ROAD EXTRACTION IN SUBURBAN AREAS BASED ON NORMALIZED CUTS

A. Grote *, M. Butenuth, C. Heipke

Leibniz Universität Hannover, Institute of Photogrammetry and GeoInformation, 30167 Hannover, Germany - (grote, butenuth, heipke)@ipi.uni-hannover.de

KEY WORDS: Segmentation, Normalized Cuts, Road extraction, Urban areas, Grouping

ABSTRACT:

This paper deals with road extraction of high resolution aerial images of suburban scenes based on segmentation using the Normalized Cuts algorithm. The aim of our project is the extraction of roads for the assessment of a road database, however, this paper is restricted to road extraction. The segmentation as our basic step is designed to yield a good division between road areas and the surroundings. We use the Normalized Cuts algorithm, which is a graph-based approach that divides the image on the basis of pixel similarities. The definition of these similarities can incorporate several features, which is necessary for the segmentation in complex surroundings such as built-up areas. The features used for segmentation comprise colour, hue, edges and road colour derived with prior information about the position of the centerline from the database. The initial segments have to be grouped due to an enforced oversegmentation. The grouping is based on the criteria of mean colour difference, edge strength of the shared borders and colour standard deviation of merged initial segments. The grouped segments are then evaluated using shape criteria in order to extract road parts. Results on some test images show that the approach provides reliable road parts. Concluding remarks are given at the end to point out further investigations concerning the evaluation of the road segments and their use in database assessment.

1. INTRODUCTION

Roads are among the most important objects that are extracted from aerial images; they are necessary for many applications, for example navigation systems or spatial planning. Extracted roads are recorded in geospatial databases. As roads are subject to frequent changes, it is necessary to check road databases frequently to eliminate errors and to add new road objects. Manual database assessment is very time-consuming, which is why automatic database assessment is a major research topic (Zhang, 2004; Gerke, 2006). In many approaches, up-to-date aerial or satellite images are used to automatically extract objects and to compare them to objects contained in the database (Baltsavias, 2004).

Many approaches for road extraction have been developed; some of them are summarized in (Mayer et al., 2006). However, only few approaches work in urban or suburban areas due to the highly complex structure found in urban scenes which complicates the task of automatic road extraction.. In (Price, 1999; Youn and Bethel, 2004) the road network is expected to be a more or less regular grid but this constraint is not suitable for many European urban areas. Another approach uses a very sophisticated road and context model and is based on grouping small extracted entities to lanes, carriageways and road networks (Hinz, 2004). It employs a large set of parameters that must be carefully adapted for different scenes. In recent work, colour properties are exploited, for example in (Zhang and Couloigner, 2006): the authors perform a pixel-based multispectral classification and use shape descriptors to reduce the number of misclassifications. But they still only have a completeness and correctness rate of approximately 50 per cent. In our opinion, a reason for this is that the multispectral classification does not take into account the spatial relations of the pixels and that colour and shape properties are treated separately.

From the above mentioned works we can deduce that a proper segmentation algorithm is essential for the extraction of roads in suburban areas and that it is important to combine several features for the segmentation. A simple line based road model as used in many road extraction approaches for rural areas is not applicable. Consequently, this paper deals with road extraction in suburban scenes with a focus on segmentation. For segmentation we use Normalized Cuts, a graph-based method which is capable of combining several features to describe pixel similarity. After the segmentation, the initial segments are grouped to reverse oversegmentation. The result are segments that are further evaluated in order to find road segments.

In section 2 the Normalized Cuts algorithm is explained. Our approach for segmentation and grouping is described in section 3, results are presented in section 4. Some conclusions and an outlook on further work are given in section 5.

2. NORMALIZED CUTS

In this section, the Normalized Cuts method, which is used in our approach as starting point, is described in brief. Normalized Cuts is a graph-based method which is used to divide an undirected graph with weighted edges into segments with similar features. The method and its use in image segmentation are described in detail in (Shi and Malik, 2000). Pixels are defined as nodes and connected by weighted edges. The weights describe the similarity between the connected pixels. Theoretically, every pixel can be connected to all other pixels, but in practice only pixels in the vicinity of one pixel are connected with weights different from zero. The similarity

* Corresponding author.

