THE PHOTOGRAMMETRIC RECORD

The Photogrammetric Record 27(137): 8–28 (March 2012)
DOI: 10.1111/j.1477-9730.2011.00670.x

ROAD NETWORK EXTRACTION IN SUBURBAN AREAS

Anne Grote (grote@ipi.uni-hannover.de)
Christian Heipke (heipke@ipi.uni-hannover.de)
Franz Rottensteiner (rottensteiner@ipi.uni-hannover.de)

Leibniz Universität Hannover, Germany

(Based on a presentation at the Ian Dowman Retirement Symposium, entitled ''Progress and Prospects for Photogrammetry and Remote Sensing in a Changing World'', held at University College London on 28th June 2010)

Abstract

In this paper, an algorithm for the extraction of road networks in suburban areas is presented. The algorithm is region-based and uses high-resolution colour infrared images as well as, optionally, a digital surface model (DSM). The road extraction starts with a segmentation using the normalised cuts algorithm; afterwards the segments are grouped. Road sections are extracted from the grouped segments. Road sections that are likely to belong to the same road are connected to subgraphs in the next step. To eliminate false connections in the subgraphs, context objects such as vehicles, buildings and trees are employed. The remaining road strings, represented by their centre lines, are connected to a road network. The process employs combinations of radiometric and geometric features, derived from knowledge about the appearance of roads in suburban areas. Results are presented for two test data-sets, acquired by different sensors. A quantitative analysis is performed for the quality of the road extraction as well as the topological quality of the extracted network.

Keywords: completeness and correctness, context, network generation, normalised cuts, road extraction, suburban areas

Introduction

Accurate and up-to-date road databases are very important for many applications. Change detection and the update of such databases is usually done manually with the help of aerial or satellite images, but it is desirable to automate this process. For open landscapes, several methods have been developed that are reliable enough for practical application, at least for change detection (Zhang, 2004; Gerke and Heipke, 2008). In urban areas, the task is more difficult because the scene is more complex so that many assumptions about roads are frequently violated. For instance, in urban and suburban areas roads do not stand out as clear elongated linear objects due to the occurrence of crossroads or due to occlusion by neighbouring objects such as buildings.

There is a great variety of automatic road extraction approaches. Many of them can only be used in a specific environment or with specific types of images. There are different approaches for open rural landscapes, for inner city areas and for semi-urban areas. An important distinguishing mark for road extraction algorithms is the underlying road model, which is closely related to the resolution of the images used for road extraction: in images having a ground sampling distance (GSD) of 1 to 2 m or larger, roads appear as thin elongated lines, so that road extraction can be based on line extraction algorithms. In images having a smaller GSD, roads appear as homogeneous areas, so that different methods have to be applied.

Road extraction algorithms based on a linear road model are frequently used in open landscapes. Several of them apply the line extraction method developed by Steger (1998); for instance, in Wiedemann (2002) the extracted lines are connected by searching for shortest paths in the road network, motivated by a road network model that postulates fast connections between distant places as the main function of a road network. Bacher and Mayer (2005) use the results of line extraction to define training areas to classify the image. This is followed by a second extraction within the road class. Both these methods are developed for rural areas. In urban areas, line-based approaches are of limited use because of the complexity of the scene. Therefore, line-based approaches for urban areas usually employ additional constraints or data sources. For example, Youn et al. (2008) locate the road centre lines in the two dominant directions at places where a long line intersects with only a few other line pixels. Hu et al. (2004) search for long straight lines by a Hough transform, supported by a digital surface model (DSM) from lidar and a vegetation mask to determine regions of interest for road extraction.

There also exists a great variety of region-based road extraction approaches, where roads are modelled as elongated regions. Zhang (2004) describes an approach for change detection in rural or semi-urban areas. Regions of interest are determined by an unsupervised classification around existing roads from a database, using a DSM as an additional source. In the region of interest, parallel edges are considered to delimit road segments. Various other features, such as road markings, are used to assist the extraction. Zhang and Couloigner (2006) describe a road extraction algorithm for urban areas with 1 m multispectral satellite images. They start with an unsupervised *k*-means classification, where each observation is assigned to the cluster whose mean is nearest to the observation. The road class is identified using a fuzzy classifier based on the assumption that the reflectance for roads in the infrared band is lower than in the other bands. The pixels belonging to the road class are further classified using a shape-based descriptor; only pixels that belong to elongated regions are kept. Hinz and Baumgartner (2003), working in dense urban areas, determine regions of interest using a DSM. Inside the regions of interest, roads are extracted based on a combination of lines and edges from high-resolution (0·2 m) images with ribbons in lower resolution. The extraction is guided by several rules and internal confidence checks; it relies quite heavily on road markings and uses vehicles as context objects. Poullis and You (2010) classify a high-resolution satellite image into road pixels and non-road pixels using pixel colour and orientation in a graph cut algorithm.

There is not one perfect road extraction method that is valid for all scenes and data-sets. The method has to be adapted to the scene and the data because roads have different appearances in different types of scene. In the EuroSDR test on road extraction (Mayer et al., 2006), the participating approaches performed significantly worse in urban scenes compared to the rural scenes for which the majority of them were developed. Line-based approaches usually do not perform well in urban areas because the roads do not stand out clearly and the whole scene content is more complex. Region-based methods (and, consequently, high-resolution images) are better suited for road extraction in suburban areas.

In this paper, a new method for the extraction of roads in a suburban context is presented. The properties of suburban scenes pose several challenges: objects close to the road or on the road can partly cover the road (for example, trees and vehicles), or buildings and trees can cast shadows on the road, all of which disturbs the appearance of the roads in an image. Additionally, several assumptions about roads and road networks that are frequently employed do not hold in such areas: road markings, used for the extraction of roadsides and for the verification of extracted roads in some approaches, are rare; additionally, the main network function of roads is not to provide fast connections, but to give access to every place. Therefore, the road network can have an irregular shape; dead ends and culs-de-sac may be frequent. The present method takes these scene properties into account. It requires very high-resolution images (approx. 0·1 m GSD) in order to deal with the complex environment of suburban areas. In contrast to other approaches, road extraction starts with a region-based segmentation of the image instead of a search for parallel edges. An unsupervised segmentation is used that does not rely on assumptions about the road surface colour and can be more easily transferred to images of different origins. Knowledge about the appearance of roads is already used in the segmentation. The extraction of roads is based on radiometric and geometric properties. The method does not rely on the presence of road markings, straight roads or a particular form of the road grid. Contextual objects such as trees, buildings and vehicles are used to assist road extraction and to prevent false extractions. A DSM can be used optionally as additional information. The main focus in this work lies on the extraction of the centre lines of the network; achieving a high geometric accuracy of these centre lines is not the primary goal.

METHOD

Overview

The approach described in this paper requires very high-resolution (approximately 0·1 m GSD) colour infrared ortho-images and, optionally, a DSM. Table I shows an overview over the steps and the features used in each step. Several geometric and radiometric features are used; to start with, the radiometric features are dominant, whereas in later steps the geometric features gain importance. Contextual objects are also used in later steps. Following a region-based road model, the road extraction process starts with a *segmentation* of the image based on the *normalised cuts* method (Shi and Malik, 2000), as outlined later. One of the advantages of this method is that it can consider knowledge about the appearance of roads in the image. This is followed by a *grouping* of the segments to compensate for any over-segmentation that has occurred in the first step. From the grouped segments, hypotheses for *road sections* are extracted by classifying the segments as road segments or non-road segments.

