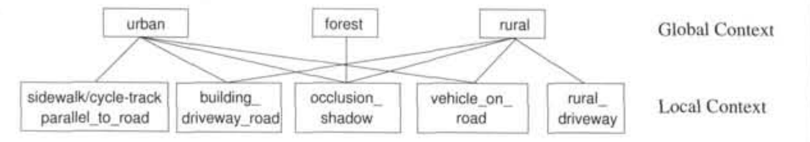


Figure 3. Global and local context.

easily and reliably, and that are ideally also useful to guide further extraction. How difficult the extraction of a certain feature is depends strongly on the context in which it is to be extracted. In urban and forest areas, knowledge about geometry and radiometry alone is often insufficient due to occlusions and shadows. On the other hand, with a simple model, relying only on attributes of the road itself, good results can be expected for rural areas.

As a consequence of these considerations, the road extraction starts in rural areas. The intermediate results after this step can then be used to guide the extraction in urban and forest areas. A segmentation of the image into the different global contexts is achieved automatically by texture analysis. In accordance with the road model, not only the original image at a resolution of about 0.2 to 0.5 m, but also a version of the image with reduced resolution, in which roads are only a few pixels wide, is used. In the reduced-resolution image, lines are extracted using the algorithm of Steger (1998) and are used to select edges from the original resolution image that are candidates for roadsides. The distance between pairs of edges must be within a certain range, depending on the road category; they have to be almost parallel; and the area enclosed by them should be quite homogeneous in the direction of the road. In addition, for each pair of candidates for roadsides, a corresponding line has to exist in the reduced resolution image. The selection of edges as roadside candidates is described in detail in Steger *et al.* (1995). From these roadsides, road segments are generated. Road segments are represented by the points of their medial axes, and are attributed by the road width. Initially, the road segments are quite short. The grouping into longer segments, i.e., the closing of gaps and the elimination of false hypotheses, is done according to the "hypothesize and test" paradigm. Hypotheses about which gaps are to be bridged are generated starting with geometric criteria (absolute and relative distance, collinearity, width ratio) and radiometric criteria (mean gray value, standard deviation). Then, the hypothetical road segments are verified in the image. The verification consists of up to three levels. In the first level, radiometric properties of the new segment are compared to the segments to be linked. The geometry of the new segment is defined by the direction at the endpoints of the segments to be linked. If the radiometric attributes do not differ too much, the connection hypothesis is accepted. If not, the verification comes to the second level. Here, a so-called "ribbon" snake is applied to the gradient image to find an optimum path for the link. In case this verification also fails, a third level is used, in which an explanation is searched for, why no evidence for a road can be found in the hypothetical road area. As possible explanations for such situations, local contexts such as *occlusion_shadow* are taken into account. This means that the local context is used as the last and apparently weakest evidence to explain and close gaps.

Simultaneously with the closing of gaps between road segments, hypotheses that are false with a high probability, i.e., short segments that cannot be connected to other segments, are eliminated. The next task is to find the intersections, i.e., the nodes of the network, in order to construct connections between the road hypotheses. Ideally, after this step all road hypotheses are connected, and there is a path between every pair of points on the extracted road network. Usually, such a result cannot be expected. First, due to the limited size of the images, some of the nodes will be outside of the image. Second, the results are not error-free. Especially in urban and forest areas only fragments of the network can be expected to be extracted. Because the extraction is reliable only in rural areas, the network characteristics of roads can neither be optimally exploited for the extraction nor for an internal evaluation. However, within a limited scope, it is possible to use topological relations to rate the importance of the roads in the network and to eliminate some of the remaining false hypotheses. By integrating knowledge about global and local context, the semantics of individual roads can be determined in more detail. For example, driveways to buildings can be distinguished from paths to fields. Roads with a certain minimum width can be expected to have road markings. This means that it is possible to focus the search for markings to the extracted roads. The existence of road markings is very good evidence for a correct hypothesis. The absence of markings can be used as a hint that the road hypotheses might be wrong and that further checks are necessary. Besides their benefit for the elimination of wrong hypotheses, markings also provide detailed information about a road, e.g., number of lanes or existence of turning lanes. Note, however, that a reliable extraction of markings, which is necessary for such specific analyses, can only be achieved at resolutions better than 0.3 m.

