AUTOMATIC EXTRACTION OF TREES IN URBAN AREAS FROM AERIAL IMAGERY

Uwe Bacher[1], Helmut Mayer[2]

[1] Tiefenbach GmbH, Munich Branch, Telematic Systems
Papinstraße 53, D-81249 Munich, Germany
uwe.bacher@gmx.de

[2] Institute for Photogrammetry and Cartography
University of the Federal Armed Forces Munich, D-85577 Neubiberg, Germany
Helmut.Mayer@UniBw-Muenchen.de
http://www.BauV.UniBw-Muenchen.de/institute/inst10

KEY WORDS: Automatic tree extraction, tree model, Hough transform, hysteresis thresholding

ABSTRACT

In this paper we propose an approach for the automatic extraction of leafless deciduous trees from high resolution, i.e., ca. 4 cm ground resolution, aerial imagery taken in spring. In analogy to approaches for building extraction, we make use of the dark shadow of the tree as well as of the fact that the vertical trunk is imaged as a straight line pointing to the nadir point. Hypotheses for the trunk are found via the Hough transform. Branches are tracked by means of hysteresis thresholding. With our approach, it is possible to determine the trunk base, height, and width of the tree. Results show the potential of the approach.

1 INTRODUCTION

The modeling of trees is a great challenge because of their complex structure and their interaction with other trees as well as with other objects, such as buildings and roads. We focus on single trees and groups of trees or trees in a straight line, but not on trees in forests. One reason for this limitation is that this kind of trees is of specific interest for cadastre and three dimensional (3D) urban modeling. Furthermore, there is only few work about the automatic extraction of single trees or groups of trees. Most work deals with forest, e.g., (Pollock, 1996, Gougeon, 1995, Krummheuer, 1998).

The two most state-of-the-art approaches are described more in-detail. In (Larsen, 1998), trees are modeled by a generalized rotational ellipsoid and a distribution function for the density of the branches and leaves. This model is used to generate templates of the trees, taking the geometry of the sensor, the time of data capture, the illumination, as well as the reflectance of the tree and the ground into account. The templates are varied with respect to sizes and shapes of the trees and are matched with the imagery, in this case a line scanner with 30 cm ground resolution. The disadvantage of the approach is that it works well only for imagery containing mostly trees, i.e., forest. Moreover, many templates are necessary if the variability of trees is high. The approach proposed in (Brandtberg, 1996) is based on color-infrared imagery with a ground pixel size of about 3 cm. The approach uses information about the structure to classify the tree. Second order derivatives result into lines which are used to describe the branching structure. From the distribution of the orientations of the lines, a parallel structure can be distinguished from a radial structure. Herewith it is possible to classify three types of trees: pine, spruce, and birch. This approach works, however, only in case the regions for the tree crowns are given.

Photogrammetric data acquisition for geographic information systems (GIS) of urban areas is in most cases based on high resolution aerial imagery. Nowadays, colored images taken primarily in spring are used. From this follows that deciduous trees, which make up the majority of the trees in urban areas, are leafless. Moreover, conifers are rarely found in public areas, and are therefore not of first interest for urban planning and tree information systems.

A model for a leafless deciduous tree, which is represented both directly and as its shadow in the imagery, is proposed in Section 2. This gives way to the strategy described in Section 3. Section 4 demonstrates the feasibility of the proposed approach. Finally, conclusions are given.

2 MODEL

Our model of a leafless deciduous tree in an aerial image with high resolution, i.e., 2-5 cm ground pixel size, is presented as semantic network in Figure 1. The basic ideas for the modeling are similar to those used for building extraction,