Due to the complexity of the scene and expected disruptions by objects such as cars, the roads are not required to be extracted in one piece from junction to junction: gaps between extracted road sections are allowed. Collinear road section hypotheses that possibly correspond to one and the same road are connected to *road subgraphs*. The term *subgraph* suggests that such a road subgraph does not represent a global road network but rather a local part of the network. *Branches* can occur in the road subgraph, especially when false extractions are present (Fig. 1). Therefore, the subgraphs are optimised with the goal of obtaining single *road strings* by removing the branches that are least likely to correspond to a road. Road strings are road subgraphs that do not contain branches. Contextual objects are used to evaluate the branches for the optimisation. In the last step, the *road network* is generated. This comprises the search for junctions and the elimination of falsely positive road strings. The final result is a road network, comprised of roads represented by their centre lines with associated widths, and

TABLE I. Overview of the features used for the steps of road extraction.

Step	*Radiometric features*	*Geometric features*	*Context objects*
Initial segmentation (similarity between pixels)	Edges (gradient image) Colour Hue NDVI	Distance	
Grouping (similarity between segments)	Edges (gradient image) Colour channel histograms NDVI	Absolute length of shared border Relative length of shared border Convexity Height	
Road section extraction (compliance of segments with road model)	Intensity Standard deviation of intensity NDVI	Area Length Elongation Convexity Width Width constancy Height	
Subgraph generation and evaluation (connection of road sections to roads)	Colour difference	Distance Direction difference Continuation smoothness Road quality Width difference	Vehicles Buildings Trees Vegetated areas Asphalt areas
Network generation (connection of roads to network)		Distance Road quality Direction difference Length	Buildings Trees Vegetated areas

NDVI is the normalised difference vegetation index.

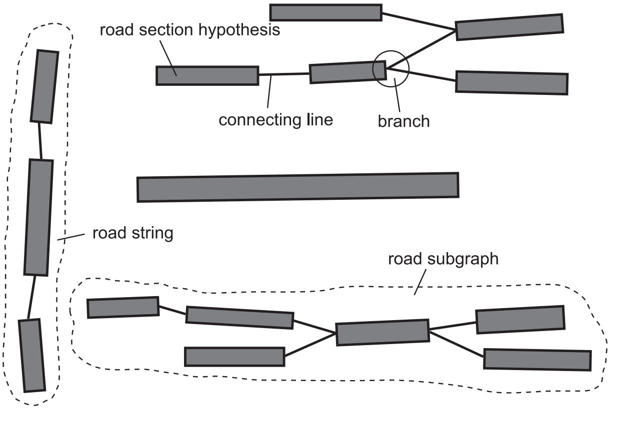

FIG. 1. Definition of road sections and road subgraphs. Grey rectangles: road section hypotheses; continuous lines: connections; dashed lines encircle one subgraph. Road strings are subgraphs without branches.

junctions represented by the points where the lines meet. Each road connects two junctions, except for dead end roads, which connect to only one junction. The individual stages of the road extraction process are explained in detail in the subsequent sections.

Segmentation and Grouping

The goal of the initial segmentation is to divide the image into segments in such a way that all road borders correspond to segment borders: ideally, one segment should only contain road pixels or non-road pixels, but not both. Several features derived from knowledge about the appearance of roads are employed because there is not a single feature that sets apart roads from all other objects. In addition, global characteristics are also used to prevent small disturbances from distorting the segmentation. Both goals are achieved by employing the normalised cuts algorithm (Shi and Malik, 2000).

In the normalised cuts framework, the image is represented by a graph whose nodes are the pixels and whose edges connect any pair of pixels in the image. Each edge between two pixels $\mathbf{x}_k$ and $\mathbf{x}_l$ is assigned a weight w_{kl} that represents the similarity of the pixels. In order to segment an image, the graph is split into parts by removing edges so that the similarity of pixels from different segments is minimised whereas, at the same time, the similarity of pixels within the individual segments is maximised. This corresponds to minimising the *normalised cut criterion*; further details can be found in Shi and Malik (2000) and Yu and Shi (2003).

Object knowledge can be inserted into the segmentation process via the definition of the edge weights w_{kl}. The definition is based on several features designed to make segment boundaries coincide with road boundaries.

(a) *Distance*: pixel pairs having a distance larger than a threshold (10 pixels) are assigned a weight $w_{dist} = 0$, whereas $w_{dist} = 1$ for all other pairs.
(b) *Colour difference*: if the distance between two pixels in colour space is long, the pixels are considered dissimilar, which is expressed by a weight function w_{col} that decreases with that distance.
(c) *Hue difference*: if the hue between two pixels is significantly different, the pixels are considered dissimilar. A weight $w_{hue} = 1$ is assigned to any pair of pixels having a hue difference smaller than a threshold, and $w_{hue} = 0$ otherwise.
(d) *Edges*: two pixels are considered dissimilar if there is an edge with a high absolute value of the gradient between them. This is used to define a weight w_{edge}.
(e) *NDVI (normalised difference vegetation index) difference*: the image is segmented into vegetated regions and non-vegetated regions by a threshold operation on the NDVI image. If one pixel of a pixel pair belongs to a vegetated region and the other belongs to a non-vegetated region, the pixels are considered dissimilar, which is expressed by a weight $w_{NDVI} = 0$. Otherwise, $w_{NDVI} = 1$.

Each similarity criterion is transformed into a weight that is in the range between 0 (no similarity) and 1 (identical), and the weights for the five features are combined into a total weight w for that pixel pair by multiplication:

$$w = w_{dist} \cdot w_{col} \cdot w_{hue} \cdot w_{edge} \cdot w_{NDVI}. \tag{1}$$

More details about the calculation of the weights and the application of the normalised cuts algorithm can be found in Grote (2011). The normalised cuts algorithm is computationally very intensive, so it cannot be computed completely in one step for the images used in the experiments. Therefore, the image is divided into rectangular tiles that are segmented independently. The number n of segments per tile is fixed and has to be specified before applying the normalised cuts segmentation. In this application, it is selected so that an over-segmentation is achieved in order to increase the chances of obtaining segment borders for all road sides. The segmentation results of the individual tiles are merged, and the combined results are used in all subsequent processes.

In order to ameliorate the effects of the over-segmentation, in the next step the initial segments are grouped into larger, more meaningful segments. As a consequence of the tiling process, some artificial boundaries between segments at the tile boundaries exist. They are treated as regular segment boundaries, but merging two segments across a boundary caused by tiling is made more likely than merging other segments. An iterative approach is used for grouping: in each step, several pairs of segments are merged based on a number of features that are calculated for each pair of initial segments. The features implicitly encode the following knowledge about the roads.

