

MULTI-SCALE ROAD EXTRACTION USING LOCAL AND GLOBAL GROUPING CRITERIA

Albert Baumgartner, Stefan Hinz

Chair for Photogrammetry and Remote Sensing
Technische Universität München, D-80290 Munich, Germany
E-mail: {albert}{hinz}@photo.verm.tu-muenchen.de
URL: <http://www.photo.verm.tu-muenchen.de>

KEY WORDS: Image Understanding, Road Extraction, Grouping.

ABSTRACT

In this paper we combine two approaches for road extraction. The first approach makes use of multiple scales to detect roads segments and employs local grouping criteria and also context information to extract the road network. This approach is suitable for aerial imagery with a resolution of 0.2–0.5 m. The second approach was designed to extract roads from satellite imagery and can be applied to resolutions of 2–5 m. It fuses lines extracted from different channels for road extraction and exploits especially the connectivity properties of roads, i.e., global network criteria for road extraction. By combining the two approaches we can reduce the effort for selecting appropriate parameters, because both help each other to get rid of some individual deficiencies. In addition, the evaluation of the extracted road network showed significant improvements compared to the results we get by applying each approach on its own.

1 INTRODUCTION

There is a big economic desire to automate the extraction of objects from aerial and satellite imagery, and there is a lot of research in this field, too. Nevertheless, fully automatic extraction of objects like buildings or roads is still an unsolved problem. At our institute two different approaches for fully automatic road extraction have been developed during the past years. The first approach makes use of multiple scales to detect roads segments and employs local grouping criteria and also context information to extract the whole road network (Baumgartner et al., 1999). The second approach focuses on the connectivity properties of roads and is able to make use of the information derived from multi-spectral satellite imagery (Wiedemann and Hinz, 1999). Whereas the first approach is restricted to gray scale imagery with a resolution of 0.2 to 0.5 m, in which roads appear as homogeneous regions, the second approach models roads as lines and is able to fuse lines extracted in multiple channels. Both approaches show individually good results – within a limited scope. In the work presented in this paper we show how both approaches can be combined and how they benefit from the strengths of each other and help to overcome their deficiencies.

Some of the basic ideas of our road extraction scheme are described in detail in earlier publications, e.g., in (Steger et al., 1995, Baumgartner et al., 1997, Steger et al., 1997, Mayer and Steger, 1998). Work related to our local approach for road extraction has been carried out, by (McKeown Jr. and Denlinger, 1988, Ruskoné et al., 1994, Airault et al., 1994). E.g., (Ruskoné, 1996) proposed a fully automatic approach for the extraction of road networks from digital aerial imagery: Hypotheses for connections between automatically detected seed points are checked based on geometrical constraints. The influence of neighboring objects on road extraction has been investigated in (Bordes et al., 1997). For our second approach the relevant previous works are (Fischler et al., 1981, Vasudevan et al., 1988)

Apart from the trend towards the integration of contextual information, there is a strong emphasis on defining evaluation criteria and developing methods to evaluate the results of automatic and semi-automatic approaches for road extraction, see e.g. (Heller et al., 1998, Heipke et al., 1998, Harvey, 1999).

The model which serves as the basis of our road extraction scheme is outlined in Section 2. In Section 3 we describe the individual approaches for road extraction which are combined in Section 4. The benefits of this combination are documented by an external evaluation of the results (Section 5). In Section 6 we draw some conclusions.

2 ROAD MODEL

We base our road extraction scheme on the road model displayed in Fig. 1. The road model comprises multiple scales and describes the road network in three different levels. The *real world* level contains the road objects (e.g., road network, junction) and their relations. At the *geometry and material* level, the 3D-shape and the material of roads are represented. In

contrast to the *image* level, the representation at the *geometry and material* level is independent from sensor characteristics. Depending on the scale, i.e., the image resolution, roads are either modeled as flat homogeneous regions, or as lines. The solid dark lines establish the connections between the concepts at the different levels.

The use of different scale is motivated by the fact, that different characteristics of roads can be best detected at different scales. At fine scale, i.e., in high resolution images, a better geometric accuracy can be achieved since the road sides can be detected very precisely. Substructures on the road, e.g., markings, can give additional hints for road extraction. At coarse scale, i.e., at resolutions where roads are only a few pixels wide, the network characteristics of roads are more clearly visible, and small objects like single vehicles or trees do not influence the extraction as heavily as they do in fine scale.

