PFG 2010 / 5, 0327 – 0338
Stuttgart, Oktober 2010

Article

Automatic Road-Tracking in Airborne Image Sequences

Mathias Koller, Matthias Butenuth, München & Markus Gerke, Enschede

Keywords: Road Tracking, Image Sequences, Image Overlap, Redundancy

Summary: This paper describes an automatic road-tracking approach for rural areas using airborne image sequences. Due to the relatively small image size, small-frame camera systems are ideal platforms for the development of automatic image analysis approaches which make use of redundant image information available through the obligatory image overlap. In the proposed approach an initial road network extracted in a first image or manually digitized by an operator is automatically continued in overlapping adjacent images. First, a supervised classification is performed: The existing road network is used as the training region in the overlapping area in order to optimally account for radiometric and illumination properties of the individual images. Second, the segments obtained in the first step are combined with extracted edges in order to derive an optimal road-tracking. The combination of areal and linear features also enables to extract road crossings. Third, a post-processing step closes gaps between extracted sections to build a complete road network. Results demonstrate that a reliable and accurate tracking of main roads including junctions and field-paths is possible with the proposed approach given a clear color separation of path and background. An open issue concerns the closing of large gaps which needs further attention in future research.

Zusammenfassung: *Automatische Verfolgung von Straßen in Luftbildsequenzen.* In diesem Beitrag wird ein Ansatz zur automatischen Straßenverfolgung in Luftbildsequenzen für ländliche Gebiete beschrieben. Trotz ihrer relativ kleinen Bildgröße sind Small-Frame Kamera-Systeme ideale Plattformen für die Entwicklung von Ansätzen zur automatischen Bildanalyse, da durch die Bildüberlappung redundante Bildinformation genutzt werden kann. Der vorgestellte Ansatz setzt ein Staßennetzwerk, entweder in einem ersten Bild extrahiert oder manuell von einem Operator digitalisiert, automatisch in überlappenden benachbarten Bildern fort. Zuerst wird eine überwachte Klassifizierung durchgeführt: Das bestehende Straßennetzwerk wird als Trainingsregion im überlappenden Bereich benutzt, um die Eigenschaften der individuellen Bilder bezüglich Radiometrie und Beleuchtung optimal zu berücksichtigen. Zweitens werden die erhaltenen Segmente mit extrahierten Kanten kombiniert, um eine optimale Straßenverfolgung herzuleiten. Drittens schließt ein Nachbearbeitungsschritt Lücken zwischen extrahierten Sektionen, um ein komplettes Straßennetzwerk zu erzeugen. Ergebnisse zeigen, dass mit diesem Ansatz eine zuverlässige und genaue Verfolgung von Hauptstraßen inklusive Kreuzungen und Feldwegen möglich ist, wobei vorausgesetzt wird, dass sich diese Wege farblich klar vom Hintergrund abheben. Ein noch nicht zufriedenstellend gelöstes Problem ist das automatische Schließen großer Lücken im extrahierten Staßennetzwerk.

1 Introduction and Related Work

The automatic extraction of roads from imagery is still an ongoing research topic. The EuroSDR test on road extraction (Mayer et al. 2006) reveals that quite a number of experimental automatic road extraction systems are nowadays mature enough, at least to extract or verify road information in rural areas, obtaining completeness and correctness rates of about 80 % or more. On the other hand, the market of image acquisition systems shows an interesting trend: More and more small- and medium-format camera systems are operational which are cheaper compared to traditional large format camera systems. This trend offers the opportunity to develop and test new methods for object extraction, which is moti-

DOI: 10.1127/1432-8364/2010/0059 1432-8364/10/0059 $ 3.00

vated in the following. Known approaches for image-based object extraction make use of orthoimage mosaics since these are standard products of most mapping agencies. The advantage of those images is the large coverage of an area and a certain guaranteed level of image accuracy by the producer. However, one disadvantage is the huge unused portion of original image information. The original images captured in airborne campaigns have at least stereo overlap, but for the orthoimage production only one image is used and, additionally, the color is balanced over the whole mosaic, biasing methods which rely on a global image statistic. If the original images would remain available, object extraction methods can directly exploit the redundant information.

The latest camera systems like the 3K allow real-time processing of georeferenced orthoimages from small-format cameras (Butenuth et al. 2009, Thomas et al. 2009). Since the orthoimage processing is done for every image separately, this kind of input data is a very interesting source for new road extraction approaches exploiting explicitly the *image overlap*. To extract roads automatically in such imagery might be of interest not only for traditional applications like the update of existing GIS data, but additionally open new application fields like the (near) real-time mapping in disaster cases. In this paper, a new approach to road extraction is presented which makes explicitly use of redundant image information in overlapping areas. The core of the strategy is a combination of image classification and edge extraction, finally enhanced by bridging gaps in the road network.