(a) *Edge strength*: two segments should not be merged if the mean absolute gradient value along the shared border of both segments is high.
(b) *Histogram difference*: if the grey value histograms of the regions are different, the regions should not be merged. For the comparison of the histograms, the sum of the differences between corresponding bins (L_1 norm of the Minkowski distance) is used. This measure gives good results when comparing segments of the same image (Rubner et al., 2001).
(c) *Absolute length of the shared border*: two segments should not be merged if the border they share is very short.
(d) *Relative length of the shared border*: this is the ratio between the length of the shared border of two regions and the minimum of the lengths of the two overall borders. Two segments should not be merged if the relative length of their shared border is short to prevent the formation of highly irregular segments.
(e) *Convexity of the merged region*: two segments should not be merged if the convexity of the merged segment, namely, the ratio between the segment area and the area of the convex hull of the segment, is low. This prevents irregular segments, such as segments with long protruding arms.
(f) *NDVI difference*: the average NDVI for both regions is compared to a threshold. Two regions must only be merged if the average NDVI is on the same side of that threshold.
(g) *Height difference*: two segments must not be merged if the difference of their average heights is larger than a threshold, so that building segments are not merged with ground segments (only used if a DSM is available).

The NDVI and the height difference are used as veto features: if a segment pair does not fulfil the respective criterion, the segments will not be merged. The values of all other features are evaluated together in order to decide whether the segments can be merged. From each feature value, membership values for fuzzy sets are calculated, indicating whether the segments can be merged with respect to the respective feature or not. They are then combined using a set of rules that assess the features against each other. For example, if at least two of the colour, edge and convexity criteria are very good and the third is still good, the criterion for the relative border length can be disregarded. The rules were found empirically by clustering region pairs that should be merged in the initial segmentation of a test scene into characteristic groups; the decision as to whether two regions should be merged was based on visual inspection. A list of the rules, as well as a more comprehensive description of the way in which they were found, is given in Grote (2011). For all segment pairs that receive a positive merge decision, a cost function C is used to sort the segment pairs. The cost function C is determined from the values of the grouping features *convexity*, *histogram difference*, *edge strength* and *relative border length*, normalised by the maximum value for each feature in the current iteration cycle:

$$C = \frac{1 - convexity}{\max convexity} + \frac{histogram\ diff}{\max histogram\ diff} + \frac{edge\ strength}{\max edge\ strength} + \frac{rel.\ border\ length}{\max rel.\ border\ length}. \quad (2)$$

In each iteration cycle, 10% of the segments with the lowest costs are merged. The iteration continues until no more segment pairs receive a positive merge decision. The result of the segmentation and grouping are segments that are relatively homogeneous in colour, and most segments belonging to road areas are large enough to be evaluated by shape descriptors in the next step.

Road Section Extraction

After grouping, the segments are classified according to whether they correspond to a road segment or not. As it is often not possible to extract a whole road completely because of disturbances in the appearance of roads, the focus lies on the extraction of reliable *road section hypotheses*. The extraction should be reliable enough to generate the network: the number of false positives should be small enough to allow their elimination in later steps, whereas there should be a sufficient number of correct road segments so that the gaps between them can be filled by subsequent processes, when more global knowledge can be exploited. For road section extraction, the compliance of each segment with several criteria based on shape and spectral characteristics of roads is checked as follows.

(a) *Intensity*: the segment's average intensity must be within a certain (relatively wide) range to exclude very dark and very bright segments.
(b) *Standard deviation of intensity*: should be relatively low for road sections.
(c) *NDVI*: the average NDVI should be low for road sections.
(d) *Area and length*: road sections should have a minimum length and a minimum area. The minimum length of a road section should be significantly longer than the average width of a road, in order to allow a proper evaluation of the other shape criteria.
(e) *Elongation and convexity*: a road section should have a high elongation. If the elongation is relatively low, the convexity (defined in the same way as for the grouping) should be high. If the elongation is sufficiently large, the convexity need not be as high to allow the extraction of curved roads.
(f) *Width and width constancy*: the width of a road section should be close to the average width of suburban roads, and it should be relatively constant.
(g) *Height*: if a DSM is available, the average height of a road segment should be similar to the surrounding ground.

These features are calculated for each region and checked against thresholds. Some of these thresholds can be derived directly from model knowledge about roads, for example, those for the minimum area, the minimum length and the acceptable width range. Others were found by manual evaluation of a number of segments comprising both acceptable road segments and non-road segments. Non-road segments have a considerably larger variation in the values than road segments; therefore, only regions fulfilling all criteria are selected as reliable road section hypotheses. The geometric criteria are image invariant and only need to be scaled according to the image resolution. The radiometric criteria are more dependent on sensor characteristics and illumination conditions. For the images used in the tests, only the NDVI threshold had to be adjusted.

After the extraction of road sections, the main direction of each road segment is determined. Then, adjacent road sections are checked to determine if they can be merged. This is beneficial because in the previous grouping step adjacent road segments are not always merged due to stricter criteria concerning the convexity and relative border length. Two adjacent road sections are merged if they have a common border of at least 1 m and sufficiently similar main directions, and if the merged road section meets the criteria described above for extracted road sections.

The features *length*, *elongation*, *width constancy* and the *deviation from the average road width* are used to determine a *road quality* measure of the road section hypothesis. They are mapped on an interval between 0 and 1 using membership functions for a fuzzy set ''road'' such that higher values suggest a higher chance for the segment to actually correspond to a road section. For the road width, the membership function is trapezoidal so that the values around the average road width are mapped to 1. For the other features, the membership functions increase linearly between a lower and an upper threshold. All transformed values are multiplied to obtain a single road quality measure.

Each road section hypothesis is represented by its region and the centre line of the region. The centre line is determined based on a distance transform of the segment borders. The boundary of the road section is split into two parts at the points farthest away from each other, and for each of these two parts a distance transform is computed. The difference of the distance transform images of both border sections is calculated, and the centre line then corresponds to the points where the distance difference is 0. The centre line thus determined is very uncertain near its end points. Hence, a certain percentage of the centre line next to both ends of the line is replaced by straight line segments having the same directions as the sections of the centre line next to the removed segments.

Road Subgraphs

After road section extraction, many roads correspond to one road section from one junction to the next. However, disturbances in the appearance of roads can interfere with the extraction and cause gaps between extracted road sections. In order to bridge the gaps, road sections are connected to their neighbours if they potentially belong to the same road, forming road subgraphs. Two road sections may belong to the same road if their geometric relations indicate that they follow the same course. The following features are used.

(a) *Distance*: both the *absolute* and the *relative* distance between two connected road sections should be lower than a threshold. The absolute distance is the distance between the two nearest end points of the centre lines. The relative distance is the ratio of the absolute distance and the length of the shorter road section.
(b) *Direction difference*: the direction of a road section is defined by the vector connecting the two end points of its centre line. The direction difference, that is, the angle between the direction vectors of two road sections, should be lower than a threshold for the two road sections to be connected.
(c) *Continuation smoothness*: the lateral offset between two connected road sections should be lower than a threshold. The lateral offset is related to the angle between the direction of a road section and the direction of the line connecting it to the other road section; this angle is determined for both road sections, and both angles should be lower than a threshold for the two road sections to be connected. If the distance between the road sections is very short, the continuation smoothness is disregarded

because at close distances the angles depend too much on the exact positions of the end points of the centre lines, which are relatively uncertain.

