T-REX: TUM RESEARCH ON ROAD EXTRACTION

A. Baumgartner[1], W. Eckstein[2], C. Heipke[3], S. Hinz[1], H. Mayer[4], B. Radig[5], C. Steger[5], C. Wiedemann[1]

1) Lehrstuhl für Photogrammetrie und Fernerkundung
Technische Universität München
Arcisstraße 21, D-80290 München

2) MVTec Software GmbH
Neherstraße 1, D-81675 München

3) Institut für Photogrammetrie und Ingenieurvermessungen
Universität Hannover
Nienburger Straße 1, D-30167 Hannover

4) Institut für Photogrammetrie und Kartographie
Universität der Bundeswehr München
D-85577 Neubiberg

5) Forschungsgruppe Bildverstehen (FG BV)
Informatik IX, Technische Universität München
Orleansstraße 34, D-81667 München

KEY WORDS: Road Extraction, Road Network, Evaluation

ABSTRACT

In this paper we report on research on road extraction at TUM (Technische Universität München). We propose a scheme for road extraction in rural areas that integrates three different modules with specific strengths. The first module employs local grouping and uses multiple scales and context to reliably extract most parts of the road network. In order to connect these parts, the second module exploits the network characteristics of roads for global grouping. The third module completes the network based on an analysis of path lengths within the network. An evaluation of the results shows that the system benefits from the integration of different types of knowledge within the road extraction scheme. In addition to the scheme for rural areas, we present first results of an approach for road extraction in urban areas that focuses on the substructure of roads (markings, lanes) and related objects (vehicles).

1 INTRODUCTION

In the past, the automation of road extraction from digital imagery has received considerable attention. Research on this issue is motivated by the increasing importance of geographic information systems (GIS) and the need for data acquisition and update for GIS.

This paper is a compilation of research on fully automatic road extraction from aerial and satellite imagery that has been carried out at TUM (Technische Universität München) and in which Heinrich Ebner has been significantly involved. For road extraction in rural areas we combine three different approaches, use them as modules of an integrated road extraction scheme, and show how they support each other. The first module uses multiple scales and context information for road extraction and is based on local grouping of lines (hypotheses for road axes) and edges (hypotheses for roadsides) (Steger et al., 1995, Baumgartner et al., 1997, Mayer and Steger, 1998, Baumgartner et al., 1999). It was developed for panchromatic aerial imagery with a resolution of approximately 0.5 m or smaller. It delivers quite reliable hypotheses for roads with a good geometric accuracy. The second module fuses linear structures from various sources and constructs a weighted graph. Pairs of seed points within this graph are selected and shortest paths between these seed pairs are extracted to construct a road network. Compared to the first module, the second module relies on more global grouping criteria (Steger et al., 1997, Wiedemann and Hinz, 1999). The third module completes the road network delivered by the second module. It generates hypotheses for missing connections and verifies these hypotheses based on the image data (Wiedemann, 1999).

Because of the higher complexity of urban scenes, road extraction is much more difficult there than in rural areas. In many cases, roadsides (edges) and road axes (lines) cannot be detected or their relation to roads cannot easily be recognized. In fact, in urban areas substructures like markings or other objects, e.g., vehicles or buildings, often provide the most important image features for road extraction. Based on this experience we are developing a module for road extraction in urban areas. This module starts with the extraction and grouping of road markings and therefore can only be applied to images with a ground resolution of 0.2 m or smaller. The markings are then used to construct lanes and roads (Hinz et al., 1999). In addition, knowledge about relations between lanes and vehicles is used to verify the hypotheses for lanes and roads.

References to the most relevant previous work for the modules used in the road extraction scheme proposed here can be found in the above cited papers. A comprehensive survey on models and strategies for state-of-the-art road extraction is given in (Mayer, 1998, Mayer et al., 1998a). The different approaches cover a wide variety of models and strategies to extract roads automatically from digital aerial and satellite imagery, or at least to automate parts of the manual extraction process. The approaches show promising parts of a road model and extraction strategy. Data from different sources is often useful. For example, in urban areas information derived from a digital suface model (DSM) may help to remove false road hypotheses. What is missing is the use of different resolutions of the image data, e.g., to eliminate disturbances like cars on the roads. Furthermore, it has proven to be very important, that the road model also incorporates contextual information about roads, e.g., the relation to other objects that potentially change the appearance of a road, such as buildings or trees, which may occlude a road or cast shadows over it. The strategy of road tracking is promising in automatic approaches to bridge gaps in the extracted road hypotheses. Local grouping is also very useful in this case. However, the function of roads is never modeled explicitly, and hence the use of global grouping seems to be an essential step to generate a correct and complete network.

Our road extraction scheme is based on the road model presented in Section 2. Section 3 describes the three modules which are used for road extraction in rural areas and evaluates the achieved results. In Section 4 first results of our approach for road extraction in urban areas are presented. A short summary concludes the paper.

2 MODELING ROADS

For the proposed approach, the road 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 represents 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.

A description of roads in the real world can be derived from their function for human beings: roads are defined as an open passage for vehicles, persons, or animals. They are important for transportation between different places. Therefore, roads are organized as a network. The denser an area is inhabited and the more intensively it is used, the denser the road network is. With respect to their importance, network components are classified into a hierarchy of different categories with different attributes. According to the different categories, roads differ with respect to curvature radius and maximum allowed slope. Some important attributes for parts of the road network 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.

The appearance of roads in aerial imagery strongly depends on the sensor's spectral sensitivity and its resolution in object space. In the proposed model only gray-scale images and scale dependencies are considered. In images with low resolution, i.e., more than approximately 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. The attainable geometric accuracy is better, but background objects like cars, trees, or buildings disturb the road extraction more severely.

In a smoothed image — which corresponds to a reduced resolution — lines representing road centerlines can be extracted in a stable manner even in the presence of these background objects, because the smoothing eliminates substructure of the road, e.g., vehicles or markings. This elimination of substructure 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).

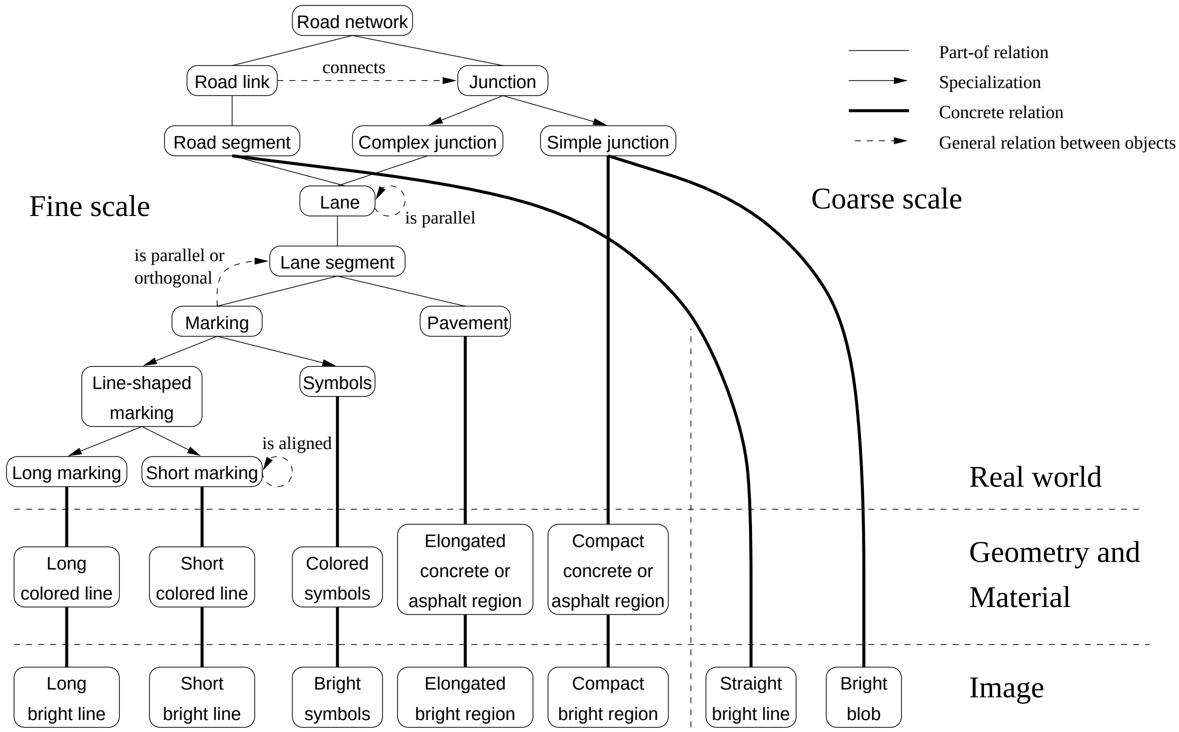


Figure 1: Road model

From the last paragraph it follows that the fusion of low and high resolution can contribute to improve the reliability of road extraction. Additionally, details like road markings, which can be recognized at a resolution of smaller than 0.2 m, can be used as further evidence to validate the detected road hypotheses. On 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 this makes it necessary to combine all features of all resolution levels into one road model.

2.1 Road Model

The road model condensed from the findings above is illustrated in Fig. 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 junctions and road links that connect junctions. While road links are directly constructed from road segments, junctions are further specialized into simple and complex junctions. In fine scale, road segments as well as complex junctions are aggregated from lanes, which consist of pavement and markings. For markings there are two specializations: Symbols and line-shaped markings. The latter define — depending on their orientation with respect to a lane segment — either the side of a lane segment (if parallel) or they define a lane segment's end (if orthogonal). The concepts of the real world are connected to the concepts of the *geometry and material* level via *concrete* relations (Tönjes, 1997), which link 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 in contrast to the *image* level it describes objects independently of sensor characteristics and viewpoint. Road segments are linked to the "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 via the "elongated 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 based on topological criteria, while details, like exact width and position, or markings, are integrated from fine scale. In this way the extraction benefits from both scales.

