A MULTISCALE APPROACH FOR THE AUTOMATIC EXTRACTION OF TREE TOPS FROM REMOTE SENSING DATA

Bernd-M. Straub

Institute of Photogrammetry and GeoInformation, University of Hannover, Germany
(SOLVing3D, Schönebecker Allee, Garbsen, Germany)

Commission III, WG 4

KEY WORDS: Forestry, Urban, Vegetation, Automation, Scale, Model

ABSTRACT:
In this paper we describe our work on the automatic extraction of trees from aerial images and digital surface models for the refinement of 3D city models with information about trees. Three different aspects of this issue are discussed in the paper. First we give a motivation for the application of the automatic approach for tree extraction. An example is given by a visual comparison of a 3D city model with trees and the same one without trees. One can see – even in the screenshots - that the impression of quality becomes much better, if tree models are placed in the scene. In the second section a short overview on recent work regarding the extraction of trees is given. Furthermore, the common elements of the published algorithms are pointed out in this section. The strategy of our multiscale approach for the automatic extraction of tree tops from remote sensing data is introduced in the third section. The paper closes with a summary, some examplary results and an outlook on further work.

1. INTRODUCTION

Obviously, trees are important 3D objects in real as well as in virtual cities, not only for the orientation and recognition of the virtual city, but also for visibility computations which are performed for modern landmark based rout descriptions (refer to (Brenner & Elias 2003)). But, even if many people seems to be interested in vegetation for 3D city models (cf. (Fuchs et al. 1998)) relatively few work was published on the automation of the extraction and integration of trees.

From the viewpoint of data capturing there are some obstacles for the extraction and integration of trees: The first obstacle comes from the conflict, that the best time for image acquisition for the extraction of vegetation is during the vegetation period. Which is again not optimal (or necessary) for the measurement of buildings and roads. The only argument against is, that one *could* extract buildings and roads from images which were recorded during the vegetation period, even if there are some problems due to exclusions and things like this. But the extraction of vegetation parameters is often *not possible* from images, which are captured outside the vegetation period.

The second drawback seems to be a little bit outdated, but it should be mentioned here: For the extraction of vegetation it is very helpful to have the optical information in the infrared band. Formerly – in the period of analogue aerial cameras - one had to decide between colour infrared (CIR) film and normal colour film. And often it was decided to capture real colour images, because the most users are more familiar with this type of imagery for interpretation and due to the advantages for visualisation. Nowadays, up-to-date digital cameras allow us to capture both types of images in one flight. The CIR images can be used for the extraction of vegetation parameters and the normal colour ones for the production of orthoimages for visualization. And as a result, this drawback is no more really relevant.

The third practical obstacle for the development of automatic approaches is perhaps the most important one. A tree can be measured interactively by a human interpreter with only one or two mouse clicks. The first one defines the tree top and the second one defines the radius of the tree. That means, even in the case that an automat would be able to detect and measure really *every* tree in the scene, only these two clicks per object can be saved. And in fact, that is not really much of one compares it with the effort for a single building or a road.

In the next short section we will give a motivation that it can make sense to spent some of effort for saving these two clicks. The third section of this paper gives an overview on recent work in the domain of "tree counting" and leads over to some relevant differences between the extraction of individual trees in forest and in urban areas. A short description of the multi-scale approach for the automatic extraction of tree tops is given afterwards. In the last section the results of a performance evaluation are mentioned, an overview is given about the scene, which is presented at the beginning of the paper. Finally some further work is proposed.

2. MOTIVATION

Obviously, it would be very helpful if an automatic algorithm would be able to detect reliable more than 99.5% percent of the visible trees in a given scene. A service provider having this algorithm would be able to provide his client with an additional nice-to-have doodad without additional costs on both sides. Automatically extracted trees as a marketing instrument, why not?