namely the utilization of dark shadows in sun direction[1], and the exploitation of vertical structures, i.e., especially the trunks, which are projected to nadir pointing straight lines (Jaynes et al., 1994, Lin and Nevatia, 1998, Shufelt, 1996). Additionally, we assume that the deciduous tree is rotational symmetric. Thus, its 3D outline can be derived from its normalized and symmetric shadow projection. Please note that conifers would require a totally different modeling. The main reason for this is that they never loose their needles resulting into a totally different appearance in the image.
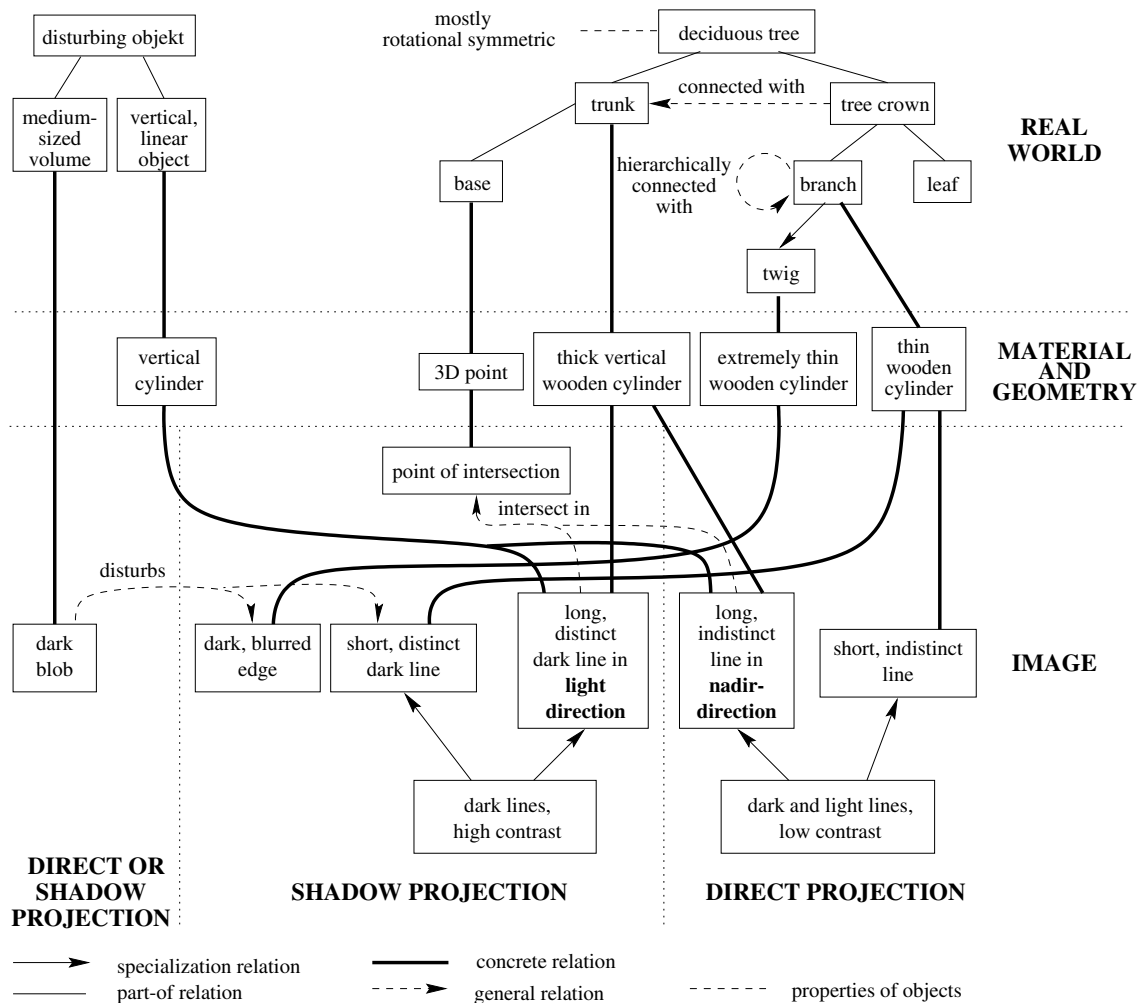

Figure 1: Model of leafless deciduous tree and disturbing object

Figure 1 represents the deciduous tree from three different points of view, i.e., levels, which are connected by the so-called concrete link (Tönjes, 1997, Mayer, 1998): In the *real world*, the nearly rotational symmetric tree consists of its trunk and crown. The base of the trunk is a part of the trunk and determines the position of the tree on the ground. The tree crown is connected to the trunk and consists of branches. The branches are connected hierarchically with decreasing thickness, ending in twigs.

The *material and geometry* level describes the objects independently of the sensor. The trunk is represented as thick vertical, the branch as thin and mostly curved, and the twig as extremely thin wooden cylinder with a bark surface. The base of the trunk is a 3D point whose coordinates can be derived from two or more images or an existing digital terrain model (DTM).