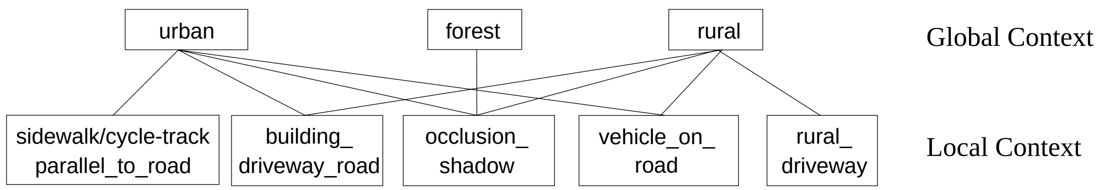


Figure 2: Global and local context

2.2 Context Model

The road model presented above comprises knowledge about radiometric, geometric, and topological characteristics of roads. This model is extended by knowledge about context. The context describes the relations between roads and background objects. Background objects, like buildings, trees, or vehicles, can support (e.g., usually there is a road to every building), but also interfere with road extraction (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*.

With the 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, a paved entrance to an agricultural field. Driveways to buildings are more likely to cause problems in suburban and urban areas. Here, buildings and sidewalks or cycle tracks are organized mostly in parallel to roads. Groups of cars or trucks occlude extended parts of roads. Depending on the employed context model, such situations are potentially hindering or supporting road extraction. Figure 3 illustrates some of the used local contexts. For example, the local context *occlusion_shadow* models a situation in which a high object occludes a part of a road or casts a shadow on a road. Other local contexts are, e.g., *building_driveway_road*, *vehicle_on_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.

Relations to background objects and their relevance for road extraction also depend on the region where they occur. As mentioned above, 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. For instance, in downtown areas, buildings typically are closer and more parallel to roads. 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. Thus, it is clear that the global context enables a more efficient use of the knowledge about roads. In Fig. 2, some frequently occurring local contexts are assigned to the global contexts.

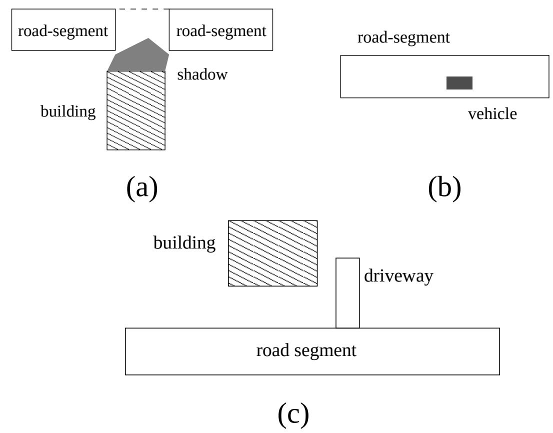


Figure 3: Local contexts: a) *occlusion_shadow*, b) *vehicle_on_road*, c) *building_driveway_road segment*.

Note, however, that the use of knowledge about local context and the verification of specific relations between local objects will in most cases be possible in high resolution imagery only, because the image features that contribute to the local context are usually not very prominent. Therefore, the local context is more tightly connected to the high resolution, whereas information about global context usually can be derived from images with a resolution $> 2\,\mathrm{m}$ and is useful to guide the road extraction in both scales.

(a) (b)

Figure 4: a) Aerial image, b) Manually plotted reference network

3 ROAD EXTRACTION IN RURAL AREAS

For road extraction in rural areas we arrange three modules into a processing chain. The first two modules were designed to extract roads from aerial or satellite imagery without any prior information, which could, for example, be provided by a GIS. Both modules can be applied independently, too. However, there are significant differences between these two modules concerning the required input data and the strategy that is applied. This is why it is useful to combine them. The third module is able to analyze an extracted or already existing road network. It generates hypotheses for missing road segments and, moreover, is able to verify these hypotheses based on image information using the second module.

The image for which we exemplify the strategy of the individual modules and the combination of the three modules is shown in Fig. 4 a). It is a pan-chromatic aerial image covering an area of about 5 km^2 with a ground resolution of approx. 0.5 m. The quality of the results is evaluated by comparing the extracted road network to a manually plotted reference (Fig. 4 b).

3.1 Strategy

The knowledge about how and when individual parts of the model can be exploited optimally is condensed into the extraction strategy. The basic idea of the proposed strategy is to focus the extraction process on those parts of the road network that can be detected most easily and reliably, and that are in addition useful to guide the 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 because of 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. According to the "easiest first" principle, salient road segments are extracted first, then connection hypotheses, i.e., the non-salient road segments, between the salient parts of the road network are captured, and finally the resulting network is analyzed and completed.

The first module (Sect. 3.2) starts with edge and line extraction and selects candidates for roadsides by locally fusing edge and line information, i.e., the scale-space behavior of roads is employed. From these roadsides initial hypotheses for road segments are derived which are then grouped using knowledge about radiometry, local geometry, and local context. The second module (Sect. 3.3) combines the results of the first module with the extracted lines and then enforces the network characteristics of roads, i.e., instead of employing purely local criteria, knowledge about global connectivity is used. The third module (Sect. 3.4) is used to check and improve the road network delivered by the second module with respect to its fitness for use, i.e., hypotheses for missing road segments in the network are generated and verified.

3.2 Module I: Local Extraction

The road extraction module described in this section combines line and edge extraction to detect hypotheses for salient roads, i.e., the parts of the road network that are clearly visible in the image (Sect. 3.2.1). By

(a) (b)

Figure 5: a) Input to the fusion: lines from coarse scale (black), edges from fine scale (white). b) Hypotheses for road segments.

applying local grouping criteria and additional information about local context, non-salient road segments are also extracted (Sect. 3.2.2). With the extraction of junctions (Sect. 3.2.3), a road network is constructed which is then used as input for the second module.

3.2.1 Salient Roads On the local level we use lines and edges as image features to construct road segments. According to the road model, apart from the original image also a version of the image with a reduced resolution is used. The lines extracted in the reduced-resolution image with a pixel size of about 2 m guide the selection of edges extracted from the original resolution that are candidates for roadsides (see Fig. 5 a). In order to reduce the amount of data, lines and edges are approximated by polygons. Here, the term “edge” is used for an individual segment of an edge polygon. Edges that are candidates for roadsides must fulfill the following criteria:

- The distance between pairs of edges must be within a certain range. Minimum and maximum distance depend on the classes of roads to be extracted.
- The edges have to be almost parallel, i.e., there is an overlap and the differences in the direction of the edges is small. The employed threshold decreases for longer edges because the direction is the better defined the longer an edge is.
- The area enclosed by a pair of parallel edges should be quite homogeneous in the direction of the road.
- In addition, in the middle of each pair of candidates for roadsides, a corresponding line must exist in the reduced resolution.

The selection of edges as roadside candidates is done by a local fusion of line and edge extraction. This local fusion is described in detail in (Steger et al., 1995). The fusion of lines from low resolution and edges from high resolution has proven to be very useful in order to obtain reliable results.

From the roadsides, road segments are constructed (Fig. 5 b). The road segments consist of quadrilaterals which are generated from parallel roadside candidates. Quadrilaterals sharing points with neighboring quadrilaterals are connected. The geometry of the road segments is represented by the points of their medial axes, attributed by the road width. These road segments are the semantic objects that are used as input for the extraction of the non-salient parts of the road network.

3.2.2 Non-salient Roads In the previous step only a small part of the knowledge about roads was exploited. The extracted road segments are clearly visible in the image, and have a high probability for being roads based on local criteria. Non-salient roads are the parts of the road network that could not be extracted due to a lack of suitable image features. For the extraction of non-salient road segments, additional knowledge about roads must be applied. We assume that non-salient roads correspond to gaps between salient road segments. Therefore, the extraction of non-salient roads is equivalent to the problem of linking salient roads extracted in Sect. 3.2.1. In addition to the extraction of non-salient road segments, we want to eliminate incorrect hypotheses for salient road segments.

Most of the road segments derived from the fusion of line and edge extraction are not directly connected and are quite short. The linking of correct and the elimination of false hypotheses is achieved by grouping the salient road segments into longer segments. The grouping is done according to the "hypothesize and test" paradigm. Hypotheses about which gaps should be bridged are generated starting with geometric criteria (absolute and relative distance, collinearity, width ratio) and radiometric criteria (mean gray value, standard deviation). Because information about only two road segments is involved, we call this local grouping, in contrast to the grouping in Sect. 3.3 which additionally uses global criteria. The hypothesized road segments are verified in the image. The verification consists of up to three stages: In the first stage, 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. Otherwise, the verification switches to the second stage. Here, a so-called "ribbon snake" is applied to the gradient image to find an optimum path for the link. If this verification also fails, a third stage is used, in which we try to derive an explanation by local context. The local context is used as last and weakest verification method to explain and close gaps.

According to the above mentioned criteria, hypotheses for connections are generated and verified iteratively. For every new iteration the allowed maximum length of a gap to be bridged is increased, while the thresholds for other criteria are only slightly relaxed. To avoid hard thresholds for a single criterion in the evaluation of a hypothesis for a connection, all criteria are combined into one value. In parallel to increasing the maximum length of the gaps that are allowed to be bridged, short and unconnected hypotheses for road segments, i.e., hypotheses that are false with a high probability, are eliminated. Figure 6 displays an intermediate result of the grouping of the road segments. This mainly collinearity-based strategy sometimes fails, especially for curved segments.

After increasing the threshold for the absolute distance, in the following iterations the constraints for collinearity are relaxed as well. During this phase of the grouping, snakes, (Kass et al., 1988), especially ribbon snakes, become increasingly important. Snakes work according to the principle of energy minimization: The so-called "internal energy" enforces geometric constraints, e.g., smoothness of a path. In contrast, the so-called "external energy" pulls the snake towards image features. By minimizing internal and external energy simultaneously, geometric properties and image information are fused. As an extension to the conventional snake, the ribbon snake has an additional parameter for the width at each line point. The image features the ribbon snake is attracted to are anti-parallel edges on both sides of its center line. By using ribbon snakes, road extraction becomes feasible for very fragmented edges and in cases where only one roadside is visible. Bridging a gap between two road segments 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-lock, but starting from both ends. , c.f., "zip-lock" snake in (Neuenschwander et al., 1995). The zip-lock behavior of the ribbon is achieved by splitting the ribbon in active and passive parts. Only the active parts are optimized. Figure 7 illustrates the zip-lock ribbon.