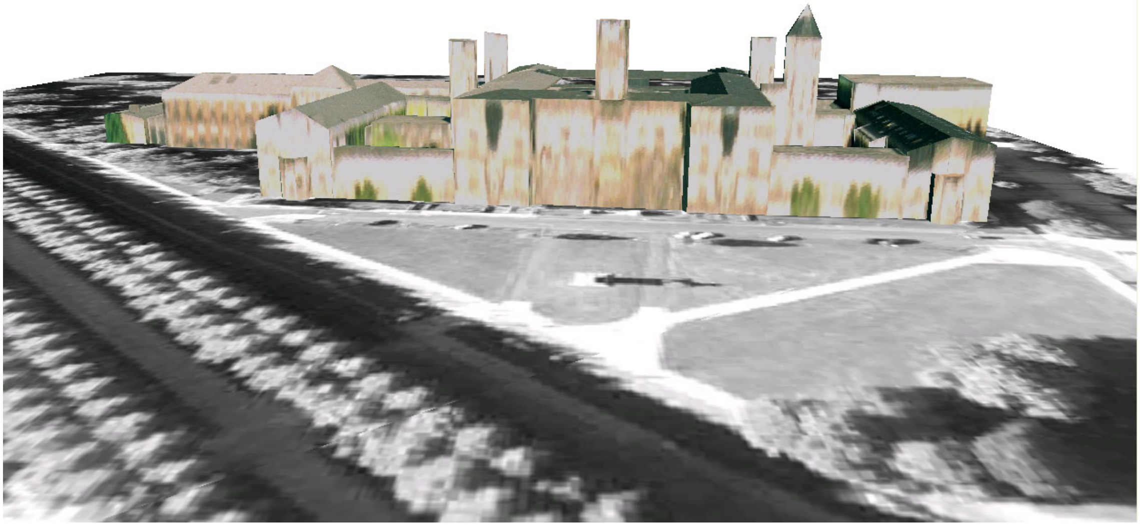

Figure 1: 3D model of the “Welfenschloß” the main building of the University of Hannover “standing” on an aerial image.

Figure 2: 3D model of the “Welfenschloß” with automatically extracted trees. The map-like image used as ground texture shows the result of an edge extractor (A Deriche operator (Deriche 1990) was applied to the aerial image, which is used as ground in the figure above. The tree template which was used for the VRML model was originally described in (Saint John 1997).)

In this section we want to show, that – even if the applied automatic algorithm works not perfect – it could make sense to have such an automatic algorithm for the extraction of trees.

The figures on the top of this page should give an impression of the impact of 3D tree models in the virtual city model. Figure 1 shows a 3D model of the “Welfenschloß”, the main building of the University of Hannover. (The reconstruction of the buildings in the city was performed manually with the inJECT software of INPHO GmbH.) The model was set onto an orthoimage, in order to give a more or less realistic impression of the scenario. With exactly the *same amount of manual work* one can produce the model which is predicted in Figure 2. The map-like texture on the ground was computed with a standard edge extractor without any manual support. The 3D tree models in the scene are the result of the approach for automatic extraction of trees from aerial images, which is explained in the following section of this paper.

The situation with the automatic approaches for the extraction of topographic objects is, that 99.5% are not achieved in the moment. Even a success rate 95% is not realistic, except in simple scenes. One of the reasons is, that one has to simplify the real world to a model, which in fact cannot cover all possibilities. And the more complicated the model, the more complex becomes the tuning of the – necessarily - increasing number of parameters. As a result, the success rate of the most automatic systems is – roughly speaking - between 70% and 90%. Significantly lower numbers point at a very complex situation or a very poor approach, and significantly higher numbers would raise suspicion, that the results of the automatic approach are a little bit sugar-coated. This holds also for the automatic extraction of trees. Let us expect a success rate of 66% for an approach which automatically extracts trees and "plants" them into the virtual city, without additional costs for the service provider. This is exactly, what we see in Figure 2. This is an interesting give-away for class A customers, or not?

In other words: As a potential customer of a 3D city model, who can select between the one product with 66% of the trees and another one without any trees in it. Both for the same amount of costs, which one would you choose?

3. STATE OF THE ART

The first trial to utilize an aerial image for forest purposes was performed in 1897 (Hildebrandt 1987). Since that time the scientific forest community is working on methods for the extraction of tree parameters from aerial images. Early work was carried out on the manual interpretation of images for forest inventory (Schneider 1974), (Lillesand & Kiefer 1994). The pioneers in the field of the automation of the interpretation task "extraction of individual trees from images" proposed first approaches about one and a half decade ago (Haenel & Eckstein 1986), (Gougeon & Moore 1988), (Pinz 1989). Recent work in the field was published in (Pollock 1996), (Brandtberg & Walter 1998), (Larsen 1999), (Andersen et al. 2002), (Persson et al. 2002), (Schardt et al. 2002).