The subgraphs are generated iteratively, starting from the road section that received the best road quality measure in the road section extraction. Two road sections are connected if the thresholds for the distance, the direction difference and the continuation smoothness are met. The search for neighbouring road sections continues until no other road section can be found that meets the connection criteria to any of the road sections in the current subgraph. Subgraph generation then continues with the road section having the best road quality measure among those not yet assigned to any road subgraph. This process is repeated until all road sections have been assigned to a subgraph. As a result, there may be subgraphs consisting of only one road section as well as subgraphs composed of several road sections. Subgraphs consisting of several road sections may have branches that correspond to competing hypotheses for the course of a road (Fig. 1). These ambiguities must be resolved in order to obtain road subgraphs consisting of one branch only, because in most cases, branches occur with falsely extracted road sections. Two aspects of the connection properties are considered to determine weights for the lines connecting two road sections: the interrelation properties of the two connected road sections, expressed by an *interrelation weight* w_I, and the properties of the gap between them, expressed by a *context weight* w_C that depends on extracted *context objects*.

The interrelation weight w_I of a connection between two road sections is a measure of the plausibility of both road sections belonging to the same road given their geometric configuration. The following features are used to calculate w_I.

(a) *Distance*: a short distance between the road sections leads to a high weight.
(b) *Direction difference*: a small angle between the two main road directions as defined above leads to a high weight.
(c) *Continuation smoothness*: a smooth continuation between the road sections as defined above leads to a high weight. In this case, the value for the larger smoothness angle is used to calculate the weight.
(d) *Road quality measure*: good quality values from road section extraction (as defined above) for both road sections lead to a high weight, which is defined as the mean value of both quality values.
(e) *Colour difference*: a low difference of the mean colour values of the road sections leads to a high weight. The colour difference between both road sections is calculated from the mean values of the colour channels. The channel with the largest difference is used to calculate the weight.
(f) *Width difference*: a low difference of the widths of both road sections leads to a high weight.

The values of all features are mapped linearly onto the interval [0, 1] such that a weight of 1 indicates that it is likely for the two road sections to belong to the same road given their respective features. For example, the value for the distance is mapped linearly onto the interval [0, 1] such that a distance of 0 corresponds to the weight 1, and the maximum distance from the subgraph generation corresponds to the weight 0. The individual weights from the six features are multiplied together to yield the interrelation weight $w_I \in [0, 1]$ for the connecting line:

$$w_I = w_{distance} \cdot w_{direction\ diff} \cdot w_{smoothness} \cdot w_{quality} \cdot w_{colour\ diff} \cdot w_{width\ diff}. \tag{3}$$

The second aspect of a connection between two road sections concerns context objects that can be found in the gap between the two road sections. The context objects are extracted automatically using very simple methods that are not optimised to yield complete and entirely correct results: the context objects can support the road extraction without being completely extracted. For brevity and because the extraction of context objects is not the focus of this work, only a short overview is given about the extraction methods; more details can be found in Grote (2011). The following context objects are considered.

(a) *Vehicles* are extracted as small homogeneous regions fulfilling size and shape criteria related to compactness, rectangularity, eccentricity and area. They may consist of several regions representing vehicle parts.
(b) *Shadows* are extracted as compact regions having low grey values.
(c) *Trees* are extracted as compact regions with a high NDVI and associated shadows; if a DSM is available, it is used to assist the tree extraction.
(d) *Buildings* are extracted as high and compact objects with low NDVI; they are only used if a DSM is available.
(e) *Vegetated areas* are areas having a high NDVI but not being trees.
(f) *Asphalt areas* are areas with the average grey values of asphalt.

After the extraction of the context objects, the context weight w_C is computed for each connection between two road sections. First, a road hypothesis is constructed for the gap bridged by a connection from the line connecting the two centre lines and the average road widths of the roads connected at either side. The context objects can contribute to the total context object weight of the connection in two ways, namely, by their relation to the road hypothesis in the gap and by the amount of occlusion caused by them. The *context relations* are described by the locations and orientations of context objects relative to the road hypothesis. Depending on the context relations and the type of the context object, a context object can support or contradict a road hypothesis. For instance, a vehicle on a road hypothesis would give support to the road hypothesis, while a building on a road hypothesis would contradict it very strongly. Buildings or trees alongside a road hypothesis, on the other hand, can give a weak support, as those arrangements are common in suburban areas. A *relation value* is assigned to each context object given its relation to the road hypothesis. This relation value reflects the strength of contradiction (negative values) or support (positive values) for the road hypothesis by the context object. An overview on all context relations and their values is given in Table II. The relation values for all objects in the gap are summed to a *context relation weight* w_{Crel}. If no relevant context object relation is found, w_{Crel} will be set to 0. If more than one context object is assigned to the same relation class, the impact of subsequent objects on the total value is decreased: the first occurrence of such a relation is considered to be more significant for the evaluation of a connection hypothesis than the subsequent ones.

For the *occlusion* analysis the context objects *vehicle*, *tree* and *shadow* are considered, because these are context objects that can occlude parts of the road or at least prevent the

TABLE II. Overview of context object relations and the corresponding relation values. Positive values indicate support for a road connection hypothesis, negative values indicate contradiction.

Relation	*Value*	*Relation*	*Value*
Building, parallel, next to road	0·4	Vehicle, perpendicular, next to road	0·3
Building, perpendicular, next to road	0·4	Tree next to road	0·4
Building, diagonal, next to road	0·1	Tree on road	−0·5
Building on road	−10·0	Vegetated area next to road	0·2
Vehicle, parallel, on road	0·5	Vegetated area on road	−10·0
Vehicle, perpendicular, on road	0·2	Asphalt area on road	0·2
Vehicle, parallel, next to road	0·3	Asphalt area next to road	0·1

correct classification of a road. For a given gap, a high degree of occlusion by these context objects supports a road hypothesis more than a low degree of occlusion. The occlusion weight w_{Cocc} is the percentage of the area of the road hypothesis that is covered by the context objects. The context relation weight w_{Crel} and the occlusion weight w_{Cocc} are summed to yield the final *context weight* w_C of a connection hypothesis:

$$w_C = w_{Crel} + w_{Cocc}. \tag{4}$$

After the calculation of both the *interrelation weight* w_I and the *context weight* w_C, both weights are combined to yield a *combined weight* w_{Conn} for the connection. The combination of weights follows a set of rules that consider the values of the respective weights as well as the length of the gap. If the gap is short, the context weights will have less impact on the combined weight than otherwise because, for short gaps, the interrelations are much more important than context objects (which may not even be present). Therefore, only the interrelation weight is used for gaps shorter than the average road width. For longer gaps, the mean of the relation weight and the context object weight is calculated. Additionally, if a building or a vegetated area was found on the road hypothesis, the hypothesis is rejected no matter how good the interrelation weight is.

The combined weights w_{Conn} of the connecting lines are used for the evaluation of all connections in a subgraph having branches. This is carried out by maximising the sum of all weights w_{Conn} of the remaining edges in the subgraph subject to the constraints that only one connecting line should be attached to each end of a road section in the subgraph. This optimisation problem is formulated as a linear program (see, for example, Dantzig and Thapa, 1997). Its solution results in a set of *road strings* that do not contain branches. More details about the optimisation for road subgraph evaluation can be found in Grote et al. (2009).