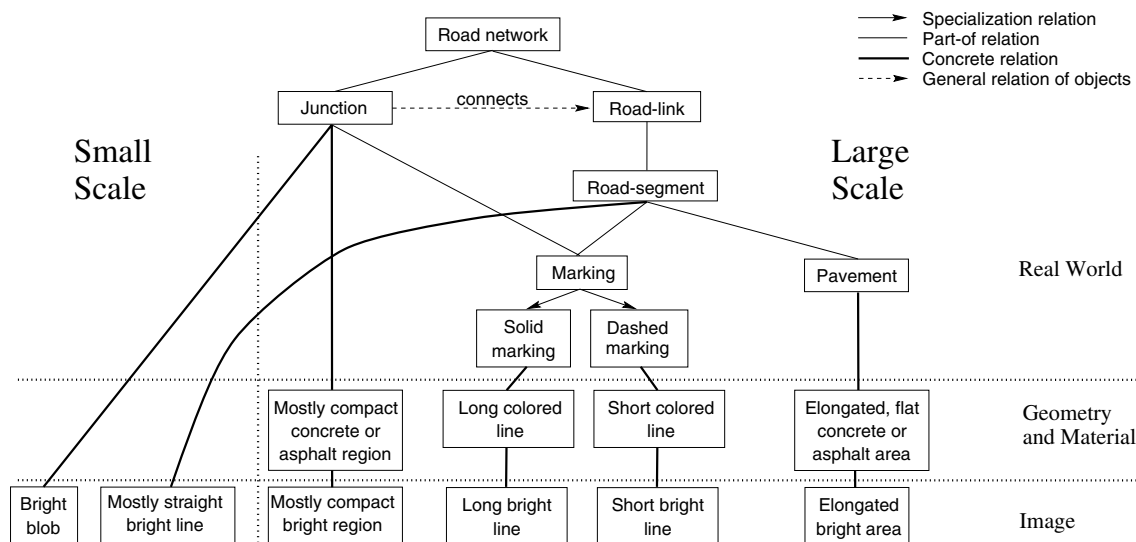


Figure 1: Road model

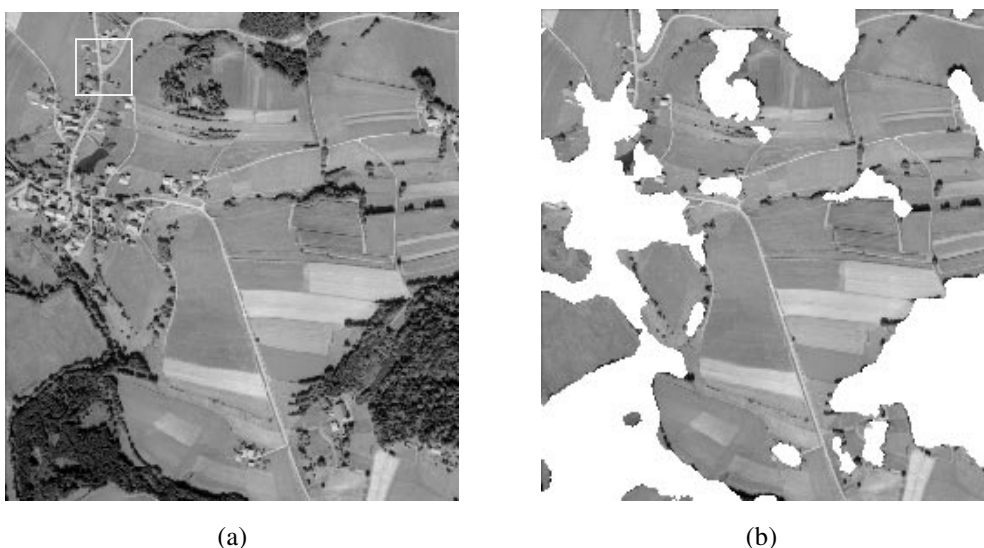


Figure 2: (a) Image (b) Segmentation of *open_rural* context

In addition to the road model in Fig. 1, which comprises knowledge about geometric, radiometric, and topological properties of roads, our model contains also relations between roads and other objects, e.g., buildings, trees, and vehicles. This type of knowledge is modeled by the context of the roads. The context model is split into *local* and *global* context. Whereas, the local context describes relations between individual objects of different types, the global context segments the image into three regions, in which the appearance of roads in imagery is completely different: *urban*, *forest*, and *open_rural* context regions. A texture based segmentation of the *open_rural* context is shown in Fig. 2. The segmentation of the different context regions provides a priori information about the typical problems which might occur during the road extraction. Therefore, we can use the global context to guide the extraction and start at places where the extraction is supposed to be easy and reliable.