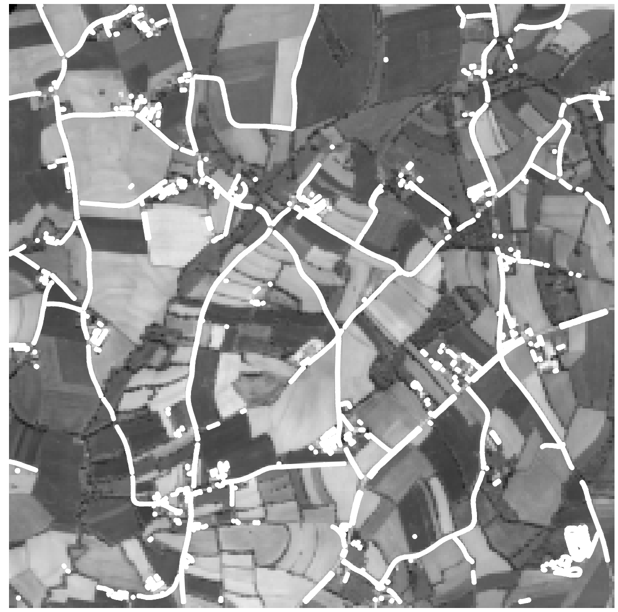
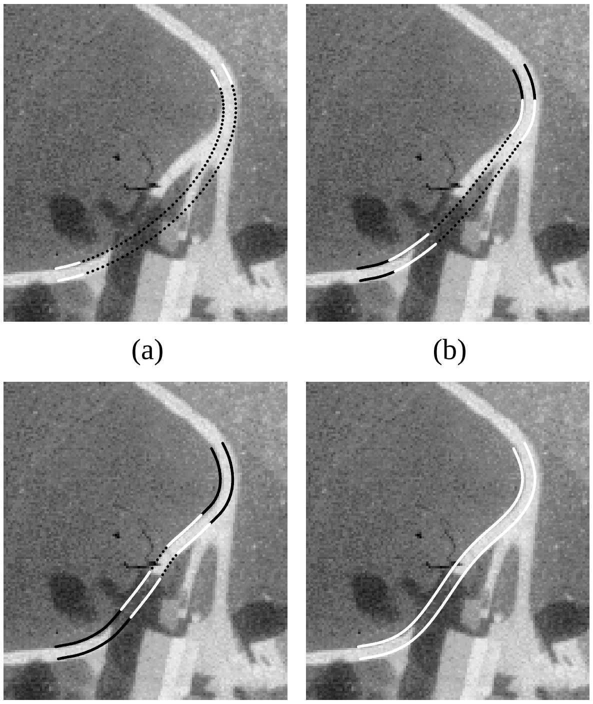


Figure 6: Road segments, intermediate result

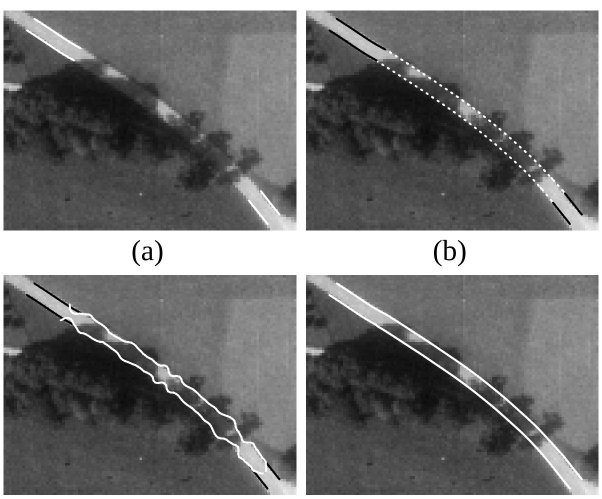
In the second phase, only the width is optimized, i.e., adapted to the image features. The hypothesis is accepted if the variance of the width is still low after this second step. Figure 8 shows that this is possible even if the roadsides do not correspond to strong edges in the image. A more detailed description of this technique is given in (Mayer et al., 1998b).

In cases in which the evidence in the image is insufficient to confirm a connection hypothesis, information about the local context of the particular road segment is considered. A plausible explanation must be given why too little evidence for a road exists in the image. If this explanation was found, the gap is



(a) (b) (c) (d)

Figure 7: Optimization steps of a “zip-lock” ribbon. (a)-(c) Dotted lines indicate the passive part of the ribbon. White parts are currently optimized. Black ends indicate the result of the optimization so far. (d) Final result

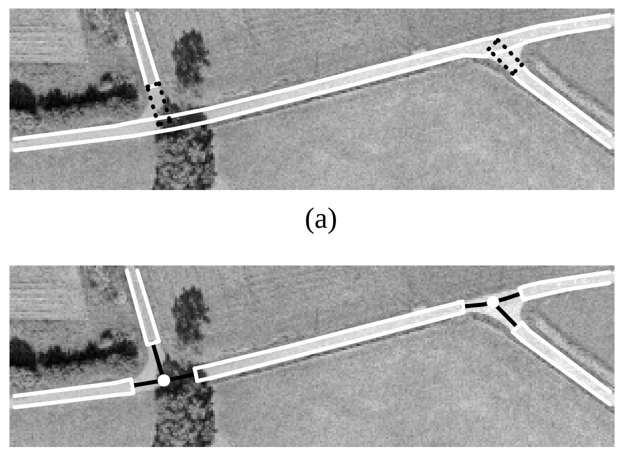


(a) (b) (c) (d)

Figure 8: Extraction of *non-salient* roads. (a) Selection of initial hypothesis (b) Optimal path; (c) Verification by optimization of width (d) Selection of hypothesis with constant width

bridged. In this case, especially the local context *occlusion_shadow* is important. The main part of the information needed can be derived from a DSM and information about when and where the image was taken (Eckstein and Steger, 1996). With this information, shadowed and occluded areas can be predicted and used to explain the gap. However, the information about background objects is not required with a high level of detail and accuracy.

3.2.3 Road Junctions After the generation of hypotheses for connections and their verification, the road network must be constructed, i.e., the junctions that link the roads must be extracted. The generation of hypotheses for junctions is mainly based on geometric calculations: Extracted road segments are extended at their unconnected end points. The length of the extension depends on 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 (Fig. 9). The verification of these new road segments is done in the same manner as for the gaps.



(a) (b)

Figure 9: a) Road segments (white) and extensions (black, dotted) b) Road segments and final junctions

Ideally, after this step all road hypotheses are connected, and there is a path between every pair of points on the extracted road network. However, such a result cannot be expected (see Fig. 10) in real examples. First, because of the limited size of the images some of the nodes will lie outside of the image. Second, the results are not error-free. Especially in urban and forest areas only fragments of the network are likely to be extracted. Because the extraction is primarily based on local information and is reliable only in rural areas, the network characteristics of roads are not optimally exploited. 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.

Up to now primarily local criteria were applied to construct a road network. In Sect. 3.3 we make use of

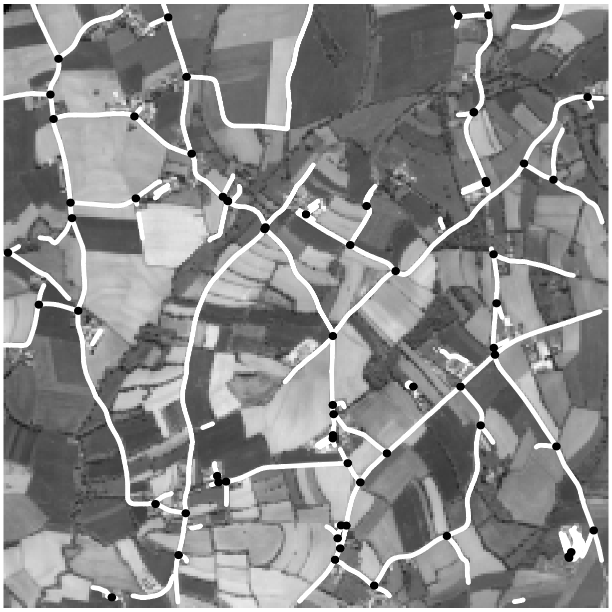


Figure 10: Road axes and junctions

global topological properties of roads to further improve the results with respect to the connectivity of the road network.

3.3 Module II: Global Extraction

The intrinsic function of roads is to connect different "important places," even if they are far away from each other. Hence, roads form a (hierarchical) network that is mostly optimized to provide an economic and convenient way for reaching different places. Because of this property, searching for the globally best connection between such places is an essential step for road extraction. Moreover, since there usually exists only one good connection between two "important places" (at least in open rural terrain) we restrict the search to the best connection between two places.

However, as described in the previous section, the first module does not make use of this property of the road network. Furthermore, a resolution smaller than approximately 0.5 m is needed for the extraction, which might not be available in every case. Thus, the system should not exclusively rely on the results of the local level. In the next step, a module is applied which was independently designed for the extraction of roads from multi-spectral satellite imagery (Wiedemann and Hinz, 1999). Due to the generally lower resolution of satellite images, only the basic shape and reflectance features of roads are used, whereas the global topology of roads, i.e., the road *network*, plays the major role.

The module incorporates a flexible design in order to integrate and process road-like structures from various sources, e.g., lines extracted from different channels having different resolutions and results of other road extraction modules. This flexibility allows us to integrate the road extraction results obtained by the first module, as described in Sect. 3.3.2.

Having these requirements in mind, the following extraction strategy has been developed: First, bright lines are extracted and processed in order to build up road segments. In contrast to Sect. 3.2, these road segments are exclusively based on coarse scale (see also Fig. 1) and therefore, they are less constrained but also less reliable. Hence, a fusion operation merges these road segments with the road segments resulting from the more precise local extraction. After fusion a weighted graph is constructed from the remaining road segments and from connection hypotheses between them. Then, pairs of "important places" are selected and the optimal path between each pair is calculated. The extracted road network results from the combination of all optimal paths computed through the graph. In the following, a detailed description of each step is given.

3.3.1 Low Level Processing Module II usually starts with line extraction from the low resolution image. It is performed using the differential geometric approach of (Steger, 1998) which returns lines as a set of pixel chains and junction points in sub-pixel precision. Additionally, local line attributes like width, direction, and contrast are obtained at each line pixel. Of course, the result is not complete and contains false alarms, i.e., some roads are not extracted and some extracted lines are not roads.

Instead of the described line extraction step, we can also use the lines extracted in Sect. 3.2. There, only the line position was needed to construct the salient road segments, whereas now we will make use of the line attributes, too.