To make this discussion more explicit, the following results illustrate the individual steps:

Information about the global context is provided by a texture-based segmentation. Figure 4 shows the segmentation of *rural* areas in an image with a reduced resolution of about 4 m. The pixel size on the ground for this example is about 0.45 m in the high resolution image. The segmentation makes use of the texture filters proposed by Laws (1980) and incorporates morphological operations to deliver regions with smooth boundaries.

From the fusion of line extraction in a reduced resolution image with about 2 m and edge extraction in a high resolution image (Figure 5), hypotheses for roadsides, and from them hypotheses for road segments, are generated by grouping (Figure 6). The employed grouping processes have a worst case complexity of $O(n^2)$, where n is the number of objects to be grouped. Therefore, because roads can extend over the whole image, problems arise with the number of lines, edges, and road segments. Fortunately, the computational effort can be reduced tremendously by generating the hypotheses for road sides and constructing road segments locally, i.e., by working on small, overlapping image patches. In a second step, the hypotheses for road segments are collected from all patches, conflicting road segment hypotheses caused by overlapping patches are examined, and only the best hypotheses are kept.

Most of the road segments derived from the fusion of line and edge extraction are not directly connected and, what is more, there are also many false hypotheses. Therefore, the main task in the next steps is the linking of correct hypotheses and the elimination of false hypotheses. According to the above-mentioned grouping cues, hypotheses for connections are generated and verified. This is done iteratively. For every new

(a) (b)

Figure 4. (a) Aerial image. (b) *Rural areas.*

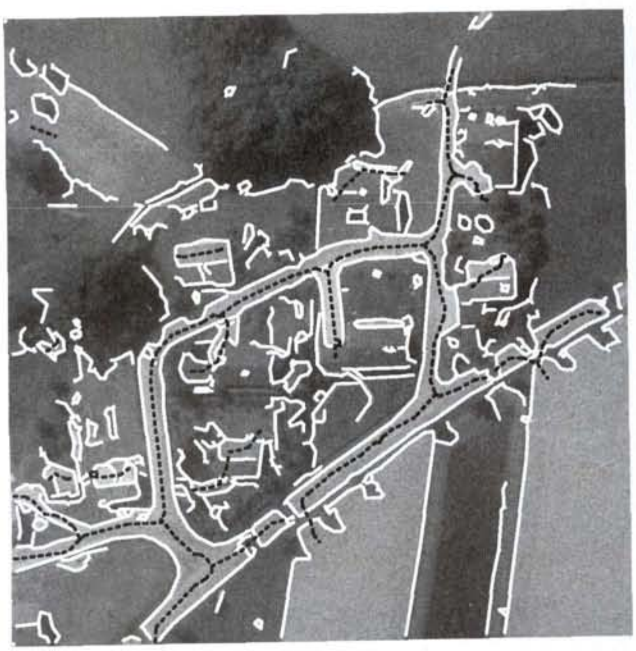


Figure 5. Input for the fusion: Lines (black, dashed), edges (white).