In the following, a brief overview on existing approaches for road extraction is given. Although a clear categorization is difficult, many existing road extraction approaches can be divided into fully automatic extraction and semi-automatic road-tracking. *Fully automatic approaches* use previous modeled knowledge for an extraction or existing GIS-data about the road network and complete it. The approach of (Hinz et al. 2000) exploits low-level extractions in different image resolutions and global/local context. In (Mena & Malpica 2005) a method based on segmentation is developed which uses knowledge from existing GIS data followed by a vectorization. The skeleton produced by the vectorization used in this approach does not detect a segmented road-area precisely at its borders. The approach of (Rosenbaum et al. 2008) is based on a fusion of existing road-knowledge with edge extraction. Prior information from a road database is used to provide a buffer region in the image where an edge extraction approach is accomplished. In (Bacher & Mayer 2005) supervised image classification is used to evaluate extracted roads, originating from a line-based road extraction approach. The classification result is mainly used to exclude wrong positive road extractions, and not for the detection and delineation of roads as such.

Road-tracking works semi-automatically with starting input and user interaction at critical points. Many approaches use areal multispectral templates of small road-sections combined with direction information to do a locally connected search. An operator determines a point of special interest in a road-network to indicate the direction and to extract the parts of the network which are locally connected with the start point. One typical example is (Vosselman & de Knecht 1995), an approach using colorprofile-matching and a Kalman-filter. A large variety of methods for road extraction in single (ortho) images exists, but there is no method available which uses the special properties of overlapping areas in a sequence of small-frame images for a fully automatic and seamless extraction. Semi-automatic tracking methods can be applied for the tasks in disaster management described before, and can theoretically operate in a sequence of single small-frame images, too. However, the automation level is low, because a great deal of user interaction is demanded. In Jing et al. 2008, a tracking method is presented which uses texture of the road and direction information for tracking. This approach allows a highly precise and robust tracking by an angular texture template with mean-value and standard deviation, but the borders of the template area are fixed. They are determined by the operator at the start of the tracking without adaptation during the tracking. The algorithm published in Hua et al. 2008, tracks roads at border-edges utilizing precise edge detection in an iterative approach.

A direction is calculated and used as criterion for a selection of edges to continue. Although those approaches show a quite high performance, a well-known and so far not solved problem of road trackers is related to junctions where the tracking stops.

The main contribution of this paper consists in enforcing the automatic road tracking by exploiting the redundant image information provided through the overlapping area of adjacent images. None of the existing approaches apply supervised classification to use redundant information in overlapping images and, additionally, improve the reliability by a fusion with gradient-edges. The combination of complementary image features – image segments from classification and image edges – also ease the inclusion of road junctions in the extraction, since the tracking does not only rely on a local template.

In the next Section 2 details of our approach are described. In Section 3 experimental results are presented and discussed. Finally, Section 4 concludes the paper.

2 Approach for Road-Tracking

This section describes the proposed approach for road-tracking. The basic idea is to perform a simple Maximum Likelihood supervised classification, combined with edge extraction by the Sobel filter. Classification supports the tracking, because of its ability to segment roads roughly depending on global image characteristics, the training information is retrieved from the previous extraction result in overlapping image parts. Edge extraction is used for the delineation of the road border, thus adding local object information.

2.1 *Model*

Firstly, the chosen road-model of the tracking approach is presented. In the real world, roads are connected objects mostly covered by asphalt or concrete. Therefore, the road model used for the road-tracking regards roads as a completely connected area, homogeneous in color and texture (Fig. 1, top left). In an image, a road is represented by homogeneous pixels and a gradient-edge at the borders. Roads have a constant width, represented through a constant distance between two opposite road-borders in the image. The image-space in the large frame of Fig. 1 shows exemplarily the area of homogeneous pixels (marked red) and the gradient-edges at the borders (marked blue).

Fig. 1: Homogeneous road area (red) and border edges (blue), original image (top left).

The tracked road-network is represented as a graph delineating the road-borders. The graph is connected and consists of one mesh surrounding the road-area, representing the borderline. Only nodes of degree 2 exist. However, nodes of degree 1 occur when a connection between two nodes is removed to add new road parts. The concept of a connected graph is used although only nodes of degree 1 and 2 occur. An alternative would be to represent the road by single lines or one mesh, respectively, but the notation of graphs allows more flexibility regarding the extension of a road at a specific node, and the conversion between line and mesh is feasible straight forward. The nodes contain the global coordinates of the border points to represent the actual shape. The spatial distance between the nodes can vary to represent the real road-area (cf. Fig. 2).