measure is chosen according to the application, it is also possible to combine several similarity criteria. After defining the graph in this way, it is divided into segments aiming at a large dissimilarity between different segments and a large similarity inside each segment. This goal is achieved by cutting the graph such that it meets the following minimization condition:

$$Ncut(A_1,...,A_n) = \sum_{i=1}^{n} \frac{link(A_i, V \setminus A_i)}{link(A_i, V)} = \min \qquad (1)$$

The graph is divided into n sets of nodes A_i. V is the set that contains all nodes in the whole graph. *Link* is the sum of all weights of the connecting edges between two sets:

$$link(P,Q) = \sum_{p \in P, q \in Q} w(p,q) \qquad (2)$$

where $w(p,q)$ is the weight between two nodes p and q belonging to the two sets P and Q. The weight assigned to each pixel pair is inserted into a similarity matrix. The matrix is symmetric, its row and column dimensions are equal to the number of pixels. The minimization is obtained by computing the eigenvectors of a matrix derived from the similarity matrix. Multiple eigenvectors are calculated for multiple segments. The details of the calculation can be found in (Shi and Malik, 2000). The result is a set of discretised eigenvectors with the same number of elements as the number of pixels. The number of segments wanted has to be specified before the calculation because each eigenvector defines one segment.

One advantage of this method is the possibility to combine several different features in one step, a property that is important in complex surroundings. The difficulties that arise from the task of combining the results of several segmentation steps can be avoided in this way. Another advantage is that the algorithm takes both local and global characteristics into account. Local characteristics are incorporated in the similarity matrix which contains the weights of neighbouring pixels. In this way the similarity of pixels in a close neighbourhood is regarded. Global characteristics are considered when the optimal cut is calculated: a global minimum criterion must be met. This is a considerable advantage of the method, because in this way, small disturbances like short or weak edges are ignored by the algorithm.

3. APPROACH

As mentioned before, road extraction in suburban areas is more complicated than in rural areas due to the inhomogeneous background. We use an area based road model and apply our strategy to high resolution CIR images.

3.1 Road Model

The road model shown in Fig. 1 is adapted from (Hinz, 2004). A road segment consists of one or more lanes that are bordered by the roadsides or road markings, these borders are parallel. Other road markings can be found on the lanes, for example arrows or zebra crossings. The road surface appears in the image as a more or less homogeneous area while the roadsides appear as edges and the line markings as bright lines. These properties are used for the definition of the pixel similarities which are used in turn for the Normalized Cuts algorithm.

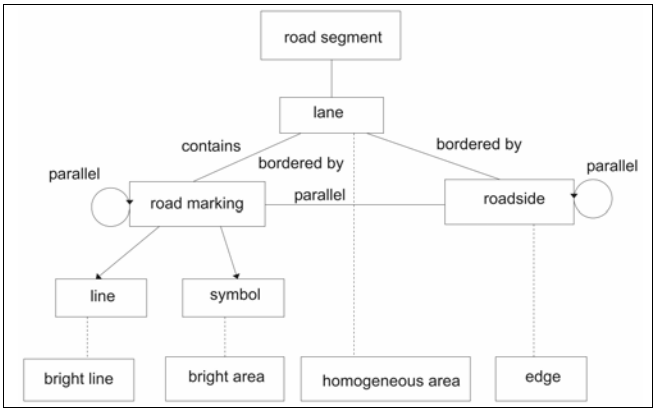

Figure 1. Road model, adapted from (Hinz, 2004).

3.2 General Strategy

The overall aim of our research is the assessment of a road database. Therefore, it is not necessary to analyse the whole image. The search space is restricted to a region of interest whose length matches the length of the road part to be assessed and whose width exceeds the expected road width in order to avoid forcing extracted segments into a road shape, thus distorting the evaluation.

The image is segmented using the Normalized Cuts method as described in detail in subsection 3.3. The result is an oversegmentation which is necessary to make sure that as many parts as possible of the road border belong to a segment border, even if the image information is weak. Therefore, the initial segments have to be grouped in a second step to coherent bigger segments before they can be evaluated and divided into road segments and non-road segments. It is not possible to evaluate the initial segments correctly based on their shape characteristics because they are too small for deriving reliable shape attributes. The grouping is currently done with a simple iterative algorithm merging initial segments with similar mean colour, weak edges at the shared border and low overall colour variance. After grouping, the segments are evaluated in order to extract road parts. The evaluation is mainly based on shape criteria, additionally the NDVI is used. The goal of the evaluation is the extraction of reliable road objects, keeping the number of false road objects to a minimum.