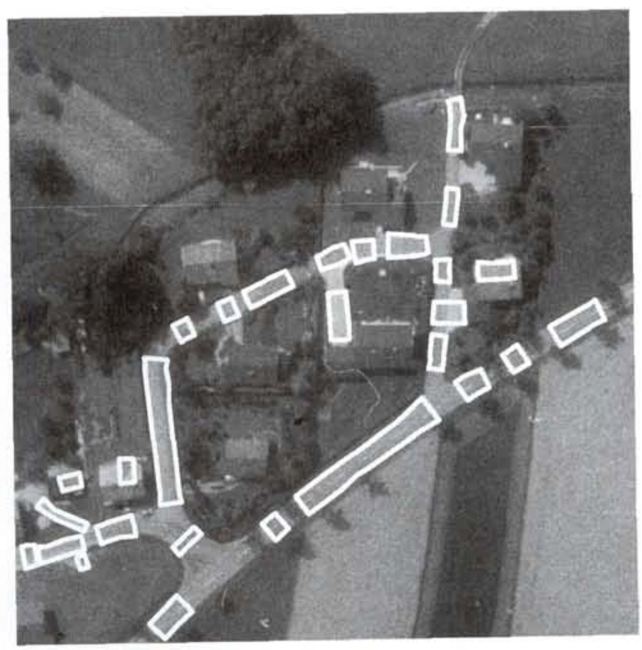


Figure 6. Hypotheses for road segments.

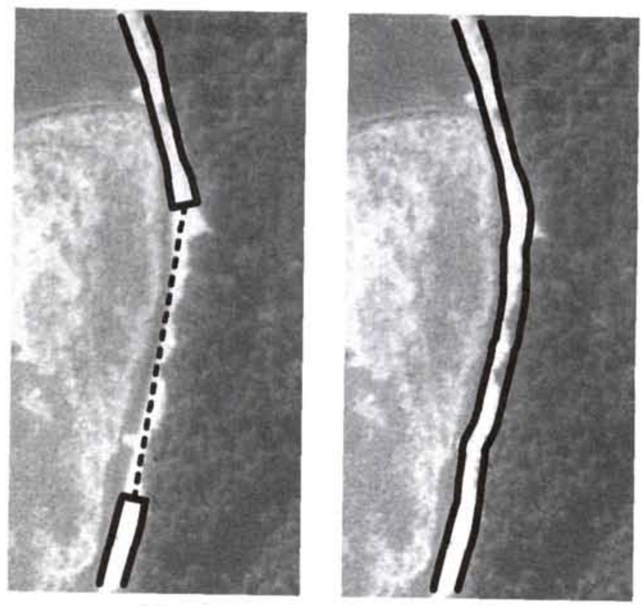
iteration, the maximum length of a gap to be bridged is increased, while the thresholds for other criteria are only slightly modified. In the first iteration, gaps of less than one road width are checked. In addition to the absolute length of a gap, its length relative to the length of the segments to be connected is also considered. This is done in order to keep short segments from bridging large gaps. To fulfill the collinearity criterion, the angle between two segments has to be below a given threshold. The same is true for the angle between the new segment that bridges the gap and the two connected segments. The angle difference is allowed to be higher only for very small gaps, i.e., smaller than half of the road width. While the maximum gap length increases during this grouping phase, the threshold on the width ratio of two segments stays almost constant. To avoid hard thresholds for a single criterion during the evaluation of a hypothesis for a connection, all criteria are combined into one value. For the verification of new segments, the image information, i.e., the gray values within the new segments and, if necessary, information about the local context, is analyzed. Parallel to increasing the length of the gaps that are allowed to be bridged, short and unconnected hypotheses for road segments are eliminated. The motivation for the elimination of short segments with the increasing number of iterations is the assumption that segments with a high probability for being a part of the road network can easily be grouped with neighboring segments. Therefore, the probability for short and unconnected segments being correct hypotheses decreases with the number of iterations. Figure 7 shows an intermediate result of this grouping process. As can be seen, the mainly collinearity-based strategy fails, especially for highly curved segments.

After increasing the threshold for the distance, in the following iterations the constraints for collinearity are relaxed. During this phase of the grouping, "snakes" (Kass *et al.*, 1988), especially ribbon snakes, become increasingly important. The benefit from using snakes is that the connection between two segments is not only defined by the geometry of these two segments, but also by image features. Snakes work according to the principle of energy minimization. Internal energy enforces geometric constraints, e.g., length and smoothness of a path. Contrary to this, external energy pushes the snake towards image features. By minimizing internal and external energy simultaneously, image information and geometric properties are fused in an effective way. In contrast to the conventional snake, the ribbon snake has an additional parameter for the width at each line point. The image features the ribbon snake deals with are anti-parallel edges on both sides of their center line. Using a ribbon snake, road extraction becomes possible for very fragmented edges and also in cases where only one roadside is visible. Bridging a gap between two road segments using a ribbon snake is performed in two phases. In the first phase, the width of the ribbon is fixed and only the position of its axis is optimized. This is done in analogy to a zip, but starting from both ends (c.f., "ziplock" snake in Neuen-schwander *et al.*, (1995)). In the second phase, only the width is optimized, i.e., adapted to the image features. The hypothesis is accepted if the variance of the width is still low after this second step. The variation of the width can be used as criterion, because random features close to wrong hypotheses cause in most cases a large variation of the width. A more detailed description of this technique is given in Mayer *et al.* (1998). Figure 8 shows an example of a gap bridged using a ribbon snake.