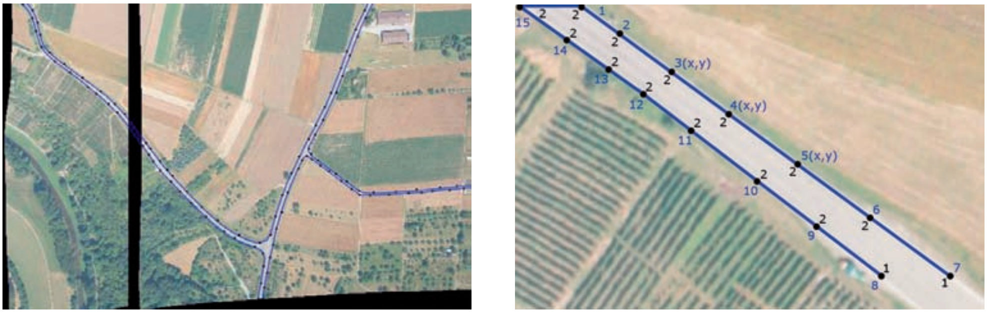

Fig. 2: Road outline superimposed on two images (left); detail properties of graph (right): node id (blue number), node degree (black number).

2.2 *System and Strategy*

In this section an overview of the proposed approach is given. The basic system for road-tracking requires the following input information: An image sequence of overlapping, georeferenced orthoimages with a minimal overlap of 30 %, a decision concerning one constant direction to process the sorted images and knowledge about the road-color. The system controls the iterative tracking of the road-network over the individual images of the sequence (cf. Fig. 3). *Roads-of-interest* are determined by manually digitizing the area of demanded roads in a *start image* with a closed polygon. Since the road is internally represented as a closed graph, the initially provided road information is called the *Start-Graph*. The edges of the graph represent the road borders, at the image-border the graph is closed in order to obtain a valid description.

The strategy is based on the processing of each image depending on extracted areas of the previous image (cf. Fig. 4). A transition to the next image is achieved by transformation of the existing graph from the previous image into the next image. The transition is either forward or backward in the image sequence, since the images are required to be georeferenced and orthoprojected. In the transformed image, called the *Current Image*, the part of the road-network contained in the overlapping area of the two images is available. The edge of the graph representing the current road outline which intersects with the current image is removed, resulting in two nodes with degree 1 and enabling an extension of the road area. A supervised Maximum Likelihood classification using training data from the overlapping area of the current image delivers initial road segments. The training area is explicitly defined by mapping the road graph into the current image. Gradient-edges in the current image, extracted by the Sobel filter, determine road-borders precisely. Tracking is performed by continuing the graph at its open sections by adaptation of the road-parts in the non-overlapping area of the current image. The adaptation is accomplished by an iterative tracking along the borders creating new nodes within the current graph as origin. The *Resulting Graph* for the road-extraction is the result of tracking and is closed at the image borders. Based on the result-graph, the next overlapping image in the sequence is chosen to perform the pre-described processes. The resulting graph is equal to the final result, if no remaining overlapping images are available (cf. Figs. 3 and 4).

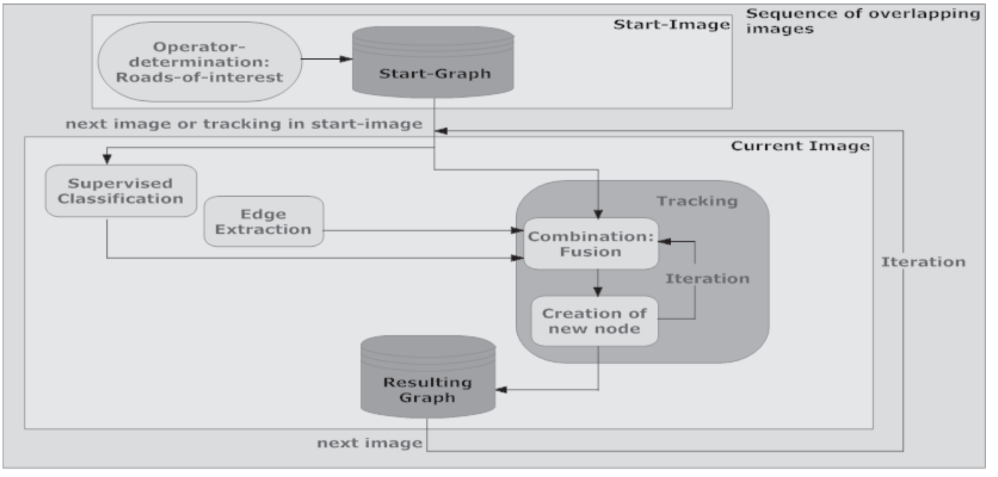

Fig. 3: System of road-tracking.