3.3 Segmentation

For the segmentation the Normalized Cuts method described in Section 2 is used. The region of interest is divided into small subsets of equal size and for each of the subsets the similarity matrix is set up individually because the similarity matrix is large even for smaller images and thus the computational resources needed become prohibitive for larger images. The aim of the segmentation is to separate the road parts of the image from the non-road parts. Each segment should contain only road pixels or non-road pixels, not a mixture of both. Therefore, the similarity criteria are derived from the road model. The similarity criteria are:

- Presence and strength of edges between two pixels

- Colour difference
- Hue difference
- Road colour derived from database information.

According to the model, roads are divided from their surroundings by edges, therefore edges are used as a criterion: if there is an edge between two pixels, the pixels are considered dissimilar. The edge criterion is incorporated into the similarity matrix in the following way: first, a Laplacian of Gaussian operator is applied to the image which yields an image of the second derivative in which edges are indicated by sign changes. For each pixel pair it is checked whether a sign change occurs along their connecting line or not. If a sign change occurs, the edge amplitude at this point is taken from an edge image calculated with the Canny operator. The edge amplitude is used to calculate the first part of the similarity measure, following the way described in (Shi and Malik, 2000):

$$w_{edge} = e^{\frac{-f^2}{2\sigma^2}} \tag{3}$$

f is the edge intensity and σ is ten per cent of the range of the edge intensity.

The second criterion is colour because roads usually have homogeneous surfaces and the pixel colour stays approximately the same. A measure for the colour similarity of two pixels is the distance between these two pixels in colour space. The colour similarity is calculated in the same way as the edge similarity in Equation 3. Here, f is the distance, and σ is defined as ten percent of the possible range of distance vector lengths.

These two similarity measures are multiplied to form a combined similarity measure:

$$w = w_{edge} \cdot w_{colour} \tag{4}$$

As a third criterion hue is used because a significant hue difference almost certainly indicates a different object. In parts darkened by shadows the hue of an object remains the same if certain conditions are met (Perez and Koch, 1994). Therefore, the weight is reduced by multiplying with a scale factor smaller than 1 (hue scale factor) if the two pixels have a hue difference that is greater than a defined threshold (hue threshold). A hard threshold is used because the weight should not be diminished at all if the hue difference is small but it should decrease significantly as soon as the difference exceeds the threshold, indicating two different objects. There is certainly some correlation between the colour and the hue criterion; however, our experiments have shown that the use of both yields the best results.

The database information is used to obtain colour information about the roads: assuming that the position of the road is approximately correct, we compute the average colour values of the road in the image for each channel from the position of the database road centerline. For every pair of pixels it is checked if both pixels lie inside an interval of the average values defined by the standard deviation. If both pixels lie inside or outside this interval in all channels, the weight is multiplied by a scale factor larger than 1 (road colour scale factor). The weight is set to 1 if it exceeds 1. By this means, it becomes more likely that segments are divided along a road border.

After the similarity matrix is set up using the defined similarity criteria, the Normalized Cuts algorithm is applied. The number of segments that has to be specified before the algorithm is started must be large enough to prevent merging of road and non-road segments.

3.4 Local Grouping

The image segmentation algorithm results in an oversegmentation. This is necessary in order to avoid the loss of any important road borders. The initial small segments are then grouped to bigger, more meaningful segments before being further evaluated. In the literature, there are only few examples dealing with grouping of image regions. One of them is (Luo and Guo, 2003): they aim at a general grouping algorithm as a bridge between image segmentation and high-level extraction algorithms. The region properties they use include, among others, the colour mean difference between two regions and the edge strength along a shared border. These two criteria are particularly suitable for our approach, because using them as merging criteria can reverse the enforced oversegmentation, which often produces segment borders at places where the image information does not justify a separation (see the results in section 4.1).

At present, we use a simple iterative approach for grouping the segments: in each iteration step two segments are merged, with regard to several criteria that are calculated for each pair of initial segments, partly based on the similarity criteria used by (Luo and Guo, 2003):

- Difference of mean colour (separate for the three bands)
- Edge strength of the intensity channel in the region around the shared border (border region)
- Joint standard deviation of colour of the regions if merged (separate for the three channels)

The border region is a seven pixel wide band along the shared border. For all criteria, the calculated values have to be below defined thresholds for the segments to be considered for merging. The thresholds are determined empirically. In each iteration step, only the two segments with the best values for all criteria (the least colour differences, the least edge strength and the least colour standard deviations) are merged. The iteration continues until the values for every segment pair exceed the thresholds.