In those cases where the evidence in the image is insufficient to confirm a connection hypothesis, information about the local context of the particular road segment is considered. In other words, a plausible explanation must be given as to why not enough evidence for a road can be found in the image but, nevertheless, the gap can be bridged. In particular, the local context *occlusion_shadow* is important in such cases. The main part of the information needed to explain such situations can be derived from a DSM and information about when and where the image was taken. With this information, shadowed and occluded areas can be predicted. For shadows, the coarse prediction can be refined in the original image. Using a DSM with a resolution of only a few meters, it is possible to detect individual high objects, and to classify them into vegetation and buildings based on their different texture in the image (Eckstein and Steger, 1996). Furthermore, the DSM is useful for eliminating wrong hypotheses for roads that lie on the roofs of buildings.

After the generation of hypotheses for connections and their verification, the road network has to be constructed, i.e., the intersections which link the roads have to be extracted. The generation of hypotheses for intersections is based mainly on geometric calculations. Extracted road segments are extended at unconnected end points. The length of the extension depends on the length and width of the particular segment. If an extension intersects an already existing road segment, a new road segment is constructed, which connects the intersection point with the extended road. Most of these new segments are shorter than five times the road width. Longer extensions are more likely to miss the actual road in the image, because of the uncertainty of the direction at the end points and the geometric variability of roads in intersection areas. The verification of these new road segments is done in the same manner as for the gaps.

Figure 7. Grouping: Intermediate result.



(a) (b)

Figure 8. (a) Connection hypothesis. (b) Gap bridged by a ribbon snake.

Results

The quality of the results depends on the global context. For the rural area, the results are almost correct and complete (Figure 9). Based on their width, agricultural paths, which are also extracted, can easily be distinguished from other roads. The most important condition for an acceptable result is a good contrast between road and background. To achieve good results in different images, radiometric parameters, e.g., thresholds for gray values, are likely to need modifications. Contrary to this, the geometric parameters for the grouping of edges into road segments do not need to be changed. For the grouping of road segments to construct a road network, minor modifications concerning the maximum length of gaps that can be bridged are usually sufficient. Figure 10 shows the result for an image of another rural scene.

In residential areas there are many problems caused by background objects. Roads appear very fragmented in the image and, therefore, right from the start there are fewer correct hypotheses for roads. Moreover, the hypotheses cannot be easily grouped because of the many gaps. Figure 11 points out the limits of the approach. Most of the roads extracted in the open rural area are ending outside the village. A thorough integration of DSMinformation, which has not yet been implemented, is expected to improve the results in this case.

A quantitative evaluation of the results according to the evaluation scheme in Heipke *et al.* (1998) was carried out with a set of test images for which reference data were plotted manually. This evaluation showed that the results for the open rural area are quite reliable and also relatively complete. More than 95 percent are really roads and 80 to 90 percent of the roads have been extracted. The geometric accuracy for the correctly extracted road axes is about one pixel, i.e., 0.3 to 0.5 m. Figure 12 may help to interpret these numbers. Within the open rural area, 91 percent of the roads have been extracted, i.e., the extracted network covers 91 percent of the reference network. In the entire image, still 76 percent of the roads have been found. For the open rural area as well as for the entire image, about 98 percent of the extracted network are really roads, i.e., 98 percent of the extracted network is covered by the reference network.