Network Generation

The generation of the road network starts with the determination of a road centre line and an average road width for each road string. The centre line of a road string is initialised by concatenating the centre lines of the individual road sections. After that, the centre line is approximated by a polygon in an iterative way, starting with a straight line between the end points of the original centre line. If the average distance of the original centre line points is higher than a threshold, the original centre line is split into two segments of equal length by a new vertex inserted in the middle of the original centre line. This procedure is recursively applied to the new polygon segments, until the average distance of the original centre line from the approximating polygon is below a threshold. Inserting new vertices at equidistant intervals proved to be a better way of eliminating unwanted bulges in the centre line compared to methods that put the new vertex at the location of the longest distance between the original and the approximation (such as in the Ramer–Douglas–Peucker algorithm outlined in Ramer (1972)). The average width of the road string is determined by calculating the mean of the average widths of the individual road sections, weighted by their lengths. In the same way, the road quality measure of the road is derived from the road quality measures of the road sections.

Subsequently, pairs of parallel roads that lie close together are searched for. If the distance between the parallel roads is shorter than a typical width of a block of houses, only one road is kept. If one road is significantly longer than the other, the longer road is kept. Otherwise, the road having the better quality measure is kept. This step is particularly important if no DSM is available as it very efficiently eliminates false extractions that lie on building roofs parallel to extracted roads.

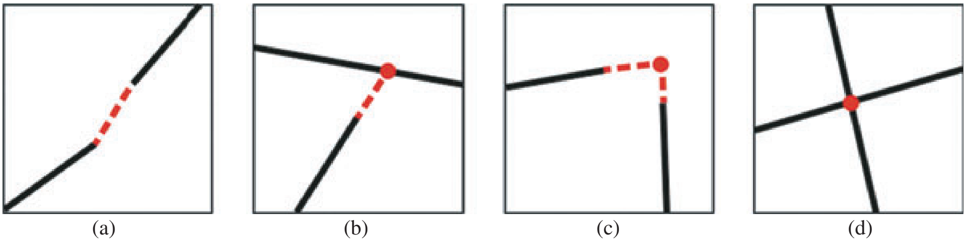

FIG. 2. Possible connections between two roads.

Next, junction connections are searched for among the remaining roads. At the end of each road, a search region is set up whose radius depends on the quality measure of the road: a road with a good quality measure receives a large search radius. If another road is found inside the search region, it is checked whether the two can be connected. Depending on whether both roads are collinear or not, the connection hypothesis is created in different ways. If the roads are nearly collinear, which means they have a small directional difference, their end points are connected if the end point of the second road lies inside the search region of the first (Fig. 2(a)). Otherwise, the junction connection is constructed from the extension of one road, if this extension intersects the other road; the junction point is located at the point of intersection (Fig. 2(b)). If no extension intersects the other road, the junction is constructed from the extensions of both roads; in this case, the junction point lies at the intersection of both extensions (Fig. 2(c)). Additionally, intersections between extracted roads are searched for: if two roads directly intersect, a junction point is created at the point of the intersection (Fig. 2(d)).

The connection hypotheses are evaluated using context objects: a connection hypothesis is discarded if buildings, trees or vegetated areas cover the region of the connection to a significant degree. This region is defined to be a rectangle whose axis is the junction connection line and whose width is the average width of the road. A connection hypothesis is discarded if buildings cover more than 10% of the connection area or if a building crosses the junction connection line. A connection hypothesis is also discarded if at least 80% of the tree area covers the connection area, if a vegetated area covers more than 20% of the connection area or if the connection area crosses the vegetated area. All the remaining connections are maintained. In the cases depicted in Figs. 2(b) and (d), this implies that road strings have to be split at the junction points.

The final stage is to detect connected components of roads in the road network. These connected components are checked for significance: in each connected component, the total length of all roads must be larger than the total length of all junction connections, and the total length of the connected component (roads and junction hypotheses together) must exceed a minimum. An exception for the last condition is made if at least two open ends of the connected component lie near the image border; then it is possible that the connected component is connected to the road network by road sections beyond the image border. The final result of the method is a network consisting of the roads, represented by their centre lines, and the junctions, represented by junction points where the road centre lines meet.

EXPERIMENTS

The method was tested on two different data-sets comprising high-resolution aerial colour infrared images and DSMs. The first data-set depicts a scene in Grangemouth, Scotland; it consists of an ortho-image with a GSD of 0·1 m derived from a scanned aerial image and a DSM derived from image matching with manual post-processing. The DSM has a GSD of 0·2 m and a height resolution of 0·1 m. The second data-set consists of an ortho-image with a GSD of 0·08 m derived from a digital aerial image and a DSM having a GSD of 0·5 m derived from lidar data. This second data-set is a part of the Vaihingen (near Stuttgart) test data-set of the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010).

For the quantitative analysis, one image subset showing suburban scene characteristics was examined for each data-set. The subset from the Grangemouth data-set comprises an area of 562 m × 485 m (5617 × 4849 pixels) and contains a road network with a total length of approximately 3·75 km and 20 junctions. The subset from the Vaihingen data-set comprises an area of 394 m × 357 m (4929 × 4465 pixels) and contains a road network with a total length of 2·41 km and 15 junctions. The road network extraction was carried out twice for each data-set, namely, with and without the DSM. The parameter settings remained the same for both data-sets, except for the threshold of the NDVI that is used in the segmentation, in the grouping phase and in the road section extraction.

Quality Measures

The quality measures as defined in Wiedemann et al. (1998) and Wiedemann and Ebner (2000) were determined for the extracted road networks, based on a comparison of the extracted road network to a reference that was generated manually. In this process, both the extraction results and the reference roads were represented by their centre lines. The first group of quality measures consists of the *completeness* and the *correctness* of the detected road centre lines. In order to determine the completeness, a buffer is constructed around each extracted road centre line, and the length l_{Comp} of all reference line segments inside the buffer is determined. The buffer width is set to ±5 m according to the typical road width in the test areas. The completeness is defined as the ratio of the length of reference roads inside the buffer l_{Comp} compared to the overall length l_{Ref} of the reference road network. In order to determine the correctness, the buffer is constructed around the reference roads, and the length l_{Corr} of all extracted road centre lines inside a buffer is determined. The correctness is the ratio of the length of extracted roads inside the buffer l_{Corr} compared to the overall length l_{Extr} of the extracted road network:

$$completeness = \frac{l_{Comp}}{l_{Ref}} \tag{5}$$

$$correctness = \frac{l_{Corr}}{l_{Extr}}. \tag{6}$$

The *root mean square error of the distance* RMS_D is a measure of the geometrical accuracy of the extracted roads. It is computed from the shortest distances d_{re} of any point on the correctly extracted roads from their nearest reference road:

$$RMS_D = \sqrt{\frac{\sum d_{re}^2}{l_{Comp}}}. \tag{7}$$

A high geometric accuracy of the road centre lines is not expected as the centre lines are approximated and the buffer for the completeness and correctness checks is relatively wide.

Selecting a smaller buffer width would result in better values for RMS_D at the cost of lower completeness and correctness values.

A third group of quality measures is related to the topology of the extracted road network. The mean detour factor F_D is the factor by which a path between two points in the extracted network is longer on average than in the reference:

$$F_D = \frac{1}{n_P} \sum_{\{i,j\} \in P} \frac{D^E_{i,j}}{D^R_{i,j}}, \tag{8}$$

where P is the set of all node pairs i, j that are connected in both the extracted network and the reference network and n_P is the number of these pairs, whereas $D^E_{i,j}$ and $D^R_{i,j}$ are the shortest paths between the nodes i and j in the extraction results and the reference, respectively.