The objects at the *material and geometry* level are connected with the image primitives at the *image* level. On this level, the deciduous tree is represented on one hand indirectly by its shadow projection and on the other hand by its direct projection. In shadow projection, the more or less vertical trunk appears as long distinct dark line in sun direction. The branches correspond to short dark lines of high contrast. The outermost thin twigs of the tree crown are not resolved at this resolution as lines, but appear as a blurred edge. In direct projection, the branches are short lines and the trunk is a long line in nadir direction. The lines are dark or light and mostly of low contrast.

---

[1] I.e., projection of a vertical line on the horizontal ground, where the sun is the projection center

To make the extraction robust, we model also disturbing objects. In the *real world*, they correspond to medium-sized volumes, e.g., cars, or they are vertical linear objects, such as poles and street-lamps. The latter can be modeled at the *material and geometry* level as vertical cylinders, which then appear in the shadow projection at the *image* level as long distinct dark lines. They are like the trunks directed in sun direction and can therefore be mixed up with them. The medium-sized volumes often appear as dark blobs disturbing the short dark lines of the branches and the blurred edge of the tree crown in shadow projection. To keep the model simple and compact, we decided not to integrate occlusions and shadows cast by buildings, though they are necessary in principle.

3   STRATEGY

The strategy on top of the model is based on "hypothesize and verify". The focus of attention is laid on objects with high contrast to make the extraction as productive as possible. We start with distinct dark lines in sun direction which are hypotheses for trunks. After the trunk is verified by integrating more and more evidence, we continue by tracking the branches, determination of the outline of the tree, and its classification.

1. **Extraction and Verification of Hypotheses for Trunks**

   First, distinct dark lines representing the shadow projections of the trunks are extracted. Knowledge about existing objects, e.g., roads or buildings, may help to limit the search space geometrically, speeding up the extraction and making it more robust. Since the shadow of the trunk is oriented in sun direction, only these lines are selected and combined to form hypotheses for trunks. The direct projection of the tree is used as evidence to verify a hypothesis by extracting lines in nadir direction. By intersecting the shadow of the trunk with its direct projection, the position of the base of a trunk can be determined more accurately and robustly. If the base is occluded, it might be recovered using only this second kind of information.

2. **Tracking of Branches and Determination of the Outline of the Tree**

   Starting from a hypothesis for a shadow projection of a trunk, the dark lines representing the branches in shadow projection are extracted. It can be seen as a further verification of the hypothesis for the trunk and it is based on the knowledge that the branches start at the trunk, branch out and become thiner and thiner. The latter can be used to separate trees with overlapping shadow projections. What is more, it can also be employed to separate trees from vertical disturbing objects. For this, the structure of the surrounding branches should be investigated.

   By using the end points of the branches together with the fact that the twigs form a blurred edge and the assumption that the shape of the tree is more or less compact and symmetric, the outline of the tree can be reconstructed using a snake-based approach (Kass et al., 1987). The diameter of the tree crown can be derived directly from the outline. The height can be computed from the known sun direction and the shadow outline projected onto a DTM. Based on the base of the extracted trunk and a DTM, the 3D position of the base of the trunk can be determined.

3. **Derivation of the Type of the Tree**

   The type of the tree is derived by classification based on the shape of the outline, the ratio of the width of the tree crown and the height of the tree, and the structure of the branches. Because of the correlation between height and age of the tree, one can determine the approximate age of a tree when its type is known (Baumann, 1957, Spurr, 1960, Hildebrandt, 1996).

The proposed strategy has been outlined for a single gray-scale image. We do not use color information, because the shadow has no color and the trunk is only weakly colored. In principle, $n \geq 2$ images might be used to find corresponding structures of the branches. We have not done this due to the following reasons: Matching the shadow projection results in a refined DTM only. In suburban areas roads and other areas can in most cases be represented by plane surfaces sufficiently accurately. By matching the direct projection of the branches in two or more images, the 3D branching structure could be reconstructed. However, there are not only large problems in line extraction, but above all, the complexity of matching these structures is extremely high. The reason for this is that because of the complex distribution of the branches in space, their order can change totally in different images. Contrary to this, the information for our application can be derived which much less effort from the shadow projection of the tree.