A in depth state of the art overview regarding the automatic extraction of trees is given in (Straub 2003a). There are mainly two common elements in the most approaches: The first one is the use of a rotationally symmetric geometric model of a tree, as it was proposed by R.J. Pollock in (Pollock 1994). A three dimensional surface which simplifies the shape of the crown to an ellipsoid of revolution (assigned as Pollock-Model in the following). The surface of a real tree is of course very noisy in comparison to this simplification. This "noise" is not caused by the measurement of the surface, it is simply a consequence of the simplification for a very complex shape like the real crown of a tree. The idea is, that the coarse shape of the crown is well modelled with such a surface description. This leads over to the next common element of the most approaches, the use of some kind of low pass filtering in order to get rid of the "noisy" fine structures. Most authors propose to apply – with good reasons - a Gaussian function as low pass filter in this early processing stage, refer to (Dralle & Rudemo 1996), (Brandtberg & Walter 1998), (Schardt et al. 2002), (Straub 2003b), and (Persson et al. 2002).

Some work with focus on the automatic extraction of trees in urban areas was also published. In (Haala & Brenner 1999) it was proposed to use node points of the region skeletons of groups of trees as hypothesis for trees. Morphological processing of automatically extracted tree groups is also used in (Straub & Heipke 2001) for the computation of tree hypotheses. Local maxima of the digital surface model are used in (Vosselman 2003) for the detection of trees. The proposed solutions are constrained to elongated regions with trees (Haala & Brenner 1999), (Straub & Heipke 2001), or less complex scenes (Vosselman 2003). But, not all the trees in urban environments are standing in rows along roads or lines of buildings. In many cases they occur in compact arrangements, which are not fare away from forest scenes (refer to Figure 3).

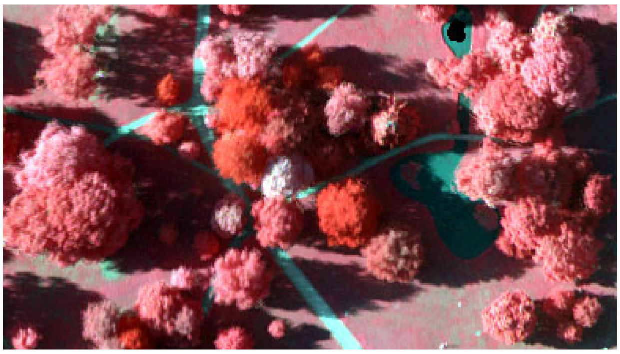

Figure 3: An "urban forest" inside of Hannover close to the University.

This was the motivation to develop a process for the automatic extraction of trees in urban environments, which should fulfil the following pre-conditions: It should be able to handle trees in *different local context*, i.e. as far as possible it should be the same algorithm for the situations single tree, row of trees, compact group of trees. Another important aspect for the extraction of trees in urban environment in contrast to a forest is, that smaller trees are not covered by the bigger ones. As the diameter of a tree can vary from two meter up to fifteen meters (cf. (Gong et al. 2002)), and in urban environment, small and big trees often stand close together. Therefore it is necessary to perform some kind of *mechanism for the selection of the locally best (or optimal) scale* for the extraction of the low level features. The scale selection is also a problem in forest areas: In (Schardt et al. 2002) it was proposed to use the scale selection mechanism proposed in (Lindeberg 1994a), which based on the maximum response after Scale-Space transformation. In our approach the scale selection is applied on a higher semantic level, i.e. after the segmentation of the image, and not before as it was proposed in (Schardt et al. 2002). This allows an internal evaluation of the segments on this semantic level, which is particularly then important if it is necessary to distinguish between trees and other objects (as well as in urban environments).

4. STRATEGY OF OUR APPROACH

In principal, there are two possibilities to build a strategy for the automatic extraction of trees from raster data. The first possibility is to model the crown in detail: one could try to detect and group the fine structures in order to reconstruct the individual crowns. The second possibility is to remove the fine structures from the data with the aim to create a surface which has the character of the Pollock-Model. In the literature examples for both strategies can be found: In (Brandtberg 1999) it was proposed to use the typical fine structure of deciduous trees in optical images for the detection of individual trees. In (Andersen et al. 2002) the fine structure of the crown is modelled as a stochastic process with the aim to detect the underlying coarse structure of the crown.

The other strategy, the removal of noise, was proposed by (Schardt et al. 2002), (Persson et al. 2002), and (Straub 2003a). The main problem of the second type of approach is the determination of an optimal low pass filter - which is of crucial importance for the segmentation - for every single tree in the image. It is kind of a chicken-and-egg problem: the optimal low pass filter depends mainly on the diameter of the individual tree one is looking for, which is not known in advance. In the case of trees this size can neither be assumed to be known nor is it constant for all trees in one image. The size of trees depends on the age, the habitat, the species and many more parameters, which cannot be modelled in advance.