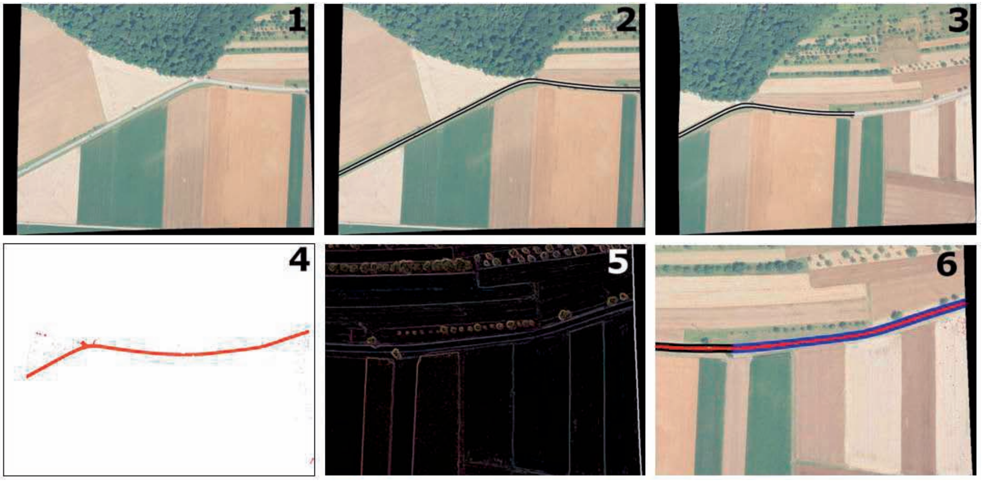

Fig. 4: Example of system process: Start image (1), determination closed start graph (black) (2), switch image (3), classification (4), edge extraction (5), continuation of graph (blue) (6).

2.3 *Combination of Classification and Edge Extraction*

One important part of the road-tracking system is the fusion of areal and linear features by a combination of image segments retrieved by the classification and the image edges. This combination is necessary for the decision on how to continue the graph at an open node. For the tracking, rules are defined to decide which source is used to extract parts of a road-border at particular points.

The rules for the combination are derived from the general road model explained above. Concerning the continuation at an open node-point, the first rule claims that the start-point of an edge must lie inside a small radius around the node because the border-lines of a road must be completely connected. The second rule demands that the edge does not deviate much in direction and collinearity, because the border-lines at straight road-sections are collinear. Since roads are supposed to have a constant width, the third rule states that an edge must build a constant width with the opposite one: The end-point of an edge and the current node must create a constant radius to its nearest node on the other roadside. Moreover, an edge must be located near the border of the classified region. Besides the rules for standard road-sections, exceptions for junctions, occlusions and road-ends are defined. At a junction, the deviation in mean-direction is allowed to be larger because the road-model regards junctions as sections with a larger change in direction. Since occlusions are modeled as interruptions in the connected borderline, the radius is allowed to be larger. The exceptions concerning occlusion over larger parts, road-entries and road-ends in the image admit the deviation from mean-direction to be larger.

Different rules have been defined to realize the described concept. The most important constraint is that the edges have to comply with the mean-direction provided by the classified region where it is connected. For the selection of a candidate first a gradient-edge candidate fulfilling the model-constraints gets connected. Edges are preferred since they determine road-borders locally more precisely. They must keep a small distance to the border of the segment from the image classification in order to guarantee the validity. In case no gradient-edge is found in the current region of interest, points on the classified region-border are searched and connected. The two separate representations and a morphologic closing of the classification result avoid the inclusion of edges with a larger deviation. To handle the exceptions junction, road-end, road-entry and partly occlusion, it is sufficient to claim that an edge has to lie near the segment border fitting to the mean-direction. As the second choice points on the segment border which can have a larger deviation from the mean-direction are permitted.

Small occluded areas, e.g., caused by single trees, are bridged through the morphologic closing, applied to the classification result, but larger gaps have to be closed in a post-processing step.

2.4 *Post-Processing: Closing of Gaps in the Network*

A post-processing step examines the graph representing the final result in order to close gaps in the network between opposite road-ends. Currently, only gaps on straight roads are found automatically. In Fig. 5 the principle is sketched. First, all edges representing road-ends are removed. In a second step collinear road borders are searched, whereas the distance between the endpoints must not exceed a certain threshold. After closing the respective gaps, the remaining road-ends are closed again to yield a closed graph.