During the next step, road segments are generated from the lines and their attributes. The attribute values are used to include additional evidence about the presence of roads into the grouping process at a later stage. Therefore, it is advantageous if the lines completely correspond either to roads or to linear structures not being roads, i.e., lines should be split at the point where they might cross the roadside. A careful analysis of the behavior of several line attributes has shown that the most significant feature for a change in the line semantics ("road"/ "non-road") is high curvature. Hence, lines are split at points in which the direction difference between two consecutive polygon points on the line exceeds a given threshold. Note that this procedure does not eliminate any part of a line. A line might be erroneously split, i.e., the line is split although it completely belongs to a highly curved road. However, there is still a high probability that the split parts are joined again at a later stage if they are found to be part of the paths computed during road network generation.

Each resulting line defines a road segment. The following attribute values are calculated to obtain an extended description of each road segment:

- length,
- straightness, i.e., standard deviation of its direction,
- width, i.e., mean width of the extracted line,
- width constancy, i.e., standard deviation of the width,
- reflectance constancy of a road segment, i.e., standard deviation of the intensity values along the segment,

3.3.2 Fusion In this step, we combine the road segments obtained from the previously extracted lines with the road segments of the local extraction. After the fusion, both types of road segments are contained in one set of linear road data. Segments or parts of segments that lie within a buffer with a suitably chosen width are candidates for a unification. In addition, if two candidates have a direction difference less than a certain threshold they are unified. Otherwise, they are checked for an intersection. Since the road segments achieved from the low resolution image are less constrained, a more complete network might be extracted compared to the results of the local extraction.

3.3.3 Graph Representation From the fused road segments, a graph is constructed. The road segments define the edges of the graph, and their end points represent the set of vertices. In case of junctions, i.e., if two or more road segments end in the same point, only one vertex of the appropriate degree is inserted in order to preserve the topology.

The attribute values of the road segments are used to weigh the graph by associating every edge with a single weight. This is done by defining linear fuzzy functions ranging from 0 to 1 for transforming the attribute values into fuzzy values. These fuzzy values are aggregated by the fuzzy *and* operation into one overall fuzzy value for each edge. An overall fuzzy value 1 stands for a road segment that matches the road model perfectly and 0 means that the road segment should not be considered any further.

The following phase of processing addresses the preparation of the graph for detecting possibly missing road junctions at a later stage. Due to deviations from the road model, some of the road junctions, especially the larger ones, might have been missed during the detection. Junctions are, however, an essential topological part of the road network. Hence, connection hypotheses should be formed in situations in which a junction might be present. For this reason, the edges of the graph are split at points which can be regarded as a priori candidates for junctions, and a new vertex of degree 2 is inserted. As an example consider point P in Fig. 11 that lies on segment S1 closest to the end of segment S2.

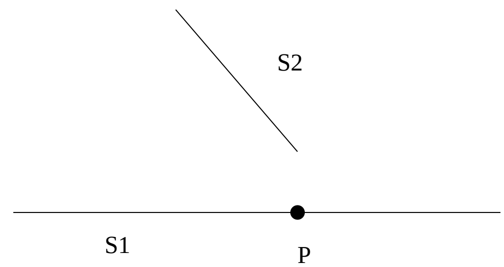


Figure 11: Candidate for a junction

A reliable decision whether a junction candidate truly represents a road junction is not possible at this stage of processing. A false candidate can, for instance, be caused by a blunder, e.g., by other linear structures close to a road like certain kinds of vegetation. However, the splitting of a road segment affects its attributes, e.g., its length, which leads to an incorrect evaluation in such cases. Therefore, the fuzzy value of a split edge is inherited.

Ideally, the resulting weighted graph comprises the whole road network (and possibly other less weighted linear structures). In practice however, some parts of the road network remain undetected, since a road model can hardly cover the full variety of a road's appearance. Hence, we generate and evaluate additional connection hypotheses between edges of the graph (see Fig. 12). The following criteria are introduced to measure the quality of a hypothesis:

- the direction difference between adjacent road segments; either collinearity (within a road) or orthogonality (at a T-junction) are assumed as reference,
- the absolute length of a connection,
- the relative length of a connection compared to the length of the adjacent road segments with the lower weight,
- an additional constraint which avoids that a connection hypothesis is assigned a higher weight than its adjacent road segments.

(a) Weighted graph of road segments, (b) with inserted connection hypotheses

Figure 12: Weighted graph (dark=high rating, bright=low rating)

As above for road segments, linear fuzzy functions are defined to obtain individual fuzzy values for each criterion, which are then aggregated into an overall fuzzy value by the fuzzy *and* operation. A special case is the evaluation of the direction difference between two road segments. In order to either prefer the continuation of a road or to support a possible road junction, a fuzzy function with two peaks is defined (e.g., at 0° and 90°), one supporting the collinearity of two segments and one supporting their connectivity with respect to a T-junction, see Fig. 13. Since a connection hypothesis can represent only one of these grouping principles, but not both at the same time, the proposed examination of the direction difference can be understood as a classification of the connection hypotheses in road connections and junction connections, whereby each connection is associated with a fuzzy value. Depending on this classification one may choose different parameter settings of some of the other fuzzy functions, e.g., for evaluating the absolute distance.

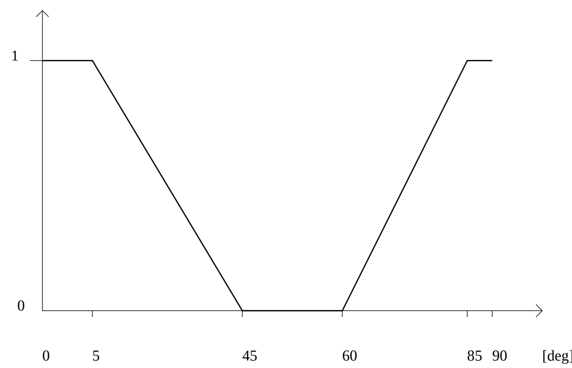


Figure 13: Fuzzy function for evaluating direction difference

3.3.4 Road Network Generation The extraction of the road network relies on the selection of "important places," i.e., seeds, which are then connected by the optimal path through the network. Such places are usually buildings, industrial areas and other sites of interest. Since this approach exclusively deals with roads without considering additional objects, i.e., context, we define "important places" as road segments that represent portions of the road network with high probability. An indication for the probability of a road segment being truly a road is its fuzzy value. Hence, all road segments yielding a high evaluation are chosen as seeds for road network generation. Note that this threshold can be derived from semantically meaningful and reasonable parameters using the same fuzzy functions as before, e.g., by demanding that a seed must have a given minimum length.

Since the image domain only covers a part of the whole road network, it might happen that several disconnected sub-networks are visible in the image instead of one complete road network. Thus, the strategy for road network extraction should be able to extract disconnected road networks but should incorporate the function of roads connecting places far away from each other. To this end, only those pairs of seeds are considered for path calculation that guarantee that the length of the resulting path exceeds some suitably chosen threshold, e.g., 1 km. A lower bound of the respective path length can be obtained without executing path calculation by summing the length of both road segments and the minimal distance between their endpoints (see Fig. 14). By this means, larger isolated parts of the road network can be detected without losing the postulated globality of the proposed grouping algorithm.

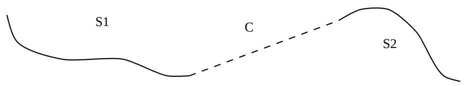


Figure 14: Minimum path length: l_{min} = l_{S1} + l_{S2} + l_C

The basic idea of the proposed algorithm for road network generation is to find the shortest paths in the weighted graph with suitably chosen distances. Therefore, the weights in the graph should reflect the true distances in the object, but depending on how a road segment or a connection hypothesis is evaluated, the distance between two vertices should be increased to make it harder to bridge obviously bad links. If a link is considered perfect, i.e., has the fuzzy value 1, the true distance between the vertices is used. If the link has the fuzzy value 0, its weight is set to ∞. The following formula is used to construct the weighted graph:

$$w_{i,j} = \begin{cases} l_{i,j}/r_{i,j} & \text{if vertices } i \text{ and } j \text{ are connected} \\ & \text{by a road segment of length } l_{i,j} \\ & \text{and } r_{i,j} > 0 \\ d_{i,j}/r_{i,j} & \text{if } i \text{ and } j \text{ are not connected in} \\ & \text{the original graph and } r_{i,j} > 0 \\ & (d_{i,j} \text{ is the Euclidean distance} \\ & \text{between vertices } i \text{ and } j) \\ \infty & \text{otherwise (no edge in the graph)} \end{cases}$$

where $w_{i,j}$ is the weight of the edge between the vertices i and j and $r_{i,j}$ is the corresponding fuzzy value.

The final step is the computation of the optimal path between each seed pair. This is carried out by using the Dijkstra-Algorithm (Knuth, 1994). The combination of all detected paths defines the extracted road network.

The road network shown in Fig. 15 has been extracted using the results of Sect. 3.2 and lines extracted at a resolution of 2 m as input for the module described in this section. When comparing Fig. 15 and Fig. 10 one may recognize that the global extraction was able to improve the connectivity of the network (see, e.g., the lower left corner of the image). On the other hand, some short portions of the road network extracted in Sect. 3.2 have been deleted because the algorithm could not connect them with the main part of the network (see, e.g., the upper right corner). It should also be noted that the resulting network is inhomogeneous with respect to the geometric accuracy since parts of the network originate from purely geometry-based gap bridging without considering the radiometric content in between. This implies that a final verification of the bridged gaps using similar criteria as in Sect. 3.2 should be used. A more detailed quantitative evaluation and comparison is given in Sect. 3.5.

Figure 15: Result after global grouping.

3.4 Module III: Network Completion

Despite the local and global grouping described in the preceeding sections, the resulting road network is in general still incomplete and fragmented. Completeness and connectivity of the extracted network can be further improved by the use of the global network structure of roads.

The division of labor in our modern business world demands a transportation network that allows fast, cheap, efficient, and secure transports. The same characteristics are expected by the people for their daily ride to work, for shopping, and for their trips to recreation areas.

Additional factors that influence the design of the transportation network are, e.g., local topography, land use, and environmental conservation. All these requirements are taken into account for the development of the road network (as part of the whole transportation network). Therefore, they can and should be used for the extraction of road networks from images.