Process of object extraction from images and/or surface models generally depends on an object model as well as a strategy for extraction of image features, their combination, and their relation to the model. A generic geometric model of a tree is used which basically consists of a function describing the tree top. Based on this model features are identified, which are used to recognise single tree tops from the image data. The basic idea for this strategy consists of two steps (cf. Figure 4). At first, the often very complex fine structures are removed from the surface model by using multiple scale levels in linear scale space. As a result of scale-space transformation the tree top can be identified in the surface model based on the coarse structure. Here, the main problem is, that on the one hand the diameter of a single tree continuously varies in reality, but also strongly influences the choice of filter parameters. To overcome this difficulty, the image data was examined at different scale levels.

The basis idea of our approach is to use a multi-scale representation of the surface model (assigned as H_{sigma} in Figure 4) and of the orthoimage (assigned as I_{sigma} in Figure 4) in order to reduce get rid of the fine structures of the tree crown, similar to the proposal described in (Persson et al. 2002). Whereas $sigma$ is the parameter of the Gaussian, which is used to create the multi-scale representation (refer to (Lindeberg 1994b) for details on Linear Scale-Space transformation).

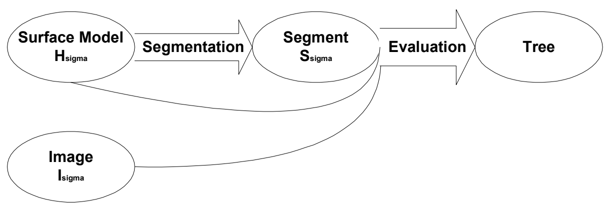

Figure 4: Strategy for the automatic extraction of trees

A Watershed transformation is used as segmentation algorithm, leading to the segments S_{sigma}. Every S_{sigma} is a hypothesis for a tree (see Figure 5, for an example). The evaluation of the segments is performed according to fuzzy membership values. A tree is an object with a defined size, circularity, convexity and vitality (NDVI value).

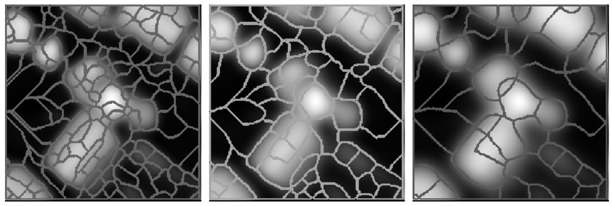

Figure 5: Segmentation results in three different scale levels, left fine scale, right coarse scale

The evaluation phase is divided in two independent steps: First, the hypotheses for trees are selected regarding their membership values (refer to Figure 6). Than, in the second step, the best hypothesis in scale-space is selected. As at one and the same spatial position in the scene, more than one valid hypothesis can exist, the best one – considering the membership value – is selected (refer to the marked segments in Figure 6).

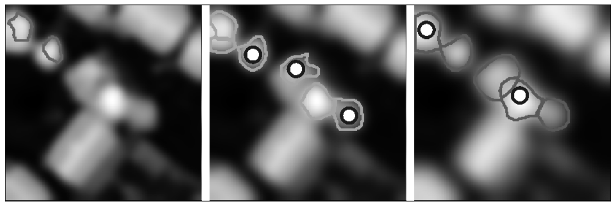

Figure 6: Valid hypotheses for trees in different scale levels, depicted are the borderlines in different grey values. Best hypotheses are marked with a white circle.

A detailed description of the approach is given in (Straub 2003c) and (Straub 2003a).

5. SUMMARY

An approach for the automatic extraction of trees from remote sensing data - aerial imagery and surface models – was shortly depicted in this paper. A detailed description of the most important considerations, leading to the development of the approach, is given: for the model of an individual tree, which is the base of the approach and for the strategy for low-level feature extraction and generation of hypotheses.