3 Results and Analysis

3.1 *Results of Road-Tracking*

The proposed road-tracking approach was tested with a dataset from the 3K-camera system, a small frame camera developed by the German Aerospace Center (DLR) (Kurz et al. 2007). The data was chosen from a test-se-

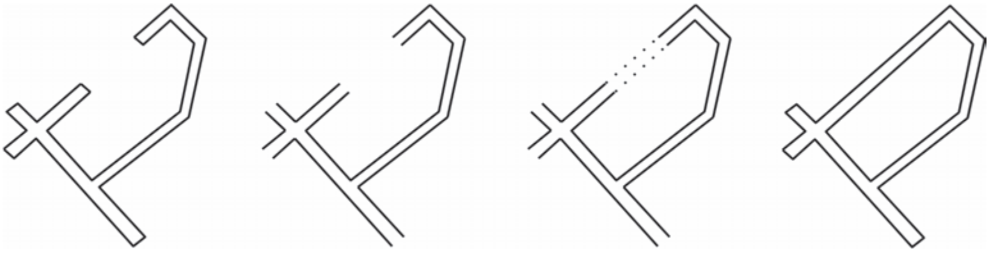

Fig. 5: Closing gaps (from left to right): Initial net with gap; removing edges in the graph not representing road borders; finding collinear road borders; final net with closed gap and closed road-ends.

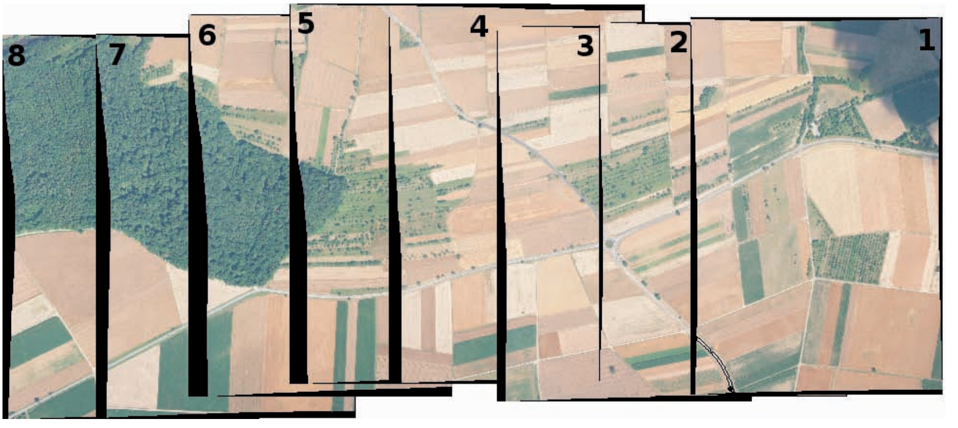

Fig. 6: Image sequence A: The start graph is delineated in image 1 (black).

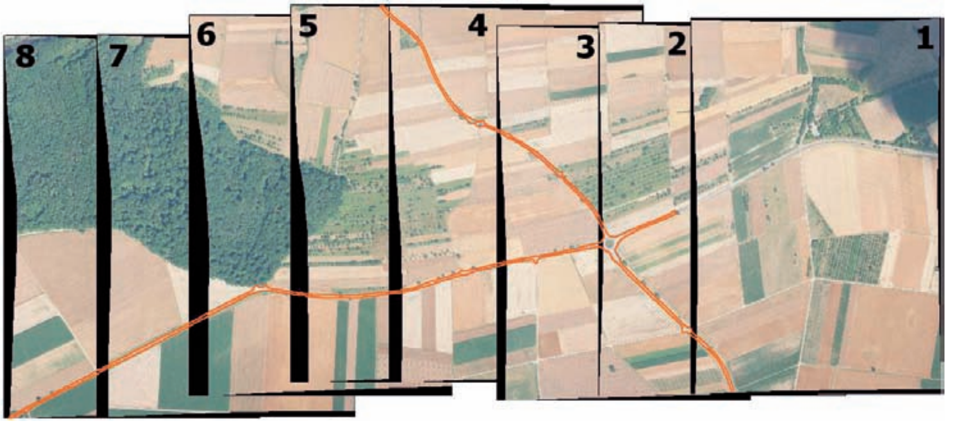

Fig. 7: Image sequence A: Result of road-tracking started from image 1 (red).

Fig. 8: Orthoimage mosaic A: Result of road-tracking started from the right (black), derived road parts up to required new initializations are color-coded (blue, red, yellow).

quence captured at a flight over Vaihingen an der Enz.

The results derived from the given dataset show that main-roads are tracked reliably. In sequence A (cf. Fig. 6), main-roads with similar color are shown. The roads are uniquely and clearly separated from the background. The road-features are long parts, junctions and road-entrances with maximal 90° deviation in mean-direction. Most of the main-roads were extracted and connected.