4   INVESTIGATIONS

Figure 2 shows a part of an image showing a suburban area of the German town Tamm near Stuttgart. The original color image was scanned in black and white at a pixel size of 14 μm corresponding to a pixel size of 4.2 cm on the ground at an image scale of 1:3000. The image shows a suburban area comprising buildings, roads, and trees. The image was taken on March 23, 1995, at about 11 o'clock in the morning. In early spring the deciduous trees are leafless. The shadows of the trees are mostly cast on roads and clearly show the structure of the branches. Figure 3a) is a part of Figure 2.
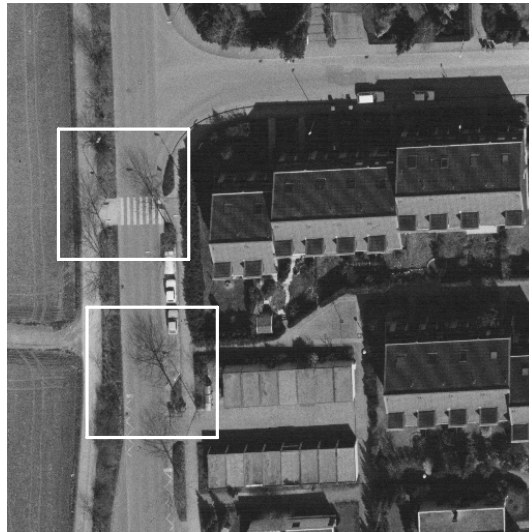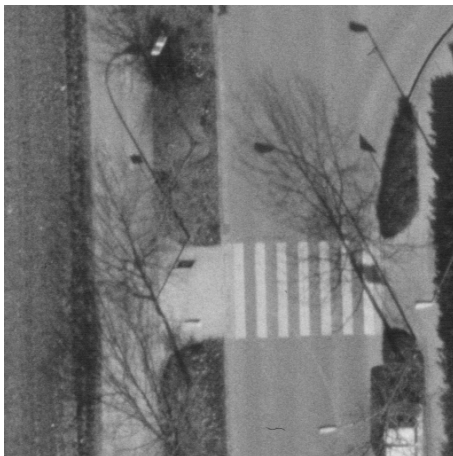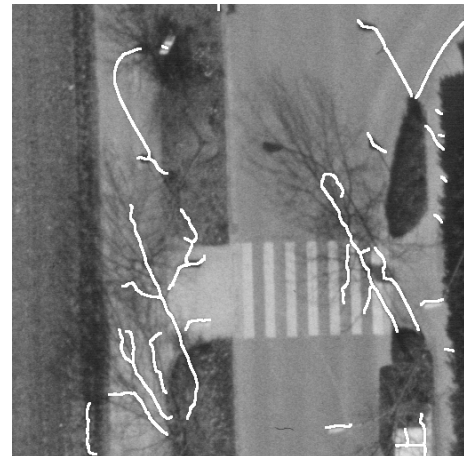
Figure 2: Part of the image Tamm (image courtesy of Photogrammetrie GmbH, Munich)

a) Part of the image   b) Distinct dark lines

Figure 3: Extraction of dark lines with high contrast

4.1 Hypotheses for Trunks

First, we extract distinct dark lines (cf. Fig. 3b), with a line extractor based on differential-geometry and the Gaussian scale-space (Steger, 1998). The parameters for smoothing and hysteresis thresholding were chosen in a way that only dark lines with high contrast were extracted.

To sort out the hypotheses for trunks, we make use of the following properties of their shadow projection: a) Trunks are represented by relatively long straight lines. b) The lines representing trunks are oriented in sun direction. As position and attitude of the image as well as the date and time of its acquisition are known, the sun direction could be computed. We have estimated it directly from the image. The detection of straight lines in a specific direction is a typical application of the Hough transform (Ballard and Brown, 1982). We apply it with the two parameters distance from the origin and orientation. For every pixel the orientation and the corresponding distance from the origin of all lines passing through it are stored in a two-dimensional (2D) array. In this array long and approximately straight lines with the same distance to the origin can be found as local maxima. Hypotheses for trunks are local maxima of an interval around the sun direction. In Figure 4a) the local maxima have been projected back into the image as straight lines. Here, the dark lines are oriented in sun direction and are called "Hough-lines" for the remainder of the paper. The light straight lines indicate those local maxima who are not oriented in sun direction.