Recently, the approach was applied on different data sets. Results of a performance evaluation of the approach are presented in (Straub 2003d) (and, more detailed in (Straub 2003a)). The test was carried out with one and the same parameter settings for all data sets in order to demonstrate its robustness and the stability of the underlying model and strategy. The Hanover example (c.f. Figure 2 and Figure 7) was produced using image and height data from Toposys Falcon system, which were acquired by Toposys GmbH in summer 2003 by order of the institute of cartography and geoinformatics (University of Hannover). An overview of the results is given in Figure 7, the automatically extracted trees are printed are depicted as white circles:

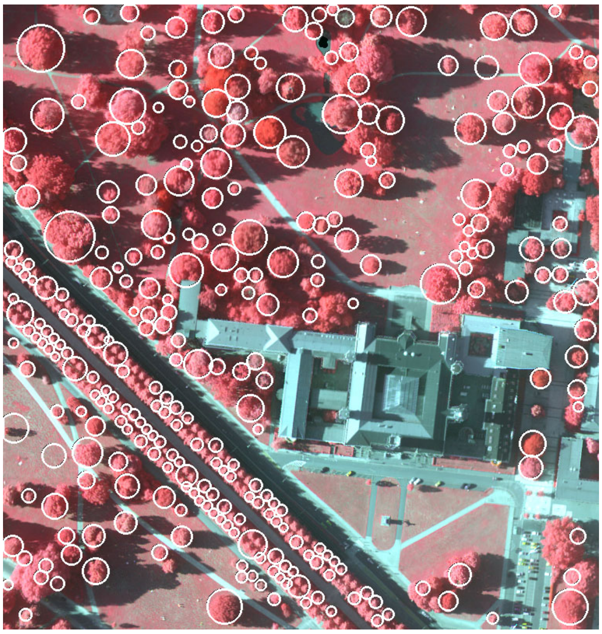

Figure 7: Results of the automatic approach for extraction of trees. Trees are depicted as white circles, the background image is a part of the CIR orthoimage. The subset is equivalent to the scene in Figure 2.

For the 3D visualisation in Figure 2 the diameter of the trees were used to scale the whole 3D model of the tree. Further work will be on a refinement of the visualisation: The shape and the colour of the extracted trees shall be used to define different classes of trees models in order to get a more realistic visualisation. As the algorithm was originally designed for the extraction of trees from image data (and the surface model, which can be computed from image using correlation techniques), the last pulse data of the laser scanner is not used at the moment. This additional information will be used to make the extraction more reliable and complete.

ACKNOWLEDGEMENT

The author would like to thank Claus Brenner, who has sponsored the laser scan from his project at the institute of cartography and geoinformatics. The buildings of the Hanover 3D city model were reconstructed by Adele Elmhorst and Ulla Wissmann. Only with their work it was possible to make this paper as it is: with the example including buildings. Many thanks go also to the students and tutors of the Seminar-3D-Team for their work in the recent project seminar.

REFERENCES

Andersen, H., Reutebuch, S. E., Schreuder, G. F., 2002. Bayesian Object Recognition for the Analysis of Complex Forest Scenes in Airborne Laser Scanner Data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (eds.) Kalliany Leberl, ISPRS, Graz, Austria, Vol. XXXIV, Nr. WG 3A, pp. 35-41.

Brandtberg, T., 1999. Automatic Individual Tree-Based Analysis of High Spatial Resolution Remotely Sensed Data. *Dissertation*, Centre for Image Analysis, Swedisch University of Agricultural Sciences, Uppsala, Nr. Silvestria 118, pp. 47.

Brandtberg, T., Walter, F., 1998. Automated Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images by Multiple Scale Analysis. *Machine Vision and Applications*, 11 (1998), pp. 64-73.

Brenner, C., Elias, B., 2003. Extracting Landmarks for Car Navigation Systems Using Existing GIS Databases and Laser Scanning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (eds.) Ebner Heipke Mayer Pakzad, ISPRS, München, Germany, Vol. XXXIV, Part 3/W8,

Deriche, R., 1990. Fast Algorithms For Low-Level Vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , 1 (12), pp. 78-87.

Dralle, K., Rudemo, M., 1996. Stem Number Estimation by Kernel Smoothing of Aerial Photos. *Canadian Journal of Forest Research*, 26 (1996), pp. 1228-1236.

Fuchs, C., Gülch, E., Förstner, W., 1998. OEEPE Survey on 3D-City Models. *OEEPE Publication*, Bundesamt für Kartographie und Geodäsie, Frankfurt, Germany, Nr. 35, pp. 9-123.

Gong, P., Sheng, Y., Biging, G., 2002. 3D Model Based Tree Measurement from High-Resolution Aerial Imagery. *Photogrammetric Engineerung & Remote Sensing*, 11 (68), pp. 1203-1212.