The tracking works reliably and reaches relatively high quality values: A completeness of 76 % and a correctness of 89 % (cf. Fig. 7). The simple shape of the road-features and the smooth borders allow a robust tracking. The extracted network includes junctions and road-entrances with a similar color to the main-roads. The classification supports the robust tracking because main-roads are unique in color compared to the background. In addition, complex junctions continuing in different directions, for example the roundabout in Fig. 7 in image 2 and image 3, are tracked correctly. The tracking along borderlines, separately and independently at each roadside, allows for an extraction independent from the shape of road-sections. However, tracking cannot be continued reliably at complete occlusion or shadows, as for example shown in Fig. 7, image 2 (right). Tracking stops because the segmented area is not completely connected.

To further examine the influence of using individual overlapping images on the approach, the same experiment was conducted a second time using an orthoimage mosaic. The result is shown in Fig. 8: Starting from the same initial segment (black) as defined in Fig. 6, the tracking stops already after a short road section (blue). In contrast to the result from the image sequence depicted in Fig. 7, two additional initial training sets have to be provided to restart the tracking process, see Fig. 8 red and yellow. This demonstrates the superior strategy of the proposed new approach using the images of the sequence directly, where for each image the classification is automatically trained exploiting the redundant information of the overlapping area. If the image segments derived through classification contain too many background pixels, the tracking may stop at one individual road side. The graph describing the road outline is then closed without further consideration of image information, which requires \ manual correction of the graph (cf. different colors of opposite road sides in Fig. 8).

Besides the tests for main roads, extractions of field-paths are performed in test sequence B (cf. Fig. 9). The example contains a successful extraction of field-paths. The width of field-paths is small and the signature of field-paths in feature space is quite similar to the surroundings, leading to a less reliable and accurate classification. The small width of field-paths is of no disadvantage for the tracking method. If a clear classification can be reached, good quality values for tracking are possible reaching a completeness of 56 % and a correctness of 88 % (cf. Fig. 10).

In this sequence the classification is accurate enough to track the road-network robustly. The borders at opposite road-sides are tracked reliably and are clearly separated from each other. However, problems occur at field-paths containing a color being too similar to the surrounding. The borders of the segments from classification are detected wrongly or changes in mean-direction of the borderlines are too large, since no clear classification of roads is possible.

The same experiment was conducted using an orthoimage mosaic and the results are shown in Fig. 11: Again, the tracking starts from the same initial segment (black) as defined in Fig. 9 and stops already after a short road section (blue). Two additional trainings for the classification are necessary to obtain results similar to the proposed approach using the images of the sequence separately.

A high redundancy by a large image overlap is of advantage for the classification. Fig. 12 shows an image which was classified with training-pixels determined by the resulting-graph of the previous image. An overlap of 60 % results in reliable segments with a completeness of 79 % and a correctness of 93 %.

A less reliable classification is gained if only a smaller overlap is available, see Fig. 13 for an example with 30 % overlap of adjacent images: Now, the quality values achieve a completeness of only 75 % and a correctness of 93 %. Less pixels of the road are classified

Fig. 9: Image sequence B: The start graph is delineated in image 1 (black).

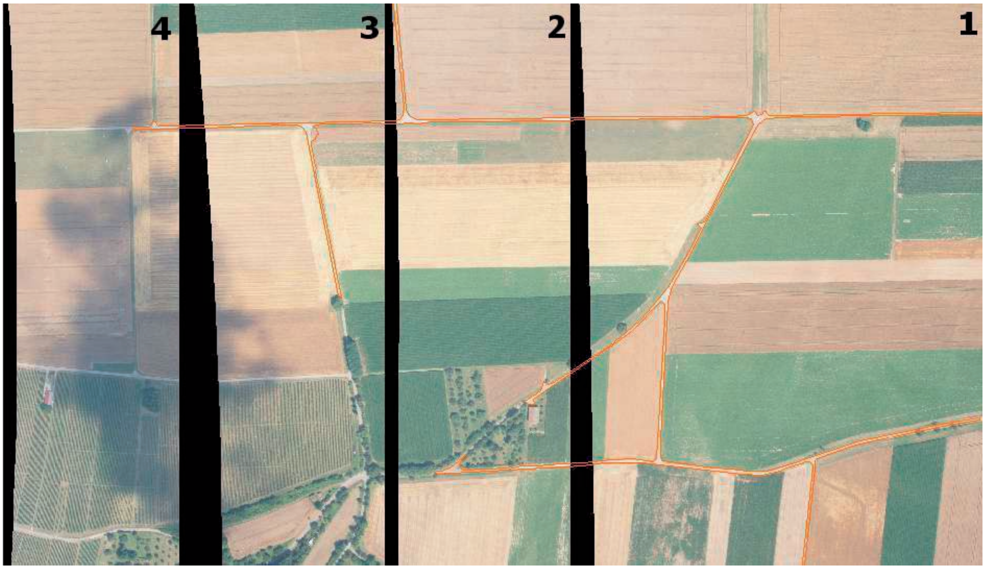

Fig. 10: Image sequence B: Result of road-tracking started from image 1 (red).