Figure 9. Extracted roads and intersections.

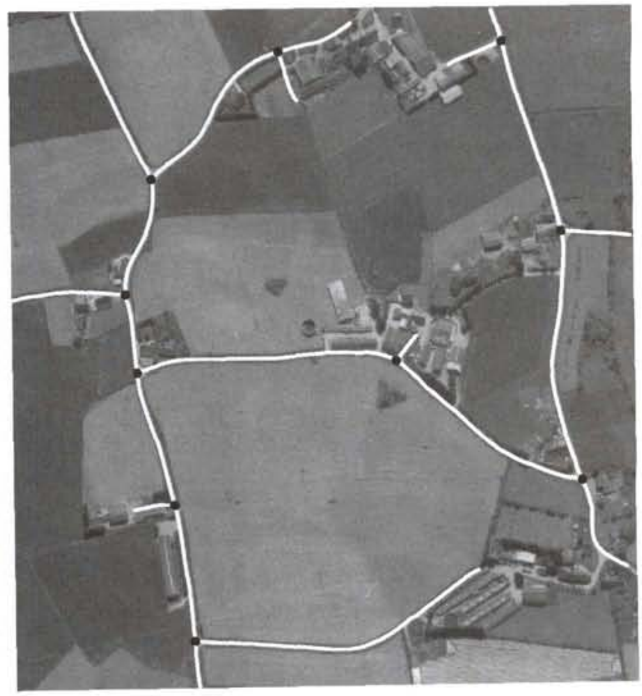


Figure 10. Result for another image.

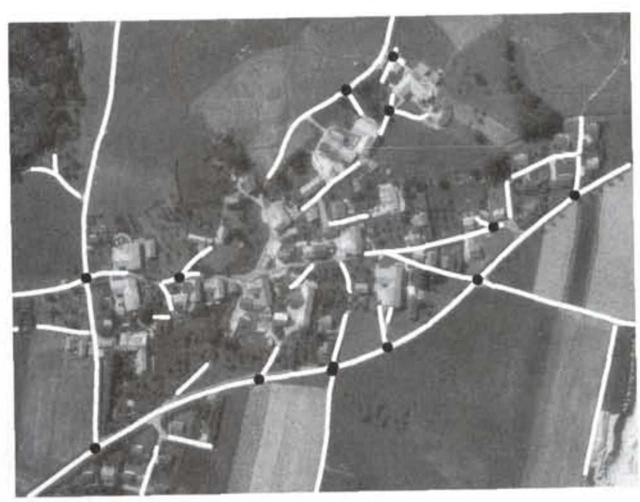
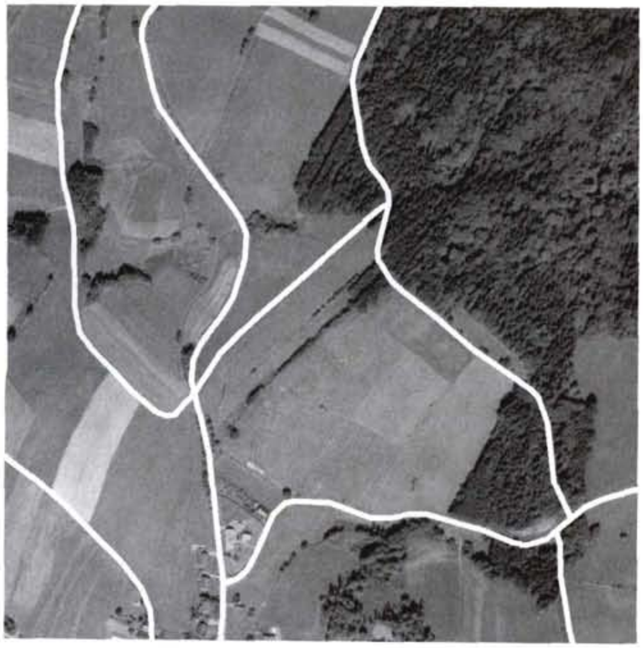