The *topological completeness* measures the ratio of connections between node pairs in the reference network that also exist in the extracted network. Similarly, the *topological correctness* is the ratio of connections in the extracted network that also exist in the reference network. In order to determine these numbers, matching of nodes between the two networks is required. With $n_{m\ conn}$ denoting the number of matched connections, $n_{ref\ conn}$ the number of reference connections and $n_{extr\ conn}$ the number of extraction connections, the quality measures are determined from

$$topological\ completeness = \frac{n_{m\ conn}}{n_{ref\ conn}} \tag{9}$$

$$topological\ correctness = \frac{n_{m\ conn}}{n_{extr\ conn}}. \tag{10}$$

The last group of quality measures is related to the extracted junctions. The junction completeness is the ratio of junctions in the reference that could be assigned to extracted junctions based on a distance criterion defined similarly to the evaluation of road centre lines. It is calculated in a similar way to the topological completeness (equation (9)), only using the respective numbers of junctions. The correctness of the junction extraction is the ratio of extracted junctions that could be assigned to junctions in the reference and is calculated in a similar manner to the topological correctness (equation (10)). The buffer for the junction evaluation is larger than for the road evaluation because the extraction of the junction points is not as accurate; the junctions are not directly extracted from the image, but often reconstructed from extensions of the ends of the road. Inaccuracies of the road centre lines are usually largest at the ends of roads, which results in an even higher offset at the ends of the extensions. Therefore, the buffer radius is set to 15 m for the junction evaluation. The geometric accuracy of the matched junctions is further evaluated by calculating the root mean square difference of the junctions RMS_J, which is calculated analogously to RMS_D (equation (7)).

Results and Evaluation

Fig. 3 shows the extracted networks of both scenes (extraction with the DSM) and the evaluation results: correctly extracted roads (true positives) are shown in green, missed roads (false negatives) in blue and erroneously extracted roads (false positives) in red. The extracted network of the Grangemouth scene consists of two connected components; the extracted network of the Vaihingen scene is more fragmented. One road near the upper right corner of

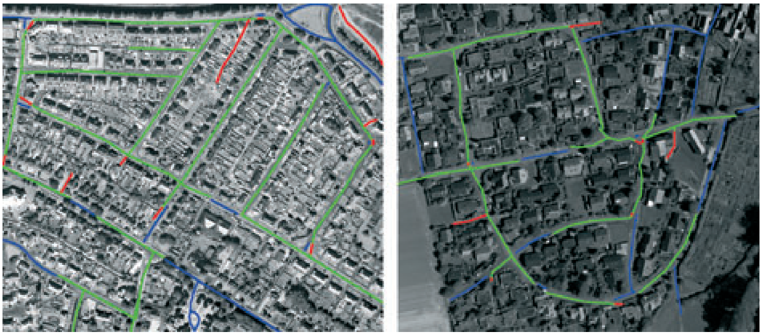

FIG. 3. Extracted road networks for the case when a DSM was used. Left: Grangemouth, with DSM. Right: Vaihingen, with DSM. True positives in green, false positives in red, false negatives in blue.

TABLE III. Results for road extraction evaluation.

	Completeness (%)	*Correctness (%)*	*RMS_D (m)*
Grangemouth with DSM	76·9	90·1	1·48
Grangemouth without DSM	72·5	83·3	1·61
Vaihingen with DSM	58·1	91·0	1·66
Vaihingen without DSM	46·5	86·5	1·47
Grangemouth with DSM, reference without tree-covered roads	82·7	89·8	1·47
Vaihingen with DSM, reference without tree-covered roads	69·1	90·1	1·66

the Grangemouth scene was deleted because it is parallel to another road and lies very close to it; this case is not supported by the employed road model.

The completeness, the correctness and the RMS_D of the extraction results for different variants are summarised in Table III. In general, the correctness of the extraction results is high. Using a DSM improves the correctness as well as the completeness of the extraction considerably. Firstly, it prevents segments at different height levels from being merged, and secondly it helps to avoid the confusion of buildings with roads. The completeness is somewhat lower than the correctness. Missed extractions are partly caused by decreasing geometric accuracy, especially towards the ends of roads. Such road ends are often geometrically inaccurate, in particular where junction connections are inserted: the orientations of end sections of extracted centre lines often differ from those of the true centre line, an effect that is aggravated by the extrapolation required in junction connection. Missed extractions also occur where no road sections could be found and where the resulting gap was too large to be bridged during the subgraph and network generation phases. This occurs if a row of trees or buildings covers the road for a significant stretch, or if the shape of the segment does not comply with the criteria for the road section extraction, for example, because it was merged with an adjacent parking area or a driveway in the grouping process. The main cause for the lower completeness in Vaihingen is the existence of some narrow roads in the reference that are almost totally covered by trees or building shadows, to an extent that they are difficult to recognise even for a human operator (note the three false negative roads in the upper right quadrant of the right image of Fig. 3). The evaluation was repeated using a reference without the roads that are, for a large part, covered by trees (Fig. 4). When the extracted roads are

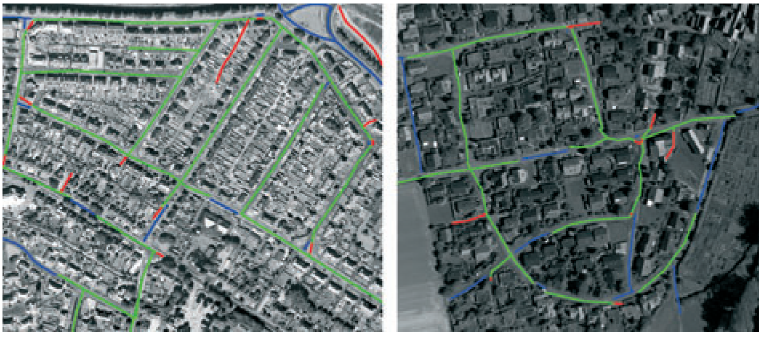

FIG. 4. Extracted road networks compared to reference without tree-covered roads. Left: Grangemouth. Right: Vaihingen. True positives in green, false positives in red, false negatives in blue.

evaluated using this reference, the completeness significantly improves, especially for the Vaihingen data-set. When a DSM is used, most false positives are caused by extracted driveways or car parks. Otherwise, some building roofs were incorrectly extracted as roads.

The RMS_D values in Table III are in the order of 1·5 m in all examples. At first glance, this looks relatively poor given the resolution of the data, but geometrical accuracy is not the most important goal of the application. The most important aspect is the correct representation of the road network. The root mean square values are caused, to a large extent, by uncertainties in the region boundaries near crossroads or in the vicinity of driveways (which have a similar colour to roads). The geometrical accuracy could be improved by a better method for centre line generation or by a post-processing step for precise road edge detection.

The results for the network quality are summarised in Table IV. In general, these results also improve with the use of a DSM. The topological correctness is good. The lower correctness achieved for Grangemouth without the DSM is caused by a false connection between two components that are not connected in the reference. For the Grangemouth example with DSM, the lower topological completeness is caused by a gap that separates two connection components of the extracted network, whereas in the Grangemouth example without the DSM there are several gaps that separate components of the network. The extracted networks in both Vaihingen examples are more fragmented, which is caused by failures in the road section extraction, partly due to trees covering the roads and partly due to segments having irregular shapes or being too small, such that no roads could be extracted.