correctly. The borders are detected less correctly compared to the results shown in Fig. 12. For example, a larger background area is segmented as road which is directly connected with the road-segment (cf. Fig. 13).

3.2 *Results of Post-Processing*

Though the basic tracking approach works well in open landscapes, occlusions caused by trees or similar background colors lead to gaps in the road network (cf. Fig. 10). Gaps are closed by the post-processing operation to re-

Fig. 11: Orthoimage mosaic B: Result of road-tracking started from the right (black), derived road parts up to required new initializations are color-coded (blue, red, yellow).

connect roads as described in Section 2.4. In Fig. 14 an example is shown where a gap of the basic tracking is closed.

4 Conclusions

In this paper, an approach for the tracking of roads in sequences of single overlapping airborne small- frame images by exploiting redundancy is presented. It combines a supervised classification with edge extraction in subsequent overlapping images in order to track roads in rural areas. The results show that the combined use of classification and edge extraction makes the method quite robust. Since training areas for the classification are defined automatically from the overlap with the previous image, the feature space is adapted ideally to the particular image. In case the road is not well separable from the background, for example at narrow field paths, the classification is not reliable and leads to an incomplete extraction. Experiments conducted with orthoimage mosaics produced from the same input images demonstrate the added value of exploiting the overlap region for automatic adaptation of the feature space: Using the same initial training region, manual intervention, including provision of additional training regions, is necessary to yield a similar result as from individual overlapping images.

The classification quality and thus the road extraction quality depend on the realized image overlap as the example shows. In practical applications, however, this is not critical as the standard overlap at least in flying direction is 60 % or more.

The proposed method to close gaps in a post-processing step has limitations, because it only considers the roads extracted so far. In addition, the current implementation is only able to complete gaps on straight road segments. A better approach would attempt to use the images again to verify and enhance the road extraction.

Further developments may go into different directions. One could use existing GIS data to support the road classification. In that case more emphasis has to be put on identifying erroneous training data. It needs to be accounted for different road surfaces, i.e., to use different road classes. For some platforms like light-weight UAV an accurate INS is not available

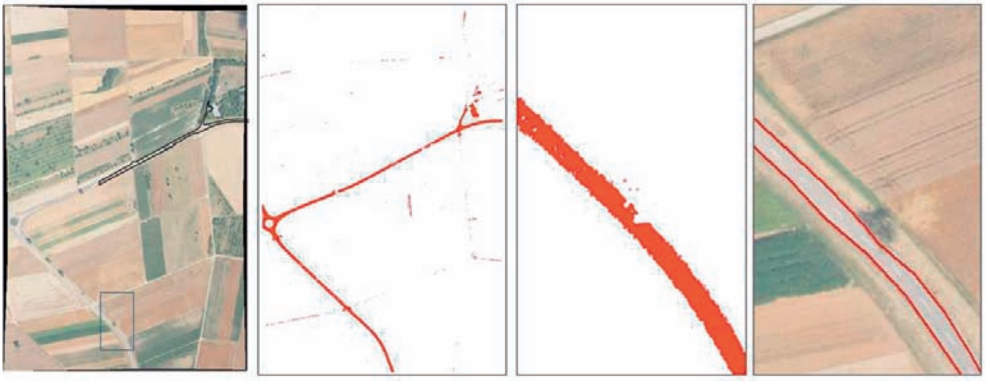

Fig. 12: Road extraction result at 60 % image overlap (from left to right): Initial road outline projected into image to delineate the training area (black); classification result; zoom shows good quality of classification; finally delineated road.