In this section, an approach is presented for the completion of extracted road networks. It is based on the function of roads as part of the transportation network. A strategy for the generation of link hypotheses is proposed which makes use of this function of roads. These hypotheses are verified based on the image data.

3.4.1 Generation of link hypotheses As an example for the proposed algorithm consider Fig. 16. Figure 16a shows a part of a sample network which consists of four nodes (A, B, C, D) and three edges (AB, BC, CD). In the first step, between all possible pairs of points which lie on the network (the nodes A, B, C, and D in the example) the distance along the shortest path within the existing network (network distance, nd) as well as the distance along a hypothetical optimal path (optimal distance, od) are calculated, where, e.g., nd_{BD} is the sum of nd_{BC} and nd_{CD} (see Fig. 16b). These distances are intended to quantify the requirements for fast and cheap transports as well as the additional factors influencing the road network design mentioned above. Therefore, the network distance depends on the actual length and road class along which the shortest path has been found. The optimal distance depends, besides the actual distance between the two points, on factors like topography, land use, and environmental conservation, provided that such information is reliably available.

In the second step, preliminary link hypotheses are defined between each possible pair of points. A so-called "detour factor" is calculated for each preliminary link hypothesis according to the following definition:

$$\textit{detour factor} = \frac{\textit{network distance}}{\textit{optimal distance}}$$

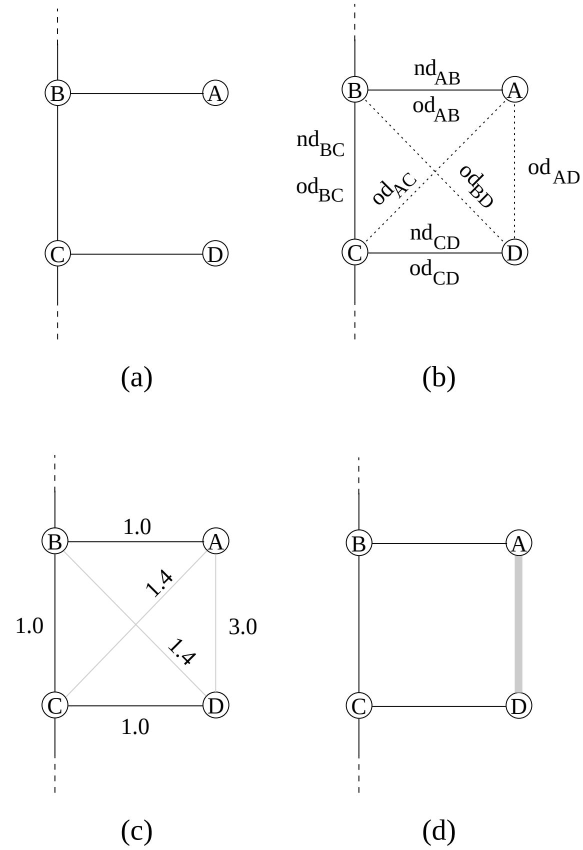


Figure 16: Hypothesis generation: a) Sample Network; b) Network distances and optimal distances; c) Detour factors for all preliminary link hypotheses; d) Link hypothesis

In Fig. 16c the detour factors for all preliminary link hypotheses are shown. For reasons of simplicity, both the network distance and the optimal distance are set to the Euclidean distance between the respective points.

The third step consists of a selection of potentially relevant link hypotheses. The selection is based on the assumptions that only links that have a locally maximum detour factor are of interest and that there is no preferred direction within the road network. Based on these assumptions, a non-maximum suppression (NMS) is performed on the set of preliminary link hypotheses: a link hypothesis is only kept if there is no competing link hypothesis that has a higher detour factor. Competing link hypotheses are preliminary link hypotheses between one end point of the preliminary link hypothesis under investigation and a point adjacent to the other end point.

In the above example only the link hypothesis AD passes the NMS (see Fig. 16d). In general, however, more than one link hypothesis will be kept. All these link hypotheses are sorted according to their detour factor and the one with the highest detour factor is sent to a module that verifies the link hypothesis based on the image data (see Sect. 3.4.2). If the link hypothesis is rejected, the next one (the one with the second highest detour factor) is sent to the verification module, and so on. If a link hypothesis is accepted (and geometrically improved according to the image data), it is inserted into the road network. This insertion changes the whole topology of the road network. Therefore, the procedure of generating link hypotheses has to be repeated from the beginning. Link hypotheses already rejected are not taken into account anymore in the following iteration. The iterative process of determining link hypotheses with maximum detour factor and verifying them can be stopped if it is likely that no further link hypothesis will be accepted by the verification module. In general, this state cannot be predicted reliably, but it can be estimated roughly, e.g., based on the highest detour factor which occurs in the current iteration.

3.4.2 Verification of link hypotheses The verification of the link hypotheses must be done based on the image data. It should provide the information if the link hypothesis can be accepted or not, and, in the case of acceptance, the exact geometry of the connecting road. Because of the modular design of the whole approach, every road extraction tool which is able to extract a road between two given points and which provides some kind of self-diagnosis in order to decide whether the connection can be accepted or has to be rejected can be used to verify a link hypothesis.

Here, the road extraction approach described in Sect. 3.3 is used for the verification of the link hypotheses. The seeds that are necessary for the extraction of roads using this approach are given by the two end points of the link hypothesis. To verify the link hypothesis, the optimal path between the end points of the link hypothesis is calculated through the weighted graph (see Sect. 3.3.4). If the graph provides no connection between the two end points, the link hypothesis is rejected. If a path can be found, the link hypothesis is accepted, geometrically improved, i.e., replaced by the extraction result, and inserted into the road network.

To avoid the extraction of an already existing connection and to reduce the amount of computation time, the search for a road which connects the two end points of a link hypothesis is performed only in a restricted region of interest (ROI) which contains both end points and which is assumed to contain the connecting road as well.

3.4.3 Insertion of accepted link hypotheses into the road network If a road which connects the two end points of the link hypothesis has been found, this new road must be inserted into the whole road network.

First, all parts of the new road that are redundant with respect to the existing road network are eliminated (see Fig. 17). In most cases, one large part of the new link will remain. This part is then inserted into the network by connecting its two end points directly with the respective nearest points of the road network. If this point is not an end point of a road, a new junction is inserted into the road network. In cases where more than one part of the new link exists, all these parts are inserted into the network as described above. If the whole new link has been eliminated, no part can be inserted into the road network, i.e., the respective link hypothesis is rejected.

3.4.4 Results Figure 18 shows the result of the approach applied to the extracted road network derived in the preceeding sections (see Fig. 15); the reference data for this road network are given in Fig. 4 b). The verification of the link hypotheses was carried out based on the image which was down-sampled to

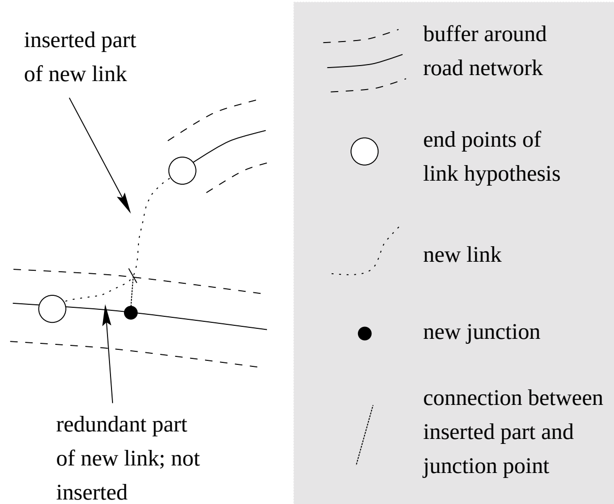


Figure 17: Insertion of a new link

Figure 18: Result of the completion of the extracted road network; the numbers refer to Fig. 19

a ground pixel size of 2 m. Most of the gaps within connected components were closed.

In Fig. 19 three link hypotheses are shown projected on the downsampled image that was used for the verification of the hypotheses. The black lines represent the incomplete road network and the two white points are the endpoints of the link hypothesis. If the link hypothesis was accepted, a white line displays the geometrically improved link that is inserted into the road network. The first example (see Fig. 19 a) shows the upper link hypothesis at the right image border of Fig. 19, which has been accepted. The initial extraction failed because of occlusions and shadows cast on the road. The verification module was able to extract the missing part of the road correctly based on the knowledge that a connection is topologically desirable. The link hypothesis from the upper left corner of Fig. 19 is displayed in Fig. 19 b). In this case, the varying road width and partly missing roadsides prevented the extraction of the missing part. Again, this link was accepted correctly. The third example (see Fig. 19 c) shows the rejected link hypothesis from the lower right image border of Fig. 19. Although it was proposed due to the network topology, no road which directly connects the two end points of the link hypothesis could be detected by the verification module. This example underlines the necessity for a verification of the link hypotheses based on image data.

3.5 Evaluation

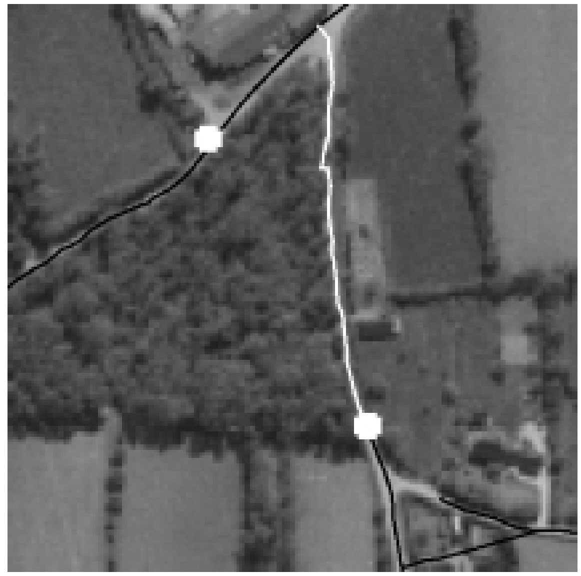
Internal self-diagnosis and external evaluation of the obtained results are essential for any automatic system. In the long run these factors are of major importance for the introduction of the system into practice. Both, internal self-diagnosis and external evaluation should yield quantitative measures which make different results commensurable. Here, we deal with the external evaluation of the automatically extracted road data by means of comparing them to manually plotted linear road axes used as reference data. In the following, the evaluation procedure is briefly described. More details can be found in (Heipke et al., 1998).