3 ROAD EXTRACTION MODULES

In this section we describe the main characteristics, the strategy, the advantages, and the deficiencies of our two basic approaches for road extraction. We will combine these two approaches and use them as complementary modules in a new road extraction scheme. In the remainder we will refer to them as “module I” and “module II”. Both modules use local and global information for road extraction. However, module I mainly focuses on local criteria, whereas module II exploits especially global criteria, i.e., the network characteristics of roads.

3.1 Module I: Local Grouping

On the local level, we use lines and edges as image features to construct road segments. According to the road model, we use apart from the original image also a version of the image with a reduced resolution. The lines extracted in the reduced-resolution image (about 2 m) are used to select edges extracted from the original resolution that are candidates for road sides. In order to be selected as road sides edges must fulfill several criteria: The distance between pairs of edges must be within a certain range. The edges have to be almost parallel, i.e., there is an overlap and the direction difference between the edges is small. The area enclosed by a pair of parallel edges should be quite homogeneous in the direction of the road. In addition, for each pair of road side candidates a corresponding line has to exist in the reduced resolution.

From these road sides, initial hypotheses for road segments are constructed (Fig. 3). The road segments consist of quadrilaterals which are generated from parallel road side 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 which are used as input for the extraction of the other parts of the road network.

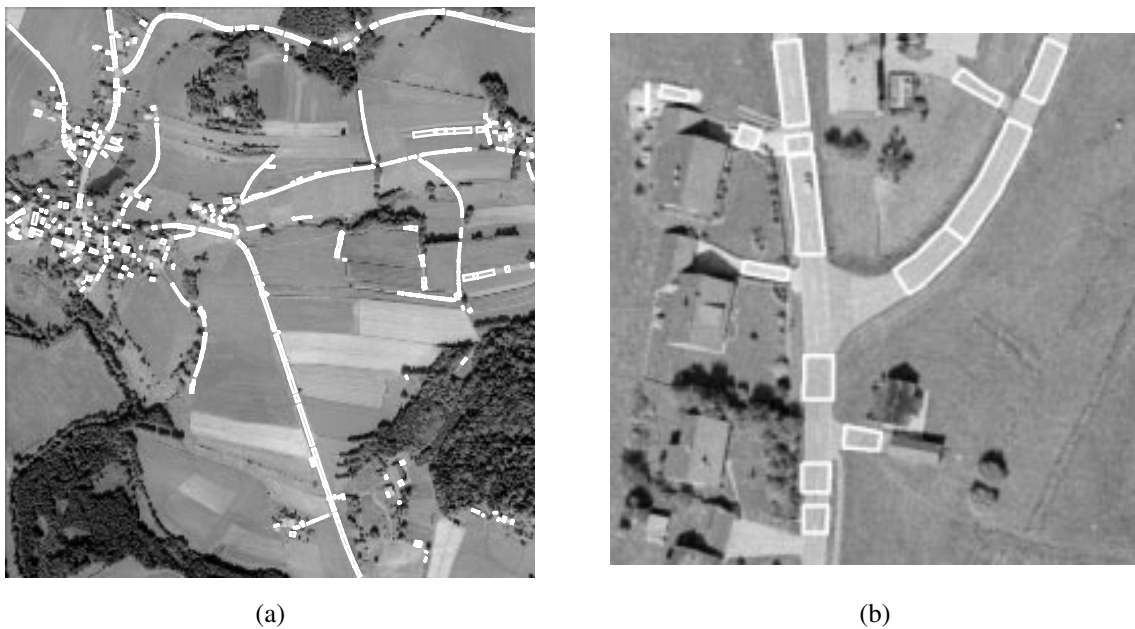


Figure 3: (a) Initial hypotheses for road segments (b) Detail

The fusion of lines from low resolution and edges from high resolution has proven to be very useful in order to get more reliable results. For easy scenes these steps are often sufficient to come up with correct hypotheses for road segments which can easily be linked into longer segments. This advantage of the combination of line and edge extraction is also confirmed by the results of (Trinder and Wang, 1998) who use a quite similar approach to fuse low and high resolution imagery for road extraction.