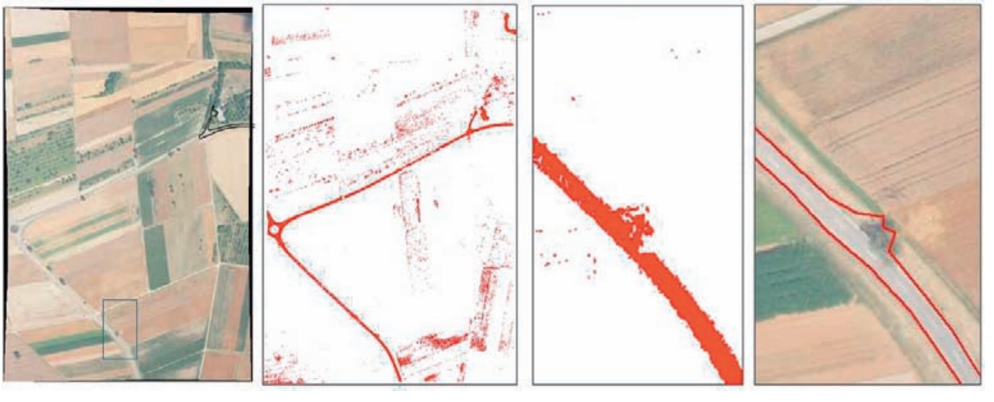

Fig. 13: Classification and tracking result at 30 % image overlap (from left to right): Smaller training area than for 60 % because of smaller overlap; classification shows much more false positive errors; zoom in shows larger error than for 60 % overlap; delineation with errors.

Fig. 14: Post-processing by topological connection of occluded road-sections: Stop at complete occlusion (left), connection through post-processing (right).

and, thus, georeferenced orthoimages can not be computed. However, for quite a number of applications, like quick inspection of UAV images, it is desired to track roads anyhow, even without georeference. Therefore, a further goal is to reach independency from real-time navigation and orthorectification. This can be achieved by computing only the relative orientation from image to image. The node points of the graph need then be transformed from one frame to the next using the particular transformation.

Acknowledgements

The 3K-camera images are provided by the German Aerospace Center (DLR).

References

Bacher, U. & Mayer, H., 2005: Automatic Road Extraction from Multispectral High Resolution Satellite Images. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3): 29–34.

Butenuth, M., Reinartz, P., Lenhart, D., Rosenbaum, D. & Hinz, S., 2009: Analysis of Image Sequences for the Detection and Monitoring of Moving Traffic. – Photogrammetrie – Fernerkundung – Geoinformation **5/09:** 421–430.

Hinz, S., Wiedemann, C. & Baumgartner, A., 2000: A scheme for road extraction in rural areas and its evaluation. – Fifth IEEE Workshop on Applications of Computer Vision, 134–139.

Hua, H., Ying, L., Xun, W., Xia-jun, Z. & Bin, X., 2008: Road Extraction in Remote Sensing Images Using a New Algorithm. – IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 779–782.

Jing, S., Xiangguo, L., Yunfei, S. & Cheng, W., 2008: Knowledge-Based Road Extraction from High Resolution Remotely Sensed Imagery. – IEEE Congress on Image and Signal Processing, 608–612.

Kurz, F., Müller, R., Stephani, M., Reinartz, P. & Schroeder, M., 2007: Calibration of a Wide-angle Digital Camera System for Near Real Time Scenarios. – ISPRS Hannover Workshop 2007, High Resolution Earth Imaging for Geospatial Information, on CD.

Mayer, H., Hinz, S., Bacher, U. & Baltsavias, E., 2006: A Test of Automatic Road Extraction Approaches. – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3): 209–214.

Mena, J. & Malpica, J.A., 2005: An Automatic Method for Road Extraction in Rural and Semiurban Areas Starting from High Resolution Satellite Imagery. – Pattern Recognition Letters **26** (9): 1201–1220.

Rosenbaum, D., Charmette, B., Kurz, F., Suri, S., Thomas, U. & Reinartz, P., 2008: Automatic Traffic Monitoring from an Airborne Wide Angle Camera System. – International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences **37** (3b): 557–562.

Thomas, U., Rosenbaum, D., Kurz, F., Suri, S. & Reinartz, P., 2009: A new Software/Hardware Architecture for Real Time Image Processing of Wide Area Airborne Camera Images. – Journal of Real-Time Image Processing **4** (3): 229–244.

Vosselman, G. & de Knecht, J., 1995: Road tracing by profile matching and Kalman filtering. – Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhaeuser, 265–274.

Addresses of the Authors:

Dipl.-Ing. Mathias Koller, Dr.-Ing. Matthias Butenuth, Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, Arcisstr. 21, D-80333 München, Tel.: +49-89-28923857, Fax: +49-(0)89-2809573, e-mail: mathias.koller@gmx.de, matthias.butenuth@bv.tum.de

Dr.-Ing. Markus Gerke, University of Twente, Faculty of Geo-Information Science and Earth Observation, P.O. Box 217, 7500 AE Enschede, The Netherlands, Tel.: +31-53-4874522, e-mail: gerke@itc.nl

Manuskript eingereicht: November 2009
Angenommen: Juli 2010