The comparison is carried out by matching the extracted data to the reference data using a so-called "buffer method," in which every portion of one network within a given distance (buffer width) from the other one is considered as matched. For the evaluation of the road extraction results, a number of quality measures is defined based on the matching results. Two questions can be answered by the quality measures: (1) how complete is the extracted road network, and (2) how correct is the extracted network? The completeness indicates how much of a given reference was successfully extracted, whereas the correctness is related to the probability of an extraction result to be indeed a road.

In addition, the geometric accuracy of the extraction is assessed. It is expressed as the RMS difference between the matched extracted and the matched reference data.

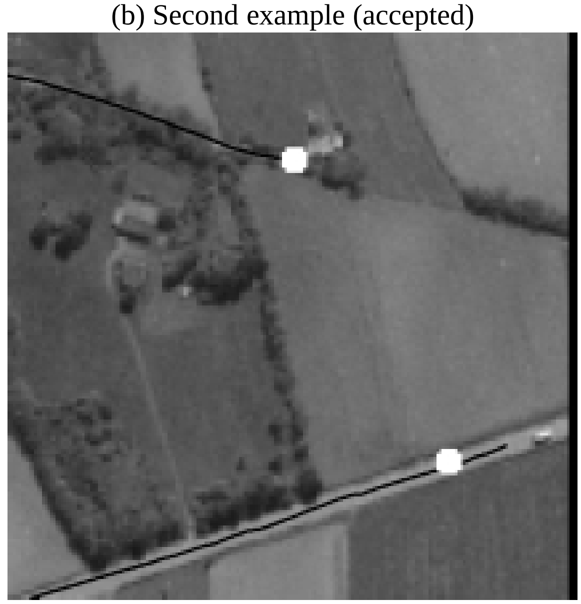
Besides the intuitively feasible measures completeness, correctness, and RMS, an evaluation of the topology of the extracted network is carried out. To this end, two new measures are introduced: *connectivity* and *mean detour factor*.

For the evaluation of the connectivity of the extracted network, a number of points P_i are defined equally



(a) First example (accepted)

(b) Second example (accepted)



(c) Third example (rejected)

Figure 19: Examples for verification of link hypotheses

distributed within the reference network. All possible pairs of these points are examined if they are connected in the reference network, i.e., if they lie within the same connected component. For these **CR** pairs **c**onnected in the **r**eference it is checked whether they are connected in the extracted network as well. This yields **CB** pairs which are **c**onnected in **b**oth networks. Based on **CR** and **CB**, *connectivity* is defined as

$$\begin{aligned} \textit{connectivity} &= \frac{CB}{CR} \\ &= \frac{\textit{\# of pairs connected in both networks}}{\textit{\# of pairs connected in reference network}} \end{aligned}$$

The optimum value of the *connectivity* is 100%. The *connectivity* decreases with an increasing fragmentation of the extracted network with respect to the reference network.

For the evaluation of the topological correctness within connected components of the extracted network, the mean detour factor with respect to the reference network is calculated: the distance along the reference network (*network distance*$^{ref}_{i,j}$) and the distance along the extracted network (*network distance*$^{extr}_{i,j}$) are calculated between all pairs $(i, j), i \neq j$, of points which are connected in both networks. The ratio $r_{i,j}$ between these two distances is calculated for each pair (i, j):

$$r_{i,j} = \frac{\textit{network distance}^{extr}_{i,j}}{\textit{network distance}^{ref}_{i,j}}$$

If $r_{i,j}$ is larger than one, the distance between points P_i and P_j along the extracted network is larger than the respective distance along the reference network. In this case it is referred to as *detour factor*$^{ref}_{i,j}$ (detour factor with respect to the reference network). The *mean detour factor* is defined as the mean of all values *detour factor*$^{ref}_{i,j}$. The optimum value for the *mean detour factor* is 1.0.

The *mean detour factor* increases with the amount of missing connections within connected components of the extracted network and with the degree of “wiggling” extraction (see Fig. 20).

In Table 1 the evaluation figures for the extracted road network of Sect. 3.4 (Fig. 18) and for the intermediate results of Sect. 3.2 (Fig. 10) and Sect. 3.3 (Fig. 15) are given. The figures show no significant changes in *completeness* and *correctness* from the intermediate results to to final result.

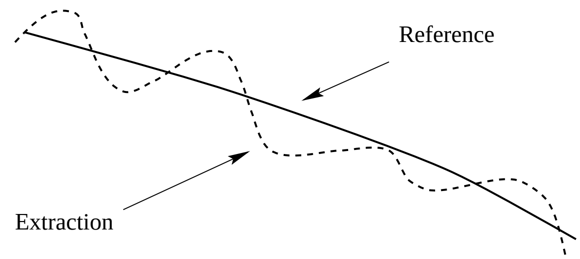


Figure 20: Wiggling extraction

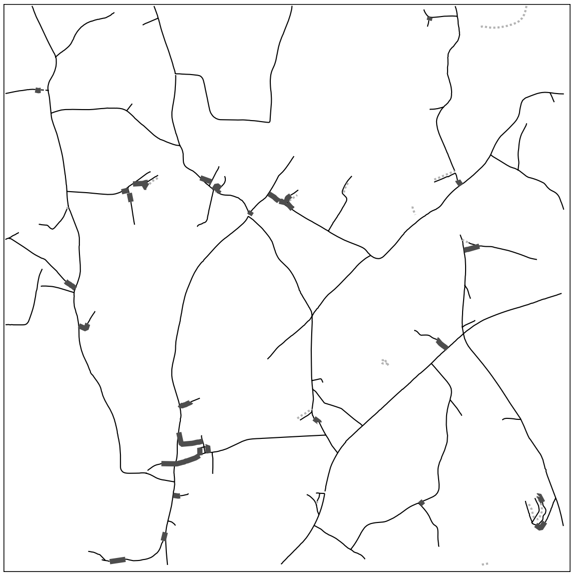
	Sect. 3.2	Sect. 3.3	Sect. 3.4
Completeness [%]	80.68	80.55	83.24
Correctness [%]	92.66	91.34	91.24
RMS [m]	1.05	1.38	1.38
Connectivity [%]	77.41	100.00	100.00
Mean detour factor	1.43	1.18	1.04

Table 1: Evaluation of intermediate and final results

The geometric accuracy of the correct parts of the final result is about 1.38 m, whereas the accuracy of the results of Sect. 3.2 is about 1.05 m. The reason for this decreasing accuracy is that all additionally detected roads result either from line extraction in low resolution (here 2 m) or from connection hypotheses between the lines which are inserted as straight segments without considering the radiometry in between. The global grouping step increases the *connectivity* up to 100%, i.e., all roads which are connected in the reference are connected in the extracted network, too. The most important contribution of the completion module (Sect. 3.4) is the significant reduction of the *mean detour factor* to 1.04. This means that all connections between two points within the extracted network have approximately the same length as in the reference. In terms of *completeness* this module causes only a small improvement. However, it adds road segments that are very important for the function of the network, e.g., for transportation purposes. The strength of the local module (Sect. 3.2) are the high *completeness* and *correctness* rates as well as the good geometric accuracy, which is obtained by the use of edge extraction in addition to the line extraction.

In Fig. 21, the changes from the intermediate results of Sect. 3.2, Sect. 3.3, and from Sect. 3.3 to the final result are marked. Added road segments are displayed as bold lines removed segments as dotted lines. The global step adds many new segments and eliminates relatively short isolated segments (Fig. 21 a). However, some of the removed segments were correctly extracted as roads. The segments added by the completion module are displayed as bold lines in Fig. 21 b).

An evaluation of results on different test images has



(a)

(b)

Figure 21: Differences between intermediate and final results (bold lines: additions; dotted lines: eliminations): a) Changes from Sect. 3.2 to Sect. 3.3 b) Changes from Sect. 3.3 to Sect. 3.4

shown that the modules described in Sect. 3.2 and Sect. 3.3 extract most of the roads in rural areas, but often fail in urban and suburban areas. The combination of these modules with the completion module enhances the quality of the result. The main reason for the problems in urban areas are fragmented or missing line-like or parallel structures, which define the road axes and roadsides, respectively. Such structures, however, are the basic features for constructing road segments in high as well as in low resolution. In other words, for a successful and reliable road extraction in suburban and, especially, urban areas, the focus must be turned to another part of the model and the strategy must be adapted accordingly.

4 ROAD EXTRACTION IN URBAN AREAS

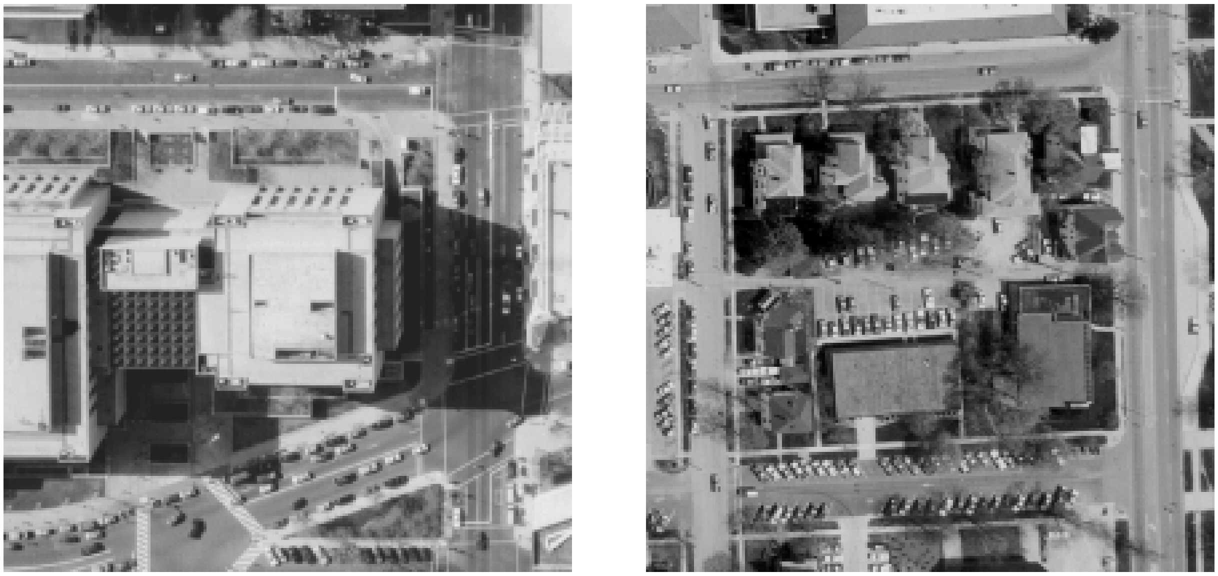
In this section, an approach for road extraction in urban areas is proposed. In Sect. 4.1 the most obvious differences that roads exhibit in rural and urban areas are exemplified. Based on these differences, an extraction strategy is proposed in Sect. 4.2, which exploits the presence of markings and vehicles on the roads. Finally, preliminary results are shown.