TABLE IV. Results for network quality evaluation.

	Detour factor	*Topological correctness (%)*	*Topological completeness (%)*
Grangemouth with DSM	1·4	100	75·2
Grangemouth without DSM	1·2	89·6	67·4
Vaihingen with DSM	1·1	100	37·3
Vaihingen without DSM	1·0	100	31·0
Grangemouth with DSM, reference without tree-covered roads	1·4	100	76·2
Vaihingen with DSM, reference without tree-covered roads	1·1	100	37·6

Table V. Results for junction evaluation.

	Junction completeness (%)	*Junction correctness (%)*	*Junction RMS (m)*
Grangemouth with DSM	70·0	73·7	6·96
Grangemouth without DSM	75·0	68·2	8·25
Vaihingen with DSM	33·3	45·5	5·37
Vaihingen without DSM	13·3	40·0	3·90
Grangemouth with DSM, reference without tree-covered roads	68·4	68·4	6·19
Vaihingen with DSM, reference without tree-covered roads	45·5	45·5	5·70

The results of the junction extraction evaluation are summarised in Table V. Falsely extracted junctions were most frequently caused by the extraction of driveways or car parks adjacent to the roads. Missed junctions are often at places where a gap close to a T-junction could not be bridged because the distance between both roads was too large, or because the verification of the junctions failed when the connection touched buildings or vegetated areas. The rather low geometric accuracy of matched junctions is mainly caused by the lower geometric accuracy of the road centre lines towards the ends of the roads. The reason for this is that the correct end points of the road centre lines can be difficult to determine from the road sections when their borders have irregular shapes.

Conclusions

The results show that the algorithm can extract roads in suburban areas with good results. The correctness is about 90% for extraction with the DSM. A common cause of false results is the extraction of driveways and car parks which are not contained in the reference but share some characteristics with roads. About half of these objects present in the scene were extracted. The completeness is not as good as the correctness; whereas more than 75% completeness was achieved in the Grangemouth scene, it drops to 58% in the Vaihingen example. Lower completeness is often caused by trees or building shadows which cover large parts of some roads, especially in the Vaihingen data-set. As Table III shows, the completeness is significantly better (69% for the Vaihingen data-set) when roads covered by trees are not considered.

The quality of the network topology is correlated with the completeness of the network. Significant decreases in the topological completeness are caused by gaps that separate components that should be connected in the network. This happens especially in the Vaihingen data-set where the extracted network consists of four separated components. The gaps between them were often not bridged because of inaccurate location of the road ends, leading to a rejection of the connection because of interfering context evidence. The comparison between the results with and without the DSM shows that the use of the DSM improves the extraction considerably. Whereas the method can still produce quite reasonable results without a DSM, the DSM should be used if it is available.

Mayer et al. (2006) claim that a completeness of at least 60% and a correctness of at least 75% are the absolute minimum for road extraction results to be considered useful in practice; for real practical considerations the completeness should be at least 70% and the correctness at least 85%. The algorithm presented here achieves these goals for the correctness when a DSM

TABLE VI. Comparison of results of different approaches. If several numbers from different data-sets were given, the average was calculated.

	Hinz (2004)	*Mena and Malpica (2005)*	*Zhang and Couloigner (2006)*	*Youn et al. (2008)*	*Poullis and You (2010)*	*Current approach (with DSM)*
Completeness	79·1%	25·0%	56·0%	80·0%	80·6%	67·5%
Correctness	96·9%	74·0%	41·0%	79·0%	75·3%	90·6%
RMS	1·9 m	1·13 m	1·52 m	2·32 m	–	1·57 m
Top. comp.	92·0%	–	–	–	–	56·3%
Top. corr.	98·1%	–	–	–	–	100·0%
GSD	< 0·2 m	1·0 m	1·0 m	0·1 m	?	0·1 m

Top. comp. is topological completeness; top. corr. is topological correctness.

is used; in terms of completeness the results of the Grangemouth data-set also fulfil these requirements. The results for Vaihingen are close to the goal for completeness, but significantly fail to achieve the goal for correctness.

Table VI compares the results of several other approaches for road extraction in urban areas with the present approach. Of course, such a comparison is not conclusive because the results were obtained using different data-sets, but it still gives an indication that the approach operates on a similar (if not better) level compared to other approaches. In the EuroSDR test on road extraction (Mayer et al., 2006), several algorithms were tested on an IKONOS suburban scene. The best results in terms of completeness were achieved by Poullis and You (2010). In terms of correctness, Hinz (2004) achieved the best result with 74%. The approach described in the current paper gives very good results for the correctness. Only Hinz (2004) (who uses images of a similar resolution as in this paper) reports better results. The results of the algorithm used in this paper in terms of completeness are not as good, but it has to be noted that Hinz (2004) only gives results for two small test sites that are, in addition, from the same scene; the current results for Grangemouth are comparable to those of Hinz (2004). The RMS value is also comparable to the other results. Evaluations of the topological quality of the road network are rarely given for urban approaches. Many approaches do not connect single extracted roads to a network at all (for example, Youn et al., 2008); others complete the network interactively (see, for example, Poullis and You, 2010). Hinz (2004) is the only one to give results on the topological quality. The topological correctness is similar to that of the present approach, for the topological completeness his results are better. However, the test sites used in Hinz (2004) have at most four junctions as opposed to, for instance, 20 junctions in the Grangemouth data-set. In summary, the results of the algorithm presented in this paper are comparable to those of other approaches; in terms of correctness, the results are better than most of them. The completeness, however, leaves room for improvement.

There are several areas where the quality of the extraction could be improved. As already noted, the most important ones are the completeness of the network and the geometric quality, especially of the junctions. The completeness can be improved by closing further gaps after the final network check. One possibility is to search for reasonable connections between the unconnected ends of roads and junction points. However, this must be done with care to prevent the connection of dead ends to other roads. A similar search could be done with the isolated roads that are currently deleted during the final network check. Improved modelling of the context objects could be used to bridge larger gaps in the network generation, for example, those caused by rows of trees. For improvement of the geometric quality of the junctions, specific models and extraction methods for junctions are required. The geometric accuracy of the network could also be improved using a snake-based approach or by using the extracted road centre lines as approximate locations that help to precisely detect the edges delineating the road.

Other important improvements are automatic parameter learning and sensitivity testing, since the algorithm uses a rather large number of parameters. The parameters are quite stable, as shown by the fact that they were tested on images from two different scenes and different sensors. Only one parameter, the NDVI threshold, had to be changed. However, in future work the aim will be a systematic training of the parameters using a stochastic model to enhance the general applicability of the algorithm.

ACKNOWLEDGEMENTS

The linear program was solved using the program lp_solve available at http://lpsolve.sourceforge.net/5.5/. The normalised cuts segmentation was implemented in C++ based on the Matlab implementation by T. Cour, S. Yu and J. Shi (http://www.seas.upenn.edu/~timothee/software_ncut/software.html). The Vaihingen data-set was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) (Cramer, 2010) (http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html).

REFERENCES

BACHER, U. and MAYER, H., 2005. Automatic road extraction from multispectral high resolution satellite images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3/W24): 29–34.