Gougeon, F., Moore, T., 1988. Individual Tree Classification Using Meis-II Imagery. *IGARSS '88 Geoscience and Remote Sensing Symposium*, IEEE, Vol. 2, pp. 927 -927.

Haala, N., Brenner, C., 1999. Extraction of Buildings and Trees in Urban Environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2-3 (54), pp. 130-137.

Haenel, S., Eckstein, W., 1986. Ein Arbeitsplatz zur automatischen Luftbildanalyse. *Mustererkennung 1986*, Springer, Berlin, Deutschland, pp. 38-42.

Hildebrandt, G., 1987. 100 Jahre forstliche Luftbildaufnahme - Zwei Dokumente aus den Anfängen der forstlichen Luftbildinterpretation. *Bildmessung und Luftbildwesen*, 55 (1987), pp. 221-224.

Larsen, M., 1999. Individual tree top position estimation by template voting. *21. Canadian Symposium on Remote Sensing*, Ottawa, Canada, 21-24 June, pp. 8.

Lillesand, T. M., Kiefer, R. W., 1994. *Remote Sensing and Image Interpretation*. John Wiley & Sons, New York Chichester Brisbane Toronto Singapore, pp. 750.

Lindeberg, T., 1994. Scale-Space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 2 (21), pp. 225-270. (b)

Lindeberg, T., 1994. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, Boston, USA, pp. 423. (a)

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and Measuring Individual Trees Using an Airborne Laser

Scanner. *Photogrammetric Engineering & Remote Sensing*, 9 (68), pp. 925-932.

Pinz, A., 1989. Final Results of the Vision Expert System VES: Finding Trees in Aerial Photographs. *Proceedings ÖAGM 13. Workshop of the Austrian Association for Pattern Recognition*, Oldenbourg Schriftenreihe Österreichische Computer Gesellschaft , Wien München, pp. 90-111.

Pollock, R. J., 1994. A model-based approach to automatically locating tree crowns in high spatial resolution images. *Image and Signal Processing for Remote Sensing*, (eds.) Desachy, SPIE, Vol. 2315, pp. 526-537.

Pollock, R. J., 1996. The Automatic recognition of Individual trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model. *Dissertation Computer Science*, The University of British Columbia, Vancouver, Canada, June 1996, pp. 170.

Saint John, R. W., 1997. Building the Perfect Tree A Lesson in Optimization an PROTO. http://www.vrmlsite.com/jun97/a.cgi/spot2.html, 3-20-2003,

Schardt, M., Ziegler, M., Wimmer, A., Wack, R., Hyyppä, R., 2002. Assessment of Forest Parameter by Means of Laser Scanning. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (eds.) Kalliany Leberl, ISPRS, Graz, Austria, Vol. XXXIV, Nr. 3A, pp. 302-309.

Schneider, S., 1974. *Luftbild und Luftbildinterpretation*. de Gruyter, Berlin New York, pp. 530.

Straub, B., 2003. A Top-Down Operator for the Automatic Extraction of Trees - Concept and Performance Evaluation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (eds.) H.-G. Maas G. Vosselmann A. Streilein, ISPRS, Dresden, Germany, Vol. WG III/4, pp. 34-39. (d)

Straub, B., 2003. Automatic Extraction of Trees from Aerial Images and Surface Models. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (eds.) Ebner Heipke Mayer Pakzad, ISPRS, München, Germany, September 17-19, Vol. XXXIV, Part 3/W8, pp. 157-164. (c)

Straub, B., 2003. Automatic Extraction of Trees from Height Data using Scale Space and Snakes. *Second International Precision Forestry Symposium*, (eds.) G. Schroeder, Precision Forestry Cooperative, University of Washington, Seattle, USA , 15.-18. Juni, pp. 75-84. (b)

Straub, B., 2003. Automatische Extraktion von Bäumen aus Fernerkundungsdaten. *Reihe C*, Deutsche Geodätische Kommission, München, Vol. 572, pp. 99. (a)

Straub, B., Heipke, C., 2001. Automatic Extraction of Trees for 3D-City Models from Images and Height Data. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Vol. 3. ed. Baltsavias Gruen van Gool, A.A.Balkema Publishers. Lisse Abingdon Exton(PA) Tokio. pp. 267-277.

Vosselman, G., 2003. 3D Reconstruction of Roads and Trees for City Modelling. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (eds.) H.-G. Maas G. Vosselmann A. Streilein, ISPRS, Dresden, Germany, Vol. WG III/4, pp. 231-236.