Figure 11. Result for *built-up area.*

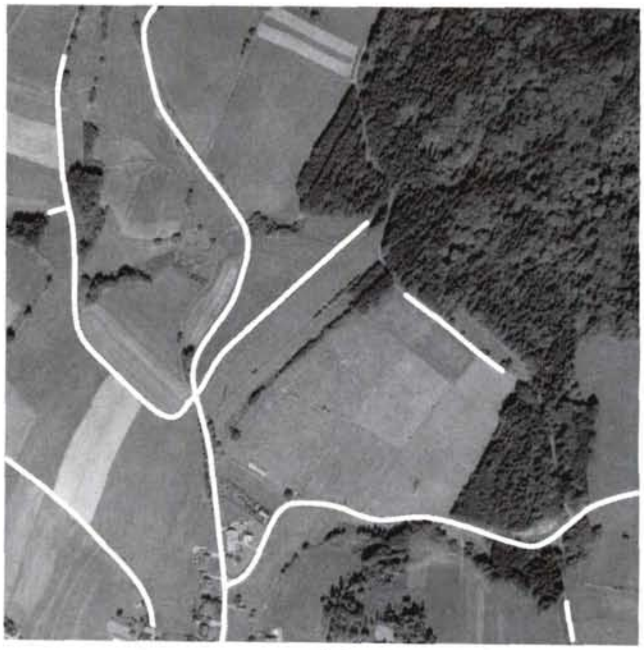
The execution time for the extraction basically depends on the image size and on the number of roads in the scene. Most of all it depends on the initial number of hypotheses for road segments and on how fast this number decreases with each iteration during grouping. On a rather old Sun Sparc 20, the extraction took about 60 minutes for Figure 9 (area: 4 km^2) and about 15 minutes for Figure 10 (area: 0.6 km^2).

Discussion and Outlook

The proposed approach is well suited for road extraction in rural areas in images with a resolution of 0.2 to 0.5 m. A resolution of less than 0.2 m results in a large number of edges and a more inhomogeneous appearance of roads. On the other hand,



(a)



(b)

Figure 12. Evaluation. (a) Reference. (b) Extraction.

for a resolution of more than 0.5 m, the regions for roads become very narrow and for most roads only line extraction is feasible. Whether or not the resolution of future satellite imagery of about 0.8 m is sufficient for the proposed approach still has to be investigated.

Because primarily the local neighbors are considered during the grouping process, it is not guaranteed that always those connections are established that are the best for the whole network. Inferring global information by generating hypotheses for connections, as proposed in Steger *et al.* (1997), could be used as a remedy.

For the extraction of roads in urban and forest areas, the extraction in the open rural area provides quite reliable starting points. However, the propagation of the road extraction into these areas requires additional extraction and tracking algorithms and must be based also on other evidence, e.g., grouping of road markings and vehicles.

At this point, the results are not absolutely reliable and complete. Hence, in operational use, a human operator would be needed to edit the results, i.e., to delete wrongly extracted roads and to insert missing parts. Nevertheless, the approach shows that good results can already be achieved based on relatively simple grouping algorithms. A noticeable improvement seems possible with a more complete integration of context information and global grouping criteria.

Acknowledgments

This work was funded by Deutsche Forschungsgemeinschaft under grant no. Eb 74/8-3.

References

Barzohar, M., M. Cohen, I. Ziskind, and D.B. Cooper, 1997. Robust Method for Automatic Aerial Detection of Occluded Roads Based on Multihypothesis Generalized Kalman Filter, *International Archives of Photogrammetry and Remote Sensing*, 32(Part 3-2W3):1–7.

Baumgartner, A., W. Eckstein, H. Mayer, C. Heipke, and H. Ebner, 1997. Context Supported Road Extraction, *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser Verlag Basel, pp. 299–308.

Bordes, G., G. Giraudon, and O. Jamet, 1997. Road Modeling Based on a Cartographic Database for Aerial Image Interpretation, *Semantic Modeling for the Aquisition of Topographic Information from Images and Maps*, Birkhäuser Verlag Basel, pp. 123–139.