3.5 Evaluation

The next step after grouping is the evaluation of the segments in order to extract road parts. The evaluation is based on shape and spectral characteristics of roads. The following characteristics are currently used for evaluation:

- Elongation
- Width
- Rectangularity
- NDVI

The elongation indicates the difference of the object from a circle. It is given by the ratio of the squared perimeter and the area of the segment. Road parts should have a high elongation and thus a high ratio. The width of a road part should not be much larger than the average width of a road. The rectangularity is a measure for the similarity of an object with a rectangle. It is calculated using the discrepancy method described in (Rosin, 1999). In this method the region of the object is compared to the region of the bounding rectangle of an ellipse with the same first- and second-order moments as the object region. The fourth criterion, the NDVI (Normalized Difference Vegetation Index), is employed because road parts should not contain vegetation. The average NDVI is calculated for each segment and for road parts the NDVI should be low.

In our tests, thresholds are defined empirically for each of these criteria and segments are extracted as road parts if they fulfil all of them. The thresholds are set rather strict in order to extract reliable road parts at the expense of missing some of them.

4. RESULTS

The approach was tested on CIR aerial orthoimages with a resolution of 0.1m. The images are from a suburban scene in Grangemouth, Scotland. Road database data were simulated by manually digitizing the visible roads in the images.

4.1 Segmentation

For segmentation with Normalized Cuts, the images are divided into subsets of approximately 200 x 200 pixels. Each subset is segmented by the Normalized Cuts algorithm yielding 20 segments, an empirical value that is suitable for this image size and scene complexity. The width of the region of interest is set to approximately three times the expected road width. The hue scale factor is set to 0.01, the hue threshold is set to 30. This value was derived from some manually taken sample objects. The road colour scale factor is set to 2.

Fig. 2 and Fig. 3 show examples of segmentation results obtained with the similarity criteria that are described in section 3.3. Segment borders are indicated by yellow lines; the green line shows the database road centerline. The results demonstrate that the segmentation in general has succeeded: road and non-road areas are in most instances clearly divided by initial segment borders. Exceptions can be found in shadow areas or where the contrast between the road and the surroundings is low, as in Fig. 3 in the right part of the image.

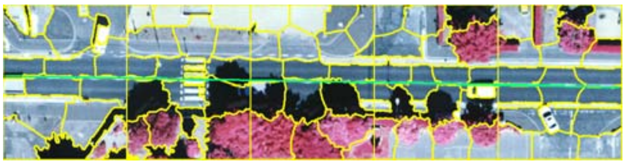

Figure 2. Segmented image, first example.

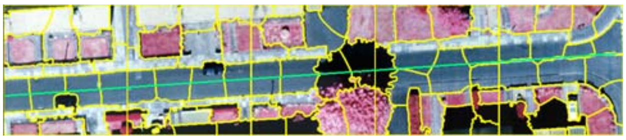

Figure 3. Segmented image, second example.

Fig. 4 shows an example where the database information is not used to obtain the road colour information as described in section 3.3. Consequently, the road colour is not considered in the segmentation. This example points out the benefit of using the prior information of the database: in Fig. 4 more road segments contain non-road areas than in Fig. 3.

Figure 4. Segmentation without using road colour information derived from the database road.

As the results are to be used for the assessment of the road database, one has to consider the case that the database information is not correct. Fig. 5 shows a segmentation result with a false database road. This example illustrates that incorporating the database information into the segmentation does not lead to wrong segmentation results: the initial segments are clearly defined by image content. Accordingly, the database information does not corrupt the results if it is wrong but can improve the segmentation results.

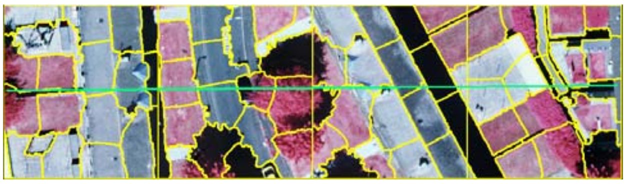

Figure 5. Segmented image with false database road.