However, the limits of this initial detection of hypotheses for road segments become clear when the approach is applied to urban or suburban areas. Inside the village the number of correct initial hypotheses for road segments decreases tremendously, in scenes with many buildings most of the hypotheses are displaced due to shadows and occlusions, or even completely wrong (cf. Fig. 3).

Putting the correct hypotheses together and eliminating false ones is the task to be solved during the next steps. In module I geometric properties of neighbored road segments are used to establish hypotheses for connections between these segments. The connection hypotheses are verified by analyzing the gap between the segments based on radiometric and geometric criteria (e.g., thresholds on mean gray value, difference in width and direction). Applying so-called

“ribbon snakes” showed to be a very useful method to find a path between two road segments and to verify the connection hypothesis. If the verification based on these criteria fails, an attempt is made to explain the gap between the neighboring segments by information about the local context, e.g., due to a shadow cast by a building. These grouping steps are applied iteratively, and the thresholds on the distance and the direction difference are relaxed step by step. Simultaneously with the relaxation of the thresholds short segments are removed. This elimination step is necessary because otherwise, due to the relaxation of the grouping thresholds and due to the limitations of the verification step, a lot of erroneous connection hypotheses would be accepted and would corrupt the further steps of the road extraction. However, correct initial hypotheses are also removed.

After the generation of hypotheses for connections and their verification, the road network is constructed (Fig. 4). Based on geometric assumptions hypotheses for junctions are generated and verified. 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, because the extraction is primarily based on local information and is reliable only in rural areas. In summary, module I uses only local information to establish connection hypotheses and to verify them. The network characteristics of roads are not optimally exploited. Therefore, its most important feature compared to module II is the aspect of local grouping. Apart from radiometric parameters which are directly linked with the quality of the image, the threshold for the elimination of the unconnected short segments is the most sensitive parameter with respect to the quality of the results obtained with this module.

3.2 Module II: Global Grouping

Module II is primarily based on the knowledge that roads have the function to connect different “important places,” even if they are far away from each other. 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 and rural terrain) the search can be restricted to the *best* connection between two places.

This module starts the extraction of the road network with the extraction of lines, calculates attributes for these lines and assesses the probability of the extracted lines to be a part of a road network. Based on local line attributes (e.g., straightness, length) which are then compared to the knowledge about shape and reflectance properties described by the road model, each line gets a quality measure. The endpoints of all lines are used as vertices of a graph. The lines which connect the endpoints are edges in this graph and for all pairs of vertices which are not connected by a line a quality measure for the shortest connection is calculated. The quality measure of the “gap-edges” in this graph depends on purely geometric considerations. The quality measures of lines and gaps are transformed by linear fuzzy functions into values ranging from 0 to 1. An overall fuzzy value of 1 means that the edge perfectly meets the properties derived from the road model.

Once the weighted graph is constructed, the next step is to select the “important places”. Since this approach knows nothing about additional objects of the real world, e.g., buildings, industrial areas, or other sites of interest, we define “important places” as lines that represent portions of the road network with high probability. Hence, instead of connecting true “important places”, we try to connect pairs of high quality line segments (seed pairs). Additionally, the



Figure 4: Results of local grouping.



Figure 5: Results of global grouping.

pairs have to be far away from each other, in order to emphasize the global network characteristics. The last step is to calculate the best path between each seed pair. The sum of all best paths is supposed to correspond to the road network.

As this approach was designed for road extraction from multi-spectral satellite imagery, it is able to fuse lines extracted from different channels. It can also be applied on a gray scale imagery, however, its capability can not be exploited very much if only one channel is used. Results for this module, applied to our example image are shown in Fig. 5.

Although the criteria for the selection of these “important places” can be derived from semantically meaningful and reasonable parameters, the selection of these points is one of the most sensitive steps within module II.

For the combination of the two modules the ability of module II to fuse line data from different images or even from different sensors is essential, because the axes of road segments delivered by module I can easily be used as an additional input “channel”. Another useful feature of module II for our purposes is that different weights can be assigned to the lines from different channels.