Next, the lines in the image window which are represented by the Hough-lines are selected, connected by morphological closing (Serra, 1982), and thinned to their medial axes. The closing allows to link disconnected lines. Short lines are eliminated because they cannot correspond to trunks. For the remaining lines approximately straight lines are computed by linear regression. The begin and the end of a line is determined from the extent of the connected component. This
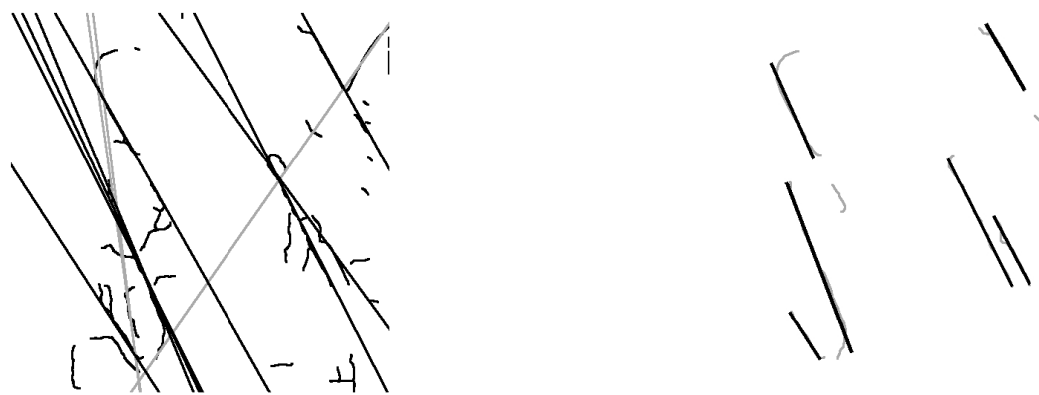
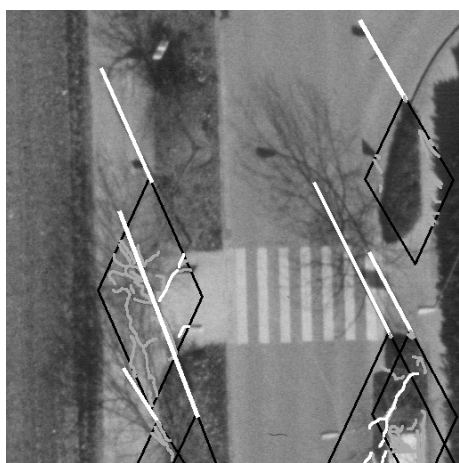a) Hough-lines (black)   b) Hypotheses for trunks

Figure 4: Result of Hough transformation and Hypotheses for trunks

results in basic hypotheses for trunks. Unfortunately, they can also correspond to the shadow projections of disturbing objects such as poles or street-lamps (cf. Fig. 4b). The result demonstrates that the extraction of the trunk is feasible, though it still has to be refined. For example, hypotheses for trunks can be tracked in sun direction to search for lines with low contrast.

4.2   Direct Projection of the Trunk in Nadir Direction

An important evidence for a hypothesis for a trunk is the existence of a trunk in direct projection. Additionally, the base of the trunk can be determined more robustly and more accurately based on the direct projection. Contrary to the shadow, which is always darker than the surroundings, the direct projection of the trunk can be both, lighter and darker. Therefore, we extract dark as well as light lines. Then, we use the same procedure based on the Hough transformation as described above for the selection of lines in nadir direction. Ideally, finally one single long line represents the direct projection of the trunk. As search area, a parallelogram is chosen. Figure 5a) shows in black the search areas for the hypotheses for trunks presented in white. One direction of the parallelogram is determined by the shadow projection, the other by the direction to the nadir point. The lengths are the same as for the hypotheses for the trunks. The extracted lines within the search areas and the extracted lines pointing in nadir direction are illustrated in grey and white, respectively. The procedure is successful for the street-lamp on the top-left and for the trunk on the bottom-right. However, a refinement is still necessary. Figure 5b) demonstrates the procedure in another part of the image, showing good results for the street-lamp on the bottom-right.
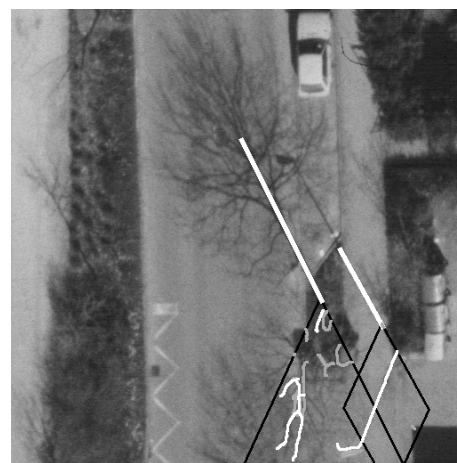
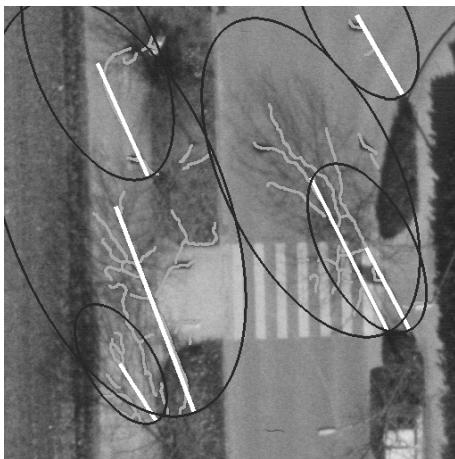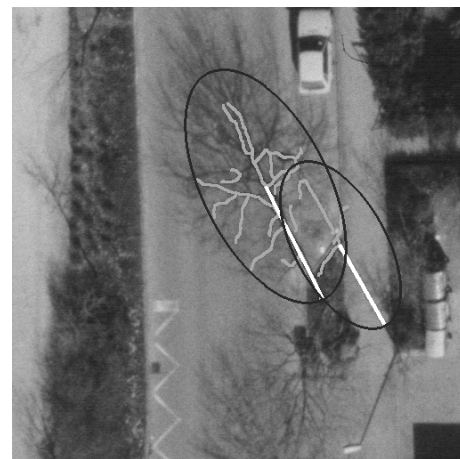a) Example 1   b) Example 2