4.2 Local Grouping

The initial segmentation results are now grouped. Fig. 6 shows the grouping result from the segments of Fig. 2, Fig. 7 shows the grouping result from the segments of Fig. 3, and Fig. 8 shows the results from Fig. 5.

The parameters for grouping are the same for all three images. The mean colour of two segments to be merged should be the same, the maximum for the mean edge strength is set to 50, and the maximum standard deviation for two merged images is set to 40.

In Fig. 6 most of the undisturbed road parts are merged into two bigger parts. These parts show some characteristics for road segments: their shape is elongated with parallel lines; their width corresponds to the road width. Parts that contain shadows, vehicles or salient road markings (zebra crossing) are not merged with the other road parts. Here, context knowledge is essential for further evaluation.

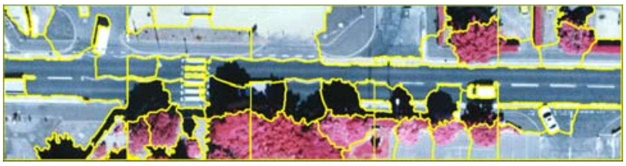

Figure 6. Grouping result, first example.

The second example, shown in Fig. 7, contains a road which is mostly undisturbed by context objects. The grouping result shows one distinctive road segment in the left part of the image. To the right, there is one big road segment that is more problematic: it contains one part of a parking lot and some parts of the pavement. The pavement parts are parts of an initial road segment and are not critical because they do not affect the overall shape of the road to a great extent. The parking lot poses more difficulties for an evaluation but such a case cannot be avoided because the parking lot has the same colour characteristics as the road and is not separated by a distinct border.

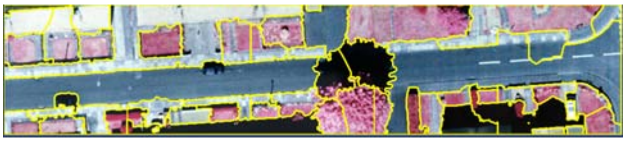

Figure 7. Grouping result, second example

Fig. 8 shows the grouping result for the example with the wrong database information. The segments here are again grouped to meaningful bigger segments. As there are no road-shaped segments along the direction of the database road, this road probably would be rejected in an evaluation step.

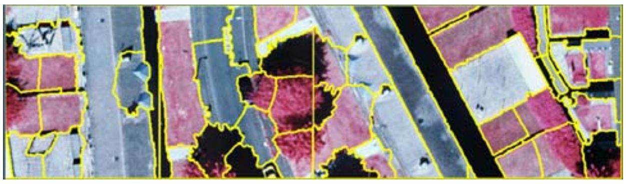

Figure 8. Grouping result, false database example

4.3 Evaluation

In the next step, the grouped segments are evaluated using the criteria described in section 3.5. The figures 9 and 10 show the experimental evaluation results for the first and second example. The thresholds used for the evaluation are: elongation more than 40, rectangularity more than 0.6, width no more than two meters above assumed road width, NDVI less than 0.

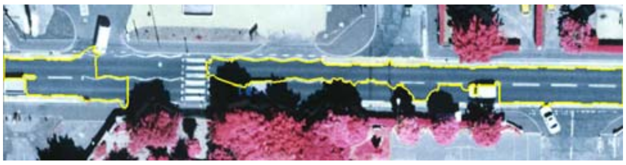

Figure 9. Evaluation result, first example.

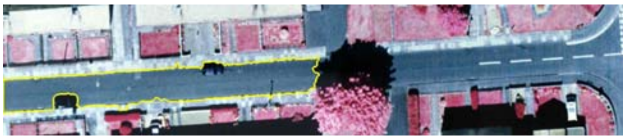

Figure 10. Evaluation result, second example.

In Fig 9, two segments were correctly extracted as road parts, in Fig. 10 one segment was found. No false road parts are extracted, which is important because the goal is a reliable rather than a complete extraction.

5. CONCLUSIONS AND OUTLOOK

The exemplary results in this paper show the general usability of the approach for the detection of roads in aerial images of suburban areas. The Normalized Cuts algorithm is suitable for the segmentation step. By considering global aspects of the image as well as local ones, the algorithm is able to ignore noise, small surface changes and weak edges, and borders are rather placed at continuing edges, as can be seen in the results. The division of the image depends on the overall image content which allows the segments to be more coherent and perceptually meaningful than segments obtained by a local segmentation only. We believe that this combination of local and global aspects is a very important characteristic of the Normalized Cuts algorithm.