CRAMER, M., 2010. The DGPF-test on digital airborne camera evaluation—overview and test design. *Photogrammetrie Fernerkundung Geoinformation*, 2010(2): 73–82.

DANTZIG, G. B. and THAPA, M. N., 1997. *Linear Programming 1: Introduction*. Springer Series in Operations Research. Springer, New York. 435 pages.

GERKE, M. and HEIPKE, C., 2008. Image-based quality assessment of road databases. *International Journal of Geographical Information Science*, 22(8): 871–894.

GROTE, A., 2011. *Automatic road network extraction in suburban areas from aerial images*. Ph.D. thesis, DGK Series C, No. 663. 96 pages.

GROTE, A., HEIPKE, C., ROTTENSTEINER, F. and MEYER, H., 2009. Road extraction in suburban areas by region-based road subgraph extraction and evaluation. *Proceedings of Urban Remote Sensing Joint Event*, Shanghai, China. 6 pages (on CD-ROM).

HINZ, S., 2004. Automatic road extraction in urban scenes—and beyond. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B3): 349–355.

HINZ, S. and BAUMGARTNER, A., 2003. Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2): 83–98.

HU, X., TAO, C. V. and HU, Y., 2004. Automatic road extraction from dense urban area by integrated processing of high resolution imagery and LIDAR data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B3): 288–292.

MAYER, H., HINZ, S., BACHER, U. and BALTSAVIAS, E., 2006. A test of automatic road extraction approaches. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3): 209–214.

MENA, J. B. and MALPICA, J. A., 2005. An automatic method for road extraction in rural and semi-urban areas starting from high-resolution satellite imagery. *Pattern Recognition Letters*, 26(9): 1201–1220.

POULLIS, C. and YOU, S., 2010. Delineation and geometric modeling of road networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(2): 165–181.

RAMER, U., 1972. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1(3): 244–256.

RUBNER, Y., PUZICHA, J., TOMASI, C. and BUHMAN, J. M., 2001. Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision and Image Understanding*, 84(1): 25–43.

SHI, J. and MALIK, J., 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8): 888–905.

STEGER, C., 1998. An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(2): 113–125.

Wiedemann, C., 2002. *Extraktion von Straßennetzen aus optischen Satellitenbilddaten*. Ph.D. thesis, TU München, Deutsche Geodätische Kommission, Reihe C, Vol. 551. 94 pages (in German).

Wiedemann, C., Heipke, C., Mayer, H. and Jamet, O., 1998. Empirical evaluation of automatically extracted road axes. In *Empirical Evaluation Techniques in Computer Vision* (Eds. K. J. Boyer & P. J. Phillips). Wiley–IEEE Computer Society Press, Los Alamitos, California, USA. 262 pages: 172–187.

Wiedemann, C. and Ebner, H., 2000. Automatic completion and evaluation of road networks. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33(B3/2): 979–986.

Youn, J., Bethel, J. S., Mikhail, E. M. and Lee, C., 2008. Extracting urban road networks from high-resolution true orthoimage and lidar. *Photogrammetric Engineering & Remote Sensing*, 74(2): 227–237.

Yu, S. X. and Shi, J., 2003. Multiclass spectral clustering. *Proceedings of the Ninth IEEE International Conference on Computer Vision*. 1525 pages: 313–319.

Zhang, C., 2004. Towards an operational system for automated updating of road databases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4): 166–186.

Zhang, Q. and Couloigner, I., 2006. Automated road network extraction from high resolution multi-spectral imagery. *Proceedings of the ASPRS Annual Conference*, Reno, Nevada, USA. 10 pages (on CD-ROM).

Résumé

Cet article présente un algorithme d'extraction du réseau routier en zone suburbaine. L'algorithme fonctionne à l'échelle de la région et utilise des images infrarouge couleur et éventuellement un modèle numérique de surface. Le processus d'extraction des routes commence par une segmentation basée sur l'algorithme des coupures normalisées (normalised cuts algorithm), après quoi les segments sont regroupés. Les tronçons de routes sont extraits à partir des segments regroupés. Les tronçons qui semblent appartenir à la route sont connectés en sous-graphes lors de l'étape suivante. Afin d'éliminer les fausses connexions dans les sous-graphes, des objets contextuels comme des véhicules, des bâtiments et des arbres sont utilisés. Les tronçons restants, représentés par leurs axes centraux, sont connectés à un réseau routier. Le processus s'appuie sur une combinaison de propriétés radiométriques et géométriques, issues de connaissances sur l'aspect des routes en zone suburbaine. Des résultats sont présentés pour deux jeux de données expérimentaux, acquis par des capteurs différents. Une analyse quantitative est réalisée pour la qualité de l'extraction des routes ainsi que pour la qualité topologique du réseau extrait.

Zusammenfassung

In diesem Beitrag wird ein Algorithmus zur Extraktion von Straßennetzen in Vorstadtgebieten vorgestellt. Der Algorithmus ist regionenbasiert und nutzt hochaufgelöste Farbinfrarotbilder sowie, optional, ein DOM. Die Straßenextraktion beginnt mit einer Segmentierung mit Hilfe des Normalised-Cuts-Algorithmus; danach werden die Segmente gruppiert. Straßenstücke werden aus den gruppierten Segmenten extrahiert. Straßenstücke, die wahrscheinlich zu der gleichen Straße gehören, werden im nächsten Schritt zu Subgraphen verbunden. Um falsche Verbindungen in den Subgraphen zu entfernen, werden Kontextobjekte wie zum Beispiel Fahrzeuge, Gebäude und Bäume verwendet. Die verbleibenden Straßen, repräsentiert durch ihre Mittellinien, werden zu einem Straßennetz verbunden. Der gesamte Ablauf nutzt Kombinationen aus radiometrischen und geometrischen Merkmalen, die aus Wissen über das Erscheinungsbild von Straßen in Vorstadtgebieten abgeleitet wurden. Ergebnisse werden für zwei Testdatensätze gezeigt, die mit verschiedenen Sensoren erstellt wurden. Eine quantitative Analyse *bezogen auf die Qualität der Straßenextraktion sowie auf die topologische Qualität des extrahierten Netzwerks wurde durchgeführt.*

Resumen

En este artículo se describe un algoritmo utilizado para la extracción de redes viarias en áreas suburbanas. El algoritmo está basado en regiones y utiliza imágenes infrarrojas en color de alta resolución así como, de forma opcional, un modelo digital de superficies. El proceso de extracción comienza con una segmentación utilizando el algoritmo de corte normalizado, tras el que se agrupan los segmentos. Las distintas partes de las vías se extraen de los segmentos agrupados. Las partes que probablemente pertenecen a la misma vía se conectan en subgrafos en la siguiente fase. Para eliminar falsas conexiones en los subgrafos se emplean objetos contextuales tales como vehículos, edificaciones y árboles. Los fragmentos de vías residuales, representados por sus centroides, se conectan a una red de vías. El proceso emplea la combinación de elementos radiométricos y geométricos, obtenidos a partir del conocimiento del aspecto de las vías en áreas suburbanas. Los resultados descritos corresponden a dos conjuntos de datos obtenidos con diferentes sensores. Finalmente se llevó a cabo un análisis cuantitativo de la calidad de la extracción de la vía así como de la calidad topológica de la red extraída.