4 COMBINATION OF LOCAL AND GLOBAL MODULE

In this section we show two examples for the combination of module I and module II. In Sect. 4.1 module I and module II are applied sequentially. This gives us a rough idea about the use of knowledge about the global connectivity of the road network, which was up to now not exploited in module I. The integration of module I and module II described in Sect. 4.2 tries to make optimal use of globally best paths which can be found by module II and keeps the good geometric accuracy of module I.

4.1 Sequential combination

The easiest way to combine the two modules is to combine them sequentially. For this combination there is no need to change anything of the internal structures of any module. The output of module I (cf. Fig. 4) is used as additional input in module II. The axes of the extracted road network are fused with the lines extracted in the images at a reduced resolution of about 2 m. By setting the weight for the axes resulting from module I much higher than the maximum weight of the extracted lines, we ensure, that no axes will be lost in the resulting network delivered by module II, i.e., we assume that the results of module I are correct. Therefore, in this case the result of the combination consists of the axes shown in Fig. 4 and some additional lines which connect the fragmented result that was delivered by module I. Comparing the combination result (Fig. 6) to the previous results, then the most significant difference to Fig. 4 is the fact that almost all parts of the extracted network are connected, and that is possible for module II to find a path through the village at the left side of the image. Compared to the stand alone result of module II (Fig. 5) there are more side-roads and blind alleys connected to the resulting network now, and the use of context information allowed to bridge the shadow in the upper part.

4.2 Integrated Combination

In this section we describe the integration of module I and module II. More exactly, we integrated module II in the extraction process of module I. As mentioned in Sect. 3.1 the elimination of wrong initial hypotheses for road segments has a strong influence on the resulting road networking. Basically, in the elimination steps module I tends to eliminate not only erroneous ones, but correct initial hypotheses, too. Thus, it achieves good correctness at the cost of completeness. However, in areas where the initial hypotheses for road segments are very fragmented and short, e.g., in urban or built-up areas, the elimination of a few correct hypotheses mostly is equivalent to significantly worsening the chances to extract the road network in this area at all. At this point module II is employed to tell module I which of the short segments that are candidates for elimination should be kept due to their importance for the connectivity of the road network.

Before initial hypotheses are eliminated by module I it sends all its road segments to module II and labels them as “good” or “bad” candidates. Module II fuses all segments with the lines extracted at reduced resolution and starts its road network extraction, using only the “good” segments as possible seed points. I.e., the task to select appropriate seed points for module II is solved by module I.



Figure 6: Results of sequential combination of local and global grouping.

Module II then returns all parts of its road network to module I which are not covered by the “good” segments. From this result module I gets information about the good segments which might be linked according to global grouping criteria, and it can easily decide which of the “bad” segments should be labeled as important for the connectivity of the road network. This procedure is invoked before every elimination step of module I.

Because the local grouping of module I bridges gaps between the road segments, relaxes its grouping thresholds, and can rate other road segments as “good” or “bad” ones, also the seed pairs for module II can change. Therefore, it is necessary that module I asks module II before every elimination step again.

In Fig. 7 this process is exemplified for an early elimination step: The road segments displayed in white are considered to be “good”. Candidates for elimination are displayed in black. The paths returned by module II are dotted white lines. In this example the overlap with a global path prevents some short road segments with a bad rating from being eliminated.

Considering, that module I is quite good at detecting correct parts of the road network, and that module II is able to provide good hypotheses for the global connectivity, both modules benefit from this combination.

After the last elimination step all paths which were provided by module II are added to the network extracted by module I. For those paths which can not be verified using the methods of module I, at least the position of the path is adapted to the features in the high resolution image. This optimization is done, because the connection hypotheses from module II are known to be quite reliable, if the seed pairs are well selected, and only the geometry of the paths is known to be less accurate. The final result of this integration of local and global grouping modules is given in Fig. 8.

5 EVALUATION

In the previous section we made some qualitative statements about the results. Now, to get more independent statements we apply a quantitative evaluation. For this evaluation we compare the extracted road networks with a reference network. The internal quality measures of module I and module II about parts of the extracted road network are not taken into account. In the following we give a short description of the external evaluation procedure we use. For more details see (Heipke et al., 1998, Wiedemann, 1998). The reference data for this evaluation is shown in Fig. 9. It was manually plotted at a resolution of 0.25 m. The width of the roads in the reference network ranges from about 3 m to 8 m.