4.1 Appearance of Roads

The discussion of the results in Sect. 3.5 has indicated that the appearance of roads in urban areas is generally different from their appearance in rural terrain. Figure 22 shows two examples of urban roads. It is obvious that these images exhibit a more complex content than scenes showing rural areas since the number of different objects and their heterogeneity is much bigger. Generally, this fact implies that more details of the road and context model must be exploited for road extraction. In dense urban areas, for instance, some of the roads comprise several lanes that are linked by complex road crossings. What is more, because of the increased number of objects the complexity of their relations grows, too. In Fig. 22 a) for instance, some parts of the roads are occluded by vehicles, especially at the roadsides. Hence in this particular case, a road is mainly defined by groups of (parking) cars and not by parallel roadsides or by a homogeneous surface. A similar relation is the occurrence of shadows cast by high buildings. A road generally appears bright in open areas, but in the case of shadows two problems for the extraction arise: (1) the surface is significantly darker, and (2) strong gray value edges of the shadow boundaries may cross the road in almost any direction disturbing the usually homogeneous reflectance (see Fig. 22 a).

Figure 22 b) shows a different kind of problem: The roof of the rectangular building in the center of the image could be incorrectly identified as a parking lot because its shape and reflectance properties match those of a road-like object almost perfectly. Only the combination with height data as given by a DSM (Digital Surface Model) or, as in this case, implicitly given by a corresponding shadow region provides enough information for avoiding this misdetection.

What follows is, that on one hand those features of a road should be selected on which the influence of the above mentioned phenomena is minimal. On the other hand, it is very important to consider the context objects, in particular different kinds of vehicles.



(a) (b)

Figure 22: Examples of roads in urban areas

Besides an (at least partially) homogeneous surface and more or less densely arranged vehicles, one obvious feature of roads in urban areas are road markings. To make use of them, we model roads and complex junctions as a combination of several lanes consisting of one or more lane segments. Dashed or solid linear markings define the border of a lane segment. The interior of a lane segment should either exhibit the typical homogeneous reflectance of the pavement, or a vehicle that occludes the pavement has to be detected. The influence of high objects is considered twice: first, roads are allowed to be partly dark because they might lie in shadowed regions, and second, they cannot lie on locally high regions like buildings or dense vegetation. Occlusions, however, are not modeled yet.

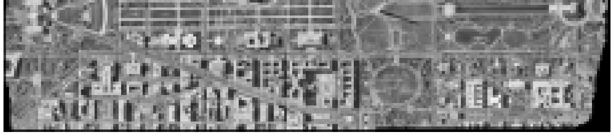
4.2 Road Extraction Based on Markings

From these components of the model, a strategy for road extraction based on markings is derived. Preconditions for a successful extraction are, of course, that markings (1) must be painted on the roads, and (2) they must be detectable in the image. Condition (1) is in fact fulfilled for many roads, especially for the larger ones in built-up areas. Condition (2) depends on the circumstances when taking the image, i.e., especially the resolution. Furthermore, it also depends on objects that might occlude road markings, e.g., large cars or trucks. Fortunately, as can be seen from Fig. 22, if the viewing angle is not too oblique, lanes are generally wide enough so that markings are visible even if cars are next to them.

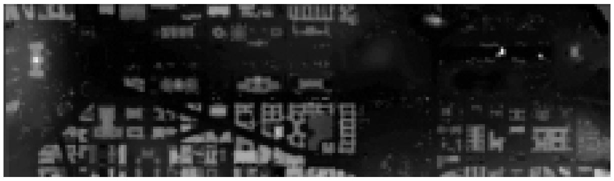
In our approach the extraction starts with the segmentation of areas of interest based on height information as given, e.g., by a DSM. Then, faint bright lines are extracted and iteratively connected to groups of markings that represent the lane sides. On both sides of every marking group a lane segment is hypothesized. Lane segments are verified by different criteria using geometric, radiometric, and context knowledge. After grouping the lane segments into lanes, the global connectivity of the lanes is checked and road junctions are constructed. In the following, the individual steps are described more in detail, and preliminary results are given.

4.2.1 Preprocessing In a first step, areas of interest are segmented using the context of roads: most buildings are higher than the road surface. Therefore, the parts that correspond to locally high regions in a DSM are removed from the image. In this example, the imagery has been down-sampled from approximately 0.25 m to 3 m. The segmentation procedure compares a smoothed version of the DSM with the original DSM and removes regions where the height difference between both DSM versions exceeds a threshold. Both parameters, the size of the smoothing mask and the minimum height difference, can be derived from the expected size and height of the buildings. Figure 23 shows the down-sampled image, the DSM image, and the segmented image.

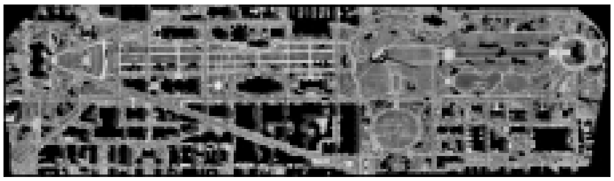
The segmentation results are then transformed to the original image resolution, and subsequently the image is partitioned into small patches (Fig. 24a). Note that, based on a DSM, a variety of segmentation techniques could be used in order to limit the search space, e.g., gray value morphology. Furthermore, a combination



(a) Original image in reduced resolution (3m)

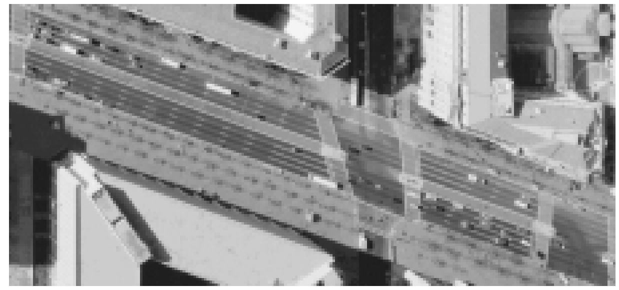


(b) DSM in reduced resolution (3m)

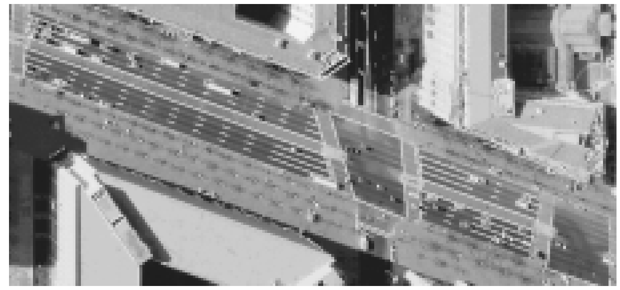


(c) Masked image

Figure 23: Segmentation of areas of interest

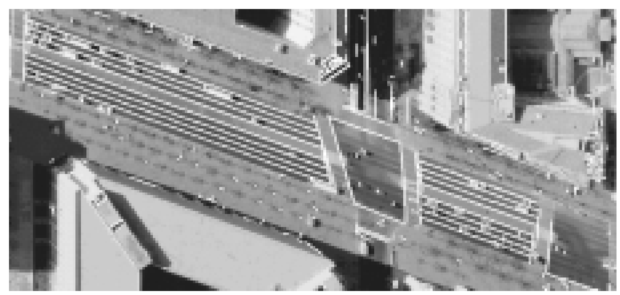


(a) Image patch

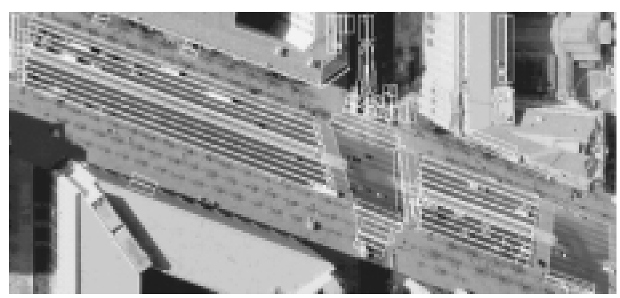


(b) Extracted lines

Figure 24: Original image patch and extracted lines



(a) Grouped lines



(b) Hypothesized lanes

Figure 25: Grouped lines and hypothesized lanes

with other road extraction approaches with which approximate roads positions can be derived easily is possible at this stage.

4.2.2 Line Extraction and Grouping During the next step, thin lines are extracted and grouped. Fig. 24b) shows the extracted lines, which are obtained using the approach of (Steger, 1998). Thereafter, lines are grouped according to the perceptual principles absolute and relative proximity and good continuation. Basically, the algorithm works in a very similar way like the one outlined in Sect. 3.3. Only the selection of seeds has been changed: from the lines and possible connection hypotheses, a weighted graph is constructed. In contrast to Sect. 3.3, the optimal path between every pair of vertices with degree 1 is calculated. By doing so, all possible groups of lines that show rather good continuation are detected. Thereafter, all paths are combined by means of deleting identical parts of different paths and splitting paths at intersections. The resulting set of unique and topologically consistent paths serves as input for the next iteration. A new graph with new connection hypotheses is constructed and the path calculation is carried out again. This procedure is repeated until no new connections are found. Figure 25a) visualizes the achieved result which represents the finally extracted groups of markings.

4.2.3 Generation of Hypotheses for Lanes The strategy for generating hypotheses for lanes is intentionally designed to be very liberal because of the following two reasons: (1) Markings usually appear as very faint lines. Thus, the line extraction may miss some of them. Such a failure, however, can be often compensated by the iterative grouping procedure described above. (2) Due to occluding objects like big trucks or trees, markings are more reliable to extract in the center of a road than on its sides. Therefore, two lane segments are hypothesized, one on each side of a detected group of markings.