Figure 5: Direct projection of trunks in nadir direction (Hypotheses for trunks from shadow projection = white; search area = black parallelogram; lines = grey; lines in nadir direction = white)

4.3 Branches in Shadow Projection

The extraction of branches in shadow projection is geometrically restricted using the hypotheses for the trunks and makes use of the properties of the chosen line extraction. The search areas are defined by ellipses with the hypotheses for the trunks defining their longer axes. The smaller axes have half the length. As the type of the tree is not yet known, the size of the ellipse, has to be chosen relatively large. Within these ellipses distinct dark lines are extracted (cf. Fig. 6a). We exploit the knowledge that branches are starting at the trunk and are of decreasing thickness by exploiting hysteresis thresholding (Canny, 1986) in the following way: The upper threshold corresponding to the starting points of the tracking is set very high to get only lines with high contrast, i.e., the trunk. On the other hand, the lower threshold controling the extent of the tracking is chosen relatively low, to track branches to their final ends. Only light smoothing is used in the image to detect also very thin lines. As the employed version of the line extractor does not work recursively, at every branching only one line is tracked. Figure 6b) shows results for the approach in another image part. Figure 7 presents the results for two different hypotheses for trunks in this image part. In one case, the hypothesis represents the trunk of a deciduous tree, but in the other case a disturbing object, i.e., a street-lamp a fact which could already be derived from the extracted branches.
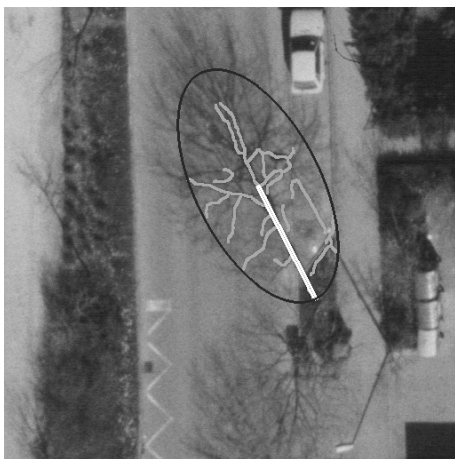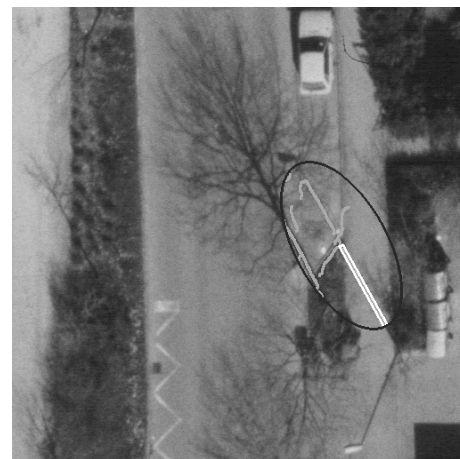
a) Example 1	b) Example 2

Figure 6: Tracking of branches (Search area = black ellipse; branches = grey)

a) Trunk-hypothesis 1	b) Trunk-hypothesis 2

Figure 7: Branches in example 2 for different trunk-hypotheses

5 CONCLUSIONS

With our investigations we have demonstrated that it is feasible to extract leafless deciduous trees from high resolution aerial imagery from their shadow and their direct projection in many cases. Even fine branches can be extracted at a ground pixel size of about 4 cm.