The evaluation scheme allows for statements about the completeness and the correctness of the extracted roads by matching the extracted data to the reference data using the so-called “buffer method”. For the correct parts of the extracted roads it provides further an RMS-value of the position of the extracted axes with respect to manually plotted reference. The completeness indicates how much is missing in the network, whereas the correctness is related to the probability of an extracted linear piece to be indeed a road.

Completeness is defined as the percentage of the reference data that is explained by the extracted data, i.e., the percentage of the reference data which lies within the buffer around the extracted data.

The correctness represents the percentage of correctly extracted road data, i.e., the percentage of the extracted data that lies within the buffer around the reference network.



Figure 7: Detail: Global link hypotheses (dotted, white) and road segments (start segments for global grouping are white, others, black).



Figure 8: Results of integrated combination of local and global module

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.

The evaluation figures given in Tab. 1 show that the sequential combination of module I and module II, i.e., of local and global grouping criteria, increases the completeness, and in the integrated combination also the good correctness and RMS-values of the results of the local level can be kept. The fact that the integration of global grouping criteria enforces the extraction of a connected road network is not expressed in the figures.

It should be noted that the networks resulting from the combined approaches are inhomogeneous with respect to the geometric accuracy since most of the resulting network originates from the hypotheses for road axes from module I and parts of the network originate from line extraction at 2 m resolution.

On principle, the evaluation results depend on the buffer width which is chosen for the matching between reference roads and extracted roads. The larger we set the buffer, the more likely extracted roads will be matched with roads in the reference. I.e., enlarging the buffer can raise the correctness and completeness figures at the cost of worse RMS-values. For the results given in Tab. 1 the buffer width was set to 3 m, i.e., about half of the road width of an average road in the given image. In our case the influence of the buffer width on the evaluation figures showed to be marginal. Only for very narrow buffers of less than 1 m the figures change significantly.

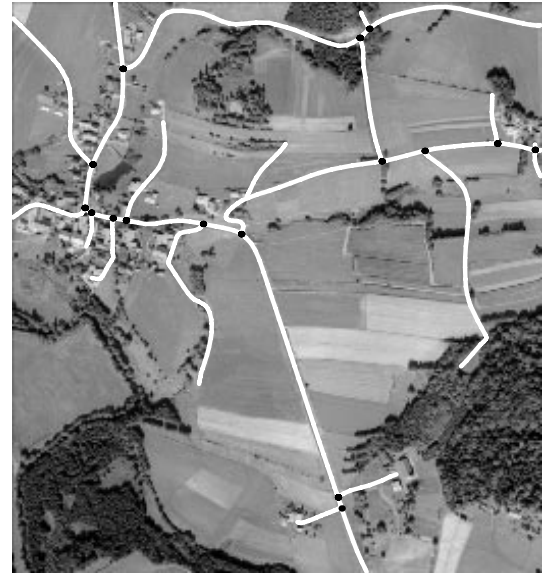


Figure 9: Manually plotted reference network

	Module I: Local (Sect. 3.1)	Module II: Global (Sect. 3.2)	I+II: Sequential (Sect. 4.1)	I+II: Integrated (Sect. 4.2)
Completeness [%]	81	71	86	87
Correctness [%]	94	88	90	93
RMS [m]	0.42	0.83	0.52	0.45

Table 1: Evaluation results for different combinations of local and global module

6 CONCLUSIONS AND OUTLOOK

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 purely local grouping. We showed that a noticeable improvement especially concerning the connectivity of the resulting road network is possible with an integration of global grouping criteria. One point which still should be improved is the weak model for junctions. By now we can handle quite simple junctions only.

In contrast to semi-automatic approaches, where the human operator can decide about the use of automation tools in each case separately and, what is even more important, he can accept or reject the results immediately, fully automatic systems must be able to decide on their own, where to start their search for specific objects and what to do next. To be an useful semi-automatic tool it is more important to optimize the interaction with the operator, than having a sophisticated self diagnosis algorithms, which are necessary for fully automatic systems. Fully automatic systems can base their decision only on knowledge about the object, which is available in the system or on the information which they can derive from the provided data. Basically, many parts the knowledge of the system can be reduced to a set of parameters, which might be hard-coded or selected by the user. The most important requirement for a fully automatic system is to be able to cope with a wide variety of data sets without major parameter modifications. The number of user selectable parameters should be reduced to a minimum and the influence of these parameters on the result of the system should be easily predictable and self-evident also for an unexperienced user. In our case the combination of the two complementary approaches for road extraction seems to be one important step towards this goal. The objection that by combining the two modules we perhaps replaced some parameters in one module with at least as many needed in the other module is not really true. Here, we were able to reduce the influence of some quite sensitive parameters, and to improve the quality of the results, too. Even if the overall number of parameters increases, for us it seems to be more important that we could eliminate some crucial parameters. Something which is not expressed by external evaluation criteria is the lower sensitivity to predefined parameter settings of the integrated combination compared to the local grouping as stand alone system. Of course these findings have to be validated on different data sets.