In order to construct lane segments from the markings, general knowledge about the geometry of lanes is used. Lanes have some lower bounds for length and curvature radius. Therefore, after polygon approximation and splitting the groups of markings at sharp bends, short polygon segments are deleted. Additionally, lanes have in general a certain constant width. Hence, lane segments are constructed as rectangular regions on each side of a group of markings (see Fig. 25b).

4.2.4 Verification of Hypotheses Since the lane segments are hypothesized in a liberal manner, a sophisticated verification is needed in order to discriminate good from bad hypotheses. To this end, not only the geometric and radiometric properties of lane segments are considered, but also knowledge about their context is included. Here, the following criteria are used to collect evidence for the presence of a lane segment:

- Long, parallel groups of markings give high evidence for the presence of a road. Therefore, lanes that have markings on each side are extracted first (see Fig. 26).

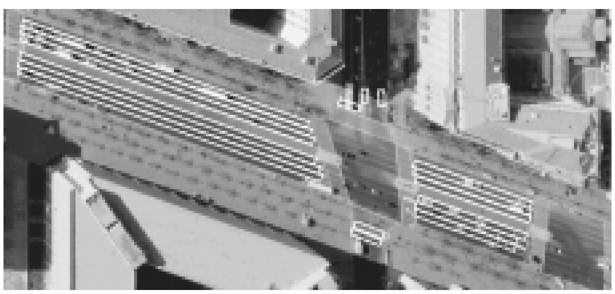


Figure 26: Parallel lane segments

- Additional markings at the margin of a lane segment are searched for by using lower thresholds than in the previous steps. A lane segment is rated depending on the percentage of dashed or solid markings (see Fig. 27).

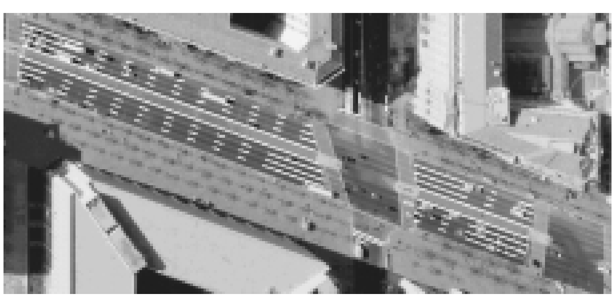


Figure 27: Additionally extracted markings

- As mentioned above, there is a low probability that markings can be detected at the sides of urban roads. However, in some cases small pieces of markings, curb-stones, and other parallel structures that can support a hypothesis might be found. Figure 28 shows all parallel edge and line structures — possibly highly fragmented — which could be extracted on this particular side of a hypothesized lane segment on which no parallel group of markings was detected.

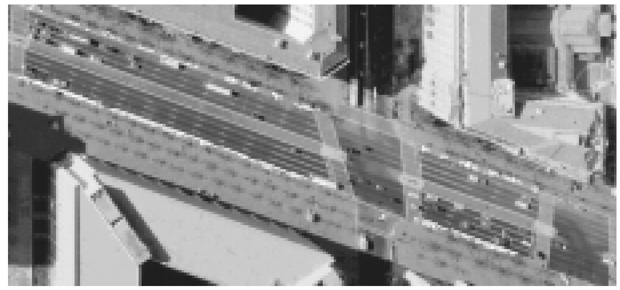
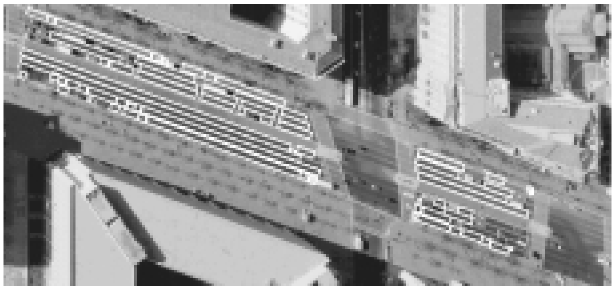


Figure 28: Extracted line and edge support

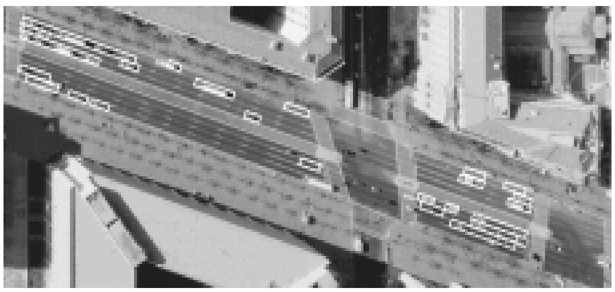
- The surface of a lane segment should be homogeneous in the direction of a lane. In regions where this criterion is not fulfilled it is a reasonable assumption that a car occludes the surface. Figures 29a) and 29b) visualize the extracted homogeneous regions and their complementary regions, respectively. We currently work on a car detection scheme which will use these regions as cues about a car's position and orientation in order to verify (or falsify) such occlusion hypotheses.

- Finally, orthogonal lines at the ends of lane segments are extracted (see Fig. 30). These lines are interpreted as cross walks or stop-lines for cars. This is on one hand useful to obtain information about the end of a lane. On the other hand, such an interpretation provides a strong cue for the presence of a junction or a T-intersection.

Note that no hypothesis is ultimately rejected at this stage of processing. A reliable decision if a lane segment belongs to a road or not is only possible when considering additional features, e.g., the connectivity of different lane segments and the global network topology. As mentioned, the results presented above should be regarded as intermediate steps of a more complex strategy. Work on this complex strategy is still in progress.



(a) Homogeneous regions inside lanes



(b) Car hypotheses

Figure 29: Radiometric analysis of lane hypotheses

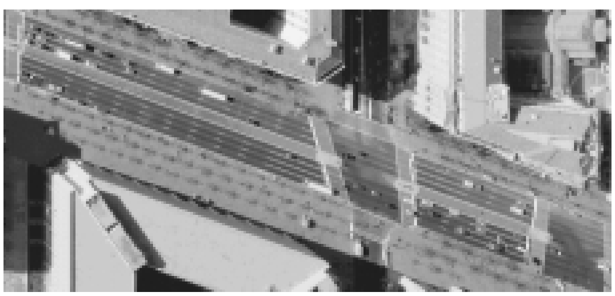


Figure 30: Extracted orthogonal lines at lane ends

5 CONCLUSIONS

The proposed approach for rural areas is suited for images with a resolution of 0.2 to 0.5 m. However, the results are not 100% reliable and complete. Hence, in operational use, a human operator is 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 grouping algorithms. By means of global grouping criteria, the knowledge about the topological properties of roads is incorporated, and we are able to overcome some deficiencies of the purely local grouping used in Sect. 3.2. The strengths of the module described in Sect. 3.2 are the integration of different scales and the use of context information. This is the main reason for the good *correctness* of the results. We showed that a noticeable improvement concerning the *connectivity* of the resulting road network is possible with an integration of global grouping criteria (Sect. 3.3). Experience has shown that the most critical point of Module II — when using it without Module I — is the selection of correct seed points for the path calculations. However, by introducing the relatively reliable results of Module I, a quite robust selection was feasible. Both modules could thus support each other, although they were originally developed independently. The main reason for the loss of geometric accuracy after global grouping is that some added road segments come only from line extraction or are not verified by image data at all. The completion of the road network (see Sect. 3.4) showed to be an adequate tool to add important portions of the road network. The benefit from this module is quantified by the reduction of the *detour factor*.

For road extraction in urban areas, markings are the most important features. DSM information has proven to be very useful to restrict the search space. Compared with the approach for rural areas, the extraction uses more knowledge about substructure of roads (markings, lanes) and relations between vehicles and lanes. The preliminary results of this approach for road extraction in urban areas are encouraging.

What is missing in our road extraction scheme is the link between urban and rural areas. The roads extracted in rural or urban areas could be used as starting points for road hypotheses in suburban areas. However, the problem in suburban areas is that partly the model for rural, but mostly the model for urban areas is valid. A consequence would be to employ a suitably extended road and context model and to employ a flexible extraction strategy.

ACKNOWLEDGMENTS

We are grateful for the vital contributions of Heinrich Ebner to our work on road extraction. He not only initiated it and gave many helpful comments, but also detected the weaknesses and always directed the work in the right way. Especially we want to thank him for the freedom he gave us and for the trust he had in all of us.

We thank Olivier Jamet (Institut Geographique National, Paris), Ram Nevatia, and Keith Price (both Institute for Robotics and Intelligent Systems at University of Southern California, Los Angeles) for useful discussions and kindly providing image data for testing our approach. In addition to the authors, Clemens Glock, Ivan Laptev, Ralph Multhammer, and Doris Reisenauer were involved in the presented work.

REFERENCES

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

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W. and Ebner, H., 1999. Automatic Road Extraction Based on Multi-Scale, Grouping, and Context. Photogrammetric Engineering & Remote Sensing 65(7), pp. 777–785.

Clément, V., Giraudon, G., Houzelle, S. and Sandakly, F., 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), pp. 779–791.

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

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

Hinz, S., Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H. and Radig, B., 1999. Road Extraction in Rural Areas and Urban Areas. In: Semantic Modeling for the Aquisition of Topographic Information from Images and Maps, pp. 133–153.

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

Knuth, D. E., 1994. The Stanford GraphBase: A Platform for Combinatorial Computing. Addison-Wesley Publishing Company, Reading, MA, USA.

Mayer, H., 1998. Automatische Objektextraktion aus digitalen Luftbildern. Deutsche Geodätische Kommission (C) 494, München.

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

Mayer, H., Baumgartner, A. and Steger, C., 1998a. Tutorial on Road Extraction from Aerial Imagery. In: CVonline: The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision. http://www.dai.ed.ac.uk/CVonline/applic.htm.

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

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

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

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

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

Tönjes, R., 1997. 3D Reconstruction of Objects from Aerial Images Using a GIS. In: International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 3-2W3, pp. 140–147.

Wiedemann, C., 1999. Completion of Automatically Extracted Road Networks Based on the Function of Roads. In: International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 3-2W5, pp. 209–215.

Wiedemann, C. and Hinz, S., 1999. Automatic Extraction and Evaluation of Road Networks from Satellite Imagery. In: International Archives of Photogrammetry and Remote Sensing, Vol. 32, Part 3-2W5, pp. 95–100.