The tracking of the branches can be extended to all possible lines. Trees and disturbing objects, such as street-lamps, can

be distinguished based on the extracted branches already at this stage. The outline can be determined based on the end points of the branches, the blurred edge defined by the shadow projection of the twigs, and the compactness of the tree crown using a snake-based approach. We expect that the parameters nowadays collected manually, such as the height of the tree, the width of the tree crown, and the base of the trunk can then be determined fully automatically. An analysis of the structure of the branches and the shape of the outline should allow a classification of the type of the tree. A further improvement of the robustness of the extraction could be achieved by employing the relations to other objects, such as buildings and roads. For instance, rows of trees can often be found along a road and the base of a tree might be occluded by a given building. On the other hand, an avenue can be defined by parallel rows of trees.

REFERENCES

Ballard, D. and Brown, C., 1982. Computer Vision. Prentice-Hall International, Englewood Cliffs, USA.

Baumann, H., 1957. Forstliche Luftbild–Interpretation. Schriftenreihe 2, Landesforstverwaltung Baden-Württemberg, Selbstverlag.

Brandtberg, T., 1996. Automated Tree Species Classification in High Resolution Aerial Images Using a Hough Transform Technique. In: Swedish Symposium on Image Analysis.

Canny, J., 1986. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6), pp. 679–698.

Gougeon, F., 1995. A System for Individual Tree Crown Classification of Conifer Stands at High Spatial Resolutions. In: Canadian Symposium on Remote Sensing, Vol. 17, pp. 635–642.

Hildebrandt, G. (ed.), 1996. Fernerkundung und Luftbildmessung für Forstwirtschaft, Vegetationskartierung und Landschaftsökologie. Wichmann Verlag, Heidelberg.

Jaynes, C., Stolle, F. and Collins, R., 1994. Task Driven Perceptual Organization for Extraction of Rooftop Polygons. In: 2nd IEEE Workshop on Applications of Computer Vision, pp. 152–159.

Kass, M., Witkin, A. and Terzopoulos, D., 1987. Snakes: Active Contour Models. International Journal of Computer Vision 1(4), pp. 321–331.

Krummheuer, F., 1998. Auswertung von Luftbildern ausgewählter Buchenbestände mit Fouriertechniken. Photogrammetrie – Fernerkundung – Geoinformation 3/98, pp. 165–176.

Larsen, M., 1998. Finding an Optimal Match Window for Spruce Top Detection Based on an Optical Tree Model. In: International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, 10.-12.02.1998, Victoria, B.C., Kanada.

Lin, C. and Nevatia, R., 1998. Building Detection and Description from a Single Intensity Image. Computer Vision and Image Understanding 72(2), pp. 101–121.

Mayer, H., 1998. Automatische Objektextraktion aus digitalen Luftbildern. Deutsche Geodätische Kommission (C) 494, München.

Pollock, R., 1996. The Automatic Recognition of Individual Trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model. PhD Thesis, The Faculty of Computer Science, The University of British Columbia, Vancouver, Kanada.

Serra, J., 1982. Image Analysis and Mathematical Morphology. Academic Press, London, Großbritannien.

Shufelt, J., 1996. Exploiting Photogrammetric Methods for Building Extraction in Aerial Images. In: International Archives of Photogrammetry and Remote Sensing, Vol. (31) B6/VI, pp. 74–79.

Spurr, S. (ed.), 1960. Photogrammetry and Photo–Interpretation. The Ronald Press Company, New York.

Steger, C., 1998. An Unbiased Extractor of Curvilinear Structures. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, pp. 113–125.

Tönjes, R., 1997. 3D Reconstruction of Objects from Aerial Images Using a GIS. In: International Archives of Photogrammetry and Remote Sensing, Vol. (32) 3-2W3, pp. 140–147.