However, also with this integration of local and global grouping methods the system is still restricted to less complex scenes. E.g., by now we do not use enough knowledge about roads and their relations to other objects to be able to extract roads in urban areas. In our opinion, in urban areas the influence of vehicles and road markings has to be considered, and especially in downtown areas we have to integrate good height data to cope with the large number of occlusions and shadows. The concept and first results for road extraction in urban scenes are described in (Hinz and Baumgartner, 2000).

ACKNOWLEDGMENTS

The work presented in this paper was carried out in cooperation of the Chair of Photogrammetry and Remote Sensing and the Department of Computer Science (Image Analysis Group) at TU München during the past six years. Funding by the Deutsche Forschungsgemeinschaft (DFG) under grant no. Eb 74/8-3 is gratefully acknowledged.

REFERENCES

- Airault, S., Ruskoné, R. and Jamet, O., 1994. Road detection from aerial images: a cooperation between local and global methods. In: *Image and Signal Processing for Remote Sensing*, Proc. SPIE 2315, pp. 508–518.
- 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.
- Bordes, G., Giraudon, G. and Jamet, O., 1997. Road Modeling Based on a Cartographic Database for Aerial Image Interpretation. In: *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*, Birkhäuser Verlag Basel, pp. 123–139.
- Fischler, M. A., Tenenbaum, J. M. and Wolf, H. C., 1981. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Computer Graphics and Image Processing* 15, pp. 201–223.
- Harvey, W. A., 1999. Performance Evaluation for Road Extraction. In: *Workshop on 3D Geospatial Data Production: Meeting Application Requirements*, Paris, pp. 175–184.
- Heipke, C., Mayer, H., Wiedemann, C. and Jamet, O., 1998. External Evaluation of Automatically Extracted Roads. *Photogrammetrie – Fernerkundung – Geoinformation* (2/98), pp. 81–94.
- Heller, A. J., Fischler, M. A., Bolles, R. C., Connolly, C. I., Wilson, R. and Pearson, J. I., 1998. An Integrated Feasibility Demonstration for Automatic Population of Geospatial Databases. In: *Image Understanding Workshop*, Monterey, CA, USA.
- Hinz, S. and Baumgartner, A., 2000. Road Extraction in Urban Areas Supported by Context Objects. In: *International Society of Photogrammetry and Remote Sensing*, Vol. 33, Part B, *International Archives of Photogrammetry and Remote Sensing*, this proceedings.
- 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.
- McKeown Jr., D. M. and Denlinger, J. L., 1988. Cooperative Methods for Road Tracking in Aerial Imagery. In: *Computer Vision and Pattern Recognition*, pp. 662–672.
- Ruskoné, R., 1996. Road Network Automatic Extraction by Local Context Interpretation: application to the production of cartographic data. PhD.-thesis, Université Marne-La-Valleé.
- Ruskoné, R., Airault, S. and Jamet, O., 1994. A Road Extraction System Using the Connectivity Properties of the Network. *Zeitschrift für Photogrammetrie und Fernerkundung* (5/94), pp. 174–180.
- 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.
- Trinder, J. C. and Wang, Y., 1998. Knowledge-Based Road Interpretation in Aerial Images. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 4, pp. 635–640.
- Vasudevan, S., Cannon, R. L. and Bezdek, J. C., 1988. Heuristics for Intermediate Level Road Finding Algorithms. *Computer Vision, Graphics, and Image Processing* (44), pp. 175–190.
- Wiedemann, C., 1998. Empirical Evaluation of Automatically Extracted Road Axes. <http://www.photo.verm.tu-muenchen.de/staff/wied/evaluation/evaluationOfRoads.html>
- 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.