

DETERMINATION OF CORRESPONDING TRUNKS IN A PAIR OF TERRESTRIAL IMAGES AND AIRBORNE LASER SCANNER DATA

Olli Jokinen, Ulla Pyysalo and Petteri Pöntinen
Institute of Photogrammetry and Remote Sensing
Helsinki University of Technology
02015 TKK, Finland
olli.jokinen@tkk.fi

ABSTRACT

A method based on extracted trunks is proposed for automatic determination of corresponding trees between a georeferenced pair of terrestrial images and airborne laser scanner data in a wooded area. Trunks in the terrestrial images are obtained by region growing segmentation between left and right edges of trunks detected within a region of interest. Trunks in a digital tree height model generated from the laser scanner data are estimated as straight vertical lines through the weighted centers of canopy points of single trees. Corresponding trunks in the left and right images are found with the help of trunks in the laser scanner data and additional attributive information. A piece of 3-D line reconstructed from the trunks in the images verifies the corresponding trunk in the laser scanner data. An experiment illustrates the performance of the method. Potential applications include augmenting a 3-D forest model derived from laser scanner data with complementary information from terrestrial images.

1. INTRODUCTION

Airborne laser scanning in forest areas provides a large amount of 3-D data from both the crown of trees and the ground. Terrestrial images viewing the scene from the side are excellent for investigation and verification where the backscattering of laser pulses has occurred in small test areas. Complementary 3-D information not easy to obtain by laser scanning can be measured from a pair of convergent images and added to a 3-D forest model derived from the laser scanner data. A related question is how to find corresponding trees in the images and laser scanner data of a rather complex scene automatically. In this short article, a solution based on extracted trunks is proposed and described with the help of an example. The method requires that enough open space exists so that a pair of convergent images can be recorded. The terrestrial imaging is thus usually performed in a clearance among the trees.

The paper is organized as follows. The previous research on correspondence matching is discussed and our contribution is highlighted in Section 2. Section 3 describes the data used in this study. The extraction of trunks from a terrestrial image and laser scanner data are addressed and the method for determination of corresponding trunks is proposed in Section 4. Test results are presented in Section 5 and conclusions are drawn in Section 6.

2. RELATED WORK

The correspondence matching between 2-D terrestrial images and 3-D airborne laser scanner data or a 3-D model derived from laser scanner data has been previously studied mostly in urban areas related to virtual city modeling or mobile mapping. In many approaches, the correspondences are

given interactively using extracted points such as vertices of buildings (Haala and Brenner, 1999). Automation of the correspondence matching is considered in (Haala and Böhm, 2003) where the silhouette of a building extracted from a terrestrial image is matched with the outline of the building obtained from an existing 3-D CAD model. The silhouette is detected and localized by a generalized Hough transform and the exterior orientation of the image is refined by spatial resection. A point cloud computed from an oblique video sequence is automatically aligned onto another nadir-view point cloud given by a 3-D sensor such as an airborne laser scanner in (Zhao *et al*, 2004). The camera poses are solved by feature based matching in consecutive image pairs. The method is suitable for texturing 3-D models in semi-urban environments including also trees.

The 2-D to 3-D correspondence problem has been investigated in a greater extent using 3-D data sets or models provided by terrestrial sensors rather than airborne laser scanning or using aerial images rather than terrestrial ones. In (Jaynes and Partington, 1999), line pencils consistent with vanishing points are extracted from a terrestrial image and matched with a wire frame model of an urban area projected onto the image. The aligned images are then used to augment the wire frame model initially reconstructed from map and survey data. In (Coors *et al*, 2000), a set of virtual views is generated from a 3-D GIS model of a city and matched with a video sequence of terrestrial images using line segments as features. A modified Hough transform is considered for solving the exterior orientation of an aerial image without known correspondences in (Habib and Kelley, 2001) where edges extracted from the image are matched with linear features in a 3-D GIS model. Line matching between edges in an aerial image and a 3-D GIS model of an urban area is also addressed in (Duong, 2003). The parallelism and orthogonality constraints present in buildings are utilized for automatic matching of terrestrial images and terrestrial range scans in (Stamos and Allen, 2001). In (Böhm and Haala, 2005), terrestrial laser scans of buildings initially georeferenced using a low-cost GPS and compass are registered using 3-D to 3-D surface matching techniques first with each others and then with a digital elevation model derived from airborne laser scanner data or with a 3-D city model measured from aerial images. Building facades in the coarse city model are augmented with 2.5-D relief images (laser maps) generated from the terrestrial laser scanner point cloud. More related works dealing with the problem of 2-D to 3-D matching in general are referenced in (Stamos and Allen, 2001).

Our approach for the correspondence problem is based on matching line segments extracted from terrestrial images and airborne laser scanner data. Related work on line matching includes (McIntosh and Krupnik, 2002) where a digital elevation model produced from airborne laser scanner data is refined at surface discontinuities using edge lines extracted and matched from a pair of aerial images. The corresponding edge in the second image is searched within a parallelogram bounded by the epipolar lines of the end points of the edge in the first image and known bounds for the disparity. Cross-correlations are computed between each pixel on the edge and its corresponding pixel on a candidate edge in the other image. The best match is the edge where the average correlation at the pixels along the edge is at its maximum. Line matching in a sequence of terrestrial images is also addressed in (Karner *et al*, 2002). For a line segment in one image, candidate matches in other images are found using the epipolar geometry as in (McIntosh and Krupnik, 2002, Schmid and Zisserman, 1997). A set of 3-D line segments is computed and projected into all the remaining images of the sequence. If image lines are found near the projected lines, then the candidate is confirmed and the correct line among multiple candidates is selected using a correlation based similarity criterion. Finally, a 3-D city model obtained from airborne laser scanning or aerial images is augmented by the terrestrial images matched against the model using vanishing points and a few control points.

In this paper, the 2-D to 3-D correspondence problem is investigated in a wooded area using a pair of calibrated terrestrial images and airborne laser scanner data. A wooded area is more difficult than an urban area in point of finding well-defined corresponding features from the data. Our contribution is to show that corresponding trunks can be established in case the exterior orientation of the images with respect to the laser scanner data is known from GPS measurements or refined interactively as in (Rönnholm *et al.*, 2003). However, the accuracy of GPS/INS positioning is limited causing uncertainties in the georeferencing of both terrestrial images and airborne laser scanner data (Haala and Böhm, 2003) and this makes the automatic search for correspondences non-trivial in addition to that the trunks extracted from different data sources are usually only approximately in the same positions even if the data sets were georeferenced very accurately into the same coordinate system. In addition to finding correspondences between the images and laser scanner data, the proposed method solves the stereo matching problem between the images with the help of additional 3-D data given by the laser scanner. The correspondences established may be used to augment a 3-D forest model with complementary information from the terrestrial images.

3. MATERIAL