One drawback of the Normalized Cuts algorithm is that the calculation is computationally expensive and the image has to be divided into smaller subsets to make the calculation possible with our current hardware. As a consequence, some segment borders are defined by the subset borders and not by image content. In many cases, this does not pose a problem because the segments are still merged in the grouping step, but not always. Another drawback is the fact that the number of segments has to be determined before starting the calculation. It is desirable to find a way to estimate the appropriate number of segments from the given data. One possibility is an iterative approach, repeating the Normalized Cuts algorithm with a varying number of segments and selecting the optimal segmentation. Possible criteria for a good segmentation are average segment size (not too small) and a satisfying homogeneity.

The grouping results show that it is possible to use the oversegmented results from the Normalized Cuts algorithm and group them to bigger segments whose shape can be assessed regarding their correspondence with the road model. The grouping works well for road parts without many disturbances by context objects. One problem are areas that are directly connected to the road and have the same colour as the road, like the parking lot in Fig. 8. Here, an additional grouping criterion, for example border continuation, could be helpful. The grouping algorithm itself, especially the combination of the different criteria, could also still be improved.

The first experimental evaluation results show that reliable road parts could be extracted, but there is still much room for improvement. For example, road parts should be close to rectangular in our current implementation of the evaluation. This can pose problems with long road parts that belong to curved roads. Therefore, we plan to change this criterion into one that requires an elongated object to have a constant width.

Our next steps will be the improvement of the grouping and evaluation steps, as indicated above. We will also investigate if the number of thresholds currently employed can be reduced and if the remaining thresholds can be estimated from the image data themselves. As the goal of our project is the assessment of a road database, we will use the extracted road parts for the assessment, adapting the strategy developed by (Gerke, 2006). In connection with the assessment, context objects like trees, vehicles, buildings and shadows will also be considered, in order to explain gaps between extracted road parts.

ACKNOWLEDGEMENTS

This project is funded by the DFG (German Research Foundation). The calculations were made with a C++ program partly adapted from a MATLAB program written by Timothée Cour, Stella Yu and Jianbo Shi. Their program can be found on http://www.seas.upenn.edu/~timothee/software_ncut/software.html (last checked July 2007).

REFERENCES

Baltsavias, E.P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. *ISPRS Journal of*

Photogrammetry and Remote Sensing, Vol. 58, No. 3-4, pp. 129-151.

Gerke, M., 2006. Automatic quality assessment of road databases using remotely sensed imagery. *Dissertation, Universität Hannover, DGK Reihe C, Nr. 599: Verlag der Bayerischen Akademie der Wissenschaften*, 105 p.

Hinz, S., 2004. Automatic road extraction in urban scenes – and beyond. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 35, part B3, pp. 349-355.

Mayer, H., Hinz, S., Bacher, U. and Baltsavias, E., 2006. A test of automatic road extraction approaches. In: *International Archives of Photogrammetry,Remote Sensing and Spatial Information Sciences*, Vol. 36, Part 3, pp. 209-214.

Luo, J. and Guo, C., 2003. Perceptual grouping of segmented regions in color images. *Pattern Recognition*, Vol. 36, No. 12, pp. 2781-2792.

Perez, F. and Koch, C., 1994. Toward color image segmentation in analog VLSI: algorithm and hardware. *International Journal of Computer Vision*, Vol. 12, No. 1, pp. 17-42.

Price, K., 1999. Road grid extraction and verification. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, part 3-2W5, pp. 101-106.

Rosin, P.L., 1999. Measuring rectangularity. *Machine Vision and Applications*, Vol. 11, No. 4, pp. 191-196.

Shi, J. and Malik, J., 2000. Normalized cuts and image segmentation. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22(8), pp. 888-905.

Youn, J. and Bethel, J.S., 2004. Adaptive snakes for urban road extraction. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 35, part B3, pp. 465-470.

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*, Vol. 58, No. 3-4, pp. 166-186.

Zhang, Q. and Couloigner, I, 2006. Automated road network extraction from high resolution multi-spectral imagery. In: *Proceedings of ASPRS Annual Conference, Reno, Nevada, May 2006*.