Clément, V., G. Giraudon, S. Houzelle, and F. Sandakly, 1993. Interpretation of Remotely Sensed Images in a Context of Multisensor Fusion Using a Multispecialist Architecture, *IEEE Transactions on Geoscience and Remote Sensing*, 31(4):779–791.

De Gunst, M., and G. Vosselman, 1997. A Semantic Road Model for Aerial Image Interpretation, *Semantic Modeling for the Aquisition of Topographic Information from Images and Maps*, Birkhäuser Verlag Basel, pp. 107–122.

Eckstein, W., and C. Steger, 1996. Fusion of Digital Terrain Models and Texture for Object Extraction, *Proceedings of the Second International Airborne Remote Sensing Conference and Exhibition*, Environmental Research Institute of Michigan, 3:1–10.

Grün, A., and H. Li, 1997. Semi-Automatic Linear Feature Extraction by Dynamic Programming and LSB-Snakes, *Photogrammetric Engineering & Remote Sensing*, 63(8):985–995.

Heipke, C., H. Mayer, C. Wiedemann, and O. Jamet, 1998. External Evaluation of Automatically Extracted Roads, *Photogrammetrie-Fernerkundung-Geoinformation*, (2/98):81–94.

Kass, M., A. Witkin, and D. Terzopoulos, 1988. Snakes: Active Contour Models, *International Journal of Computer Vision*, 1(4):321–331.

Laws, K., 1980. *Texture Image Segmentation*, Ph.D. thesis, Dept. of Engineering, University of Southern California, Los Angeles.

Mayer, H., and C. Steger, 1998. Scale- Events and Their Link to Abstraction for Road Extraction, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(2):62–75.

Mayer, H., I. Laptev, and A. Baumgartner, 1998. Multi-Scale and Snakes for Automatic Road Extraction, *Fifth European Conference on Computer Vision*, Freiburg, Germany, pp. 720–733.

McKeown, Jr., D.M., and J.L. Denlinger, 1988. Cooperative Methods for Road Tracking in Aerial Imagery, *Computer Vision and Pattern Recognition*, IEEE Computer Society Press, Washington, pp. 662–672.

Merlet, N., and J. Zerubia, 1996. New Prospects in Line Detection by Dynamic Programming, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):426–431.

Neuenschwander, W., P. Fua, G. Székely, and O. Kubler, 1995. From Ziplock Snakes to Velcro™ Surfaces, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser Verlag Basel, pp 105–114.

Plietker, B., 1997. Automatisierte Methoden zur ATKIS-Fortführung auf der Basis von digitalen Orthophotos, *Photogrammetric Week '97*, Wichmann Verlag, Karlsruhe, pp. 135–146.

Ruskoné, R., S. Airault, and O. Jamet, 1994. A Road Extraction System Using the Connectivity Properties of the Network, *Zeitschrift für Photogrammetrie und Fernerkundung*, (5/94):174–180.

Ruskoné, R., L. Guigues, S. Airault, and O. Jamet, 1996. Vehicle Detection on Aerial Images: A Structural Approach, *13th International Conference on Pattern Recognition*, Vienna, Austria, pp. 3:900–904.

Steger, C., 1998. An Unbiased Detector of Curvilinear Structures, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(2):113–125.

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

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

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

Tönjes, R., 1997. 3D Reconstruction of Objects from Aerial Iamges Using a GIS, *International Archives of Photogrammetry and Remote Sensing*, 32(Part 3-2W3):140–147.

Trinder, J.C., and Y. Wang, 1998. Knowledge-Based Road Interpretation in Aerial Images, *International Archives of Photogrammetry and Remote Sensing*, 32(Part 4):635–640.

Vosselman, G., and J. de Knecht, 1995. Road Tracing by Profile Matching and Kalman Filtering, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser Verlag Basel, pp. 265–274.

Wiedemann, C., and H. Mayer, 1996. Automatic Verification of Roads in Digital Images Using Profiles, *Mustererkennung 1996*, Springer-Verlag, Berlin, pp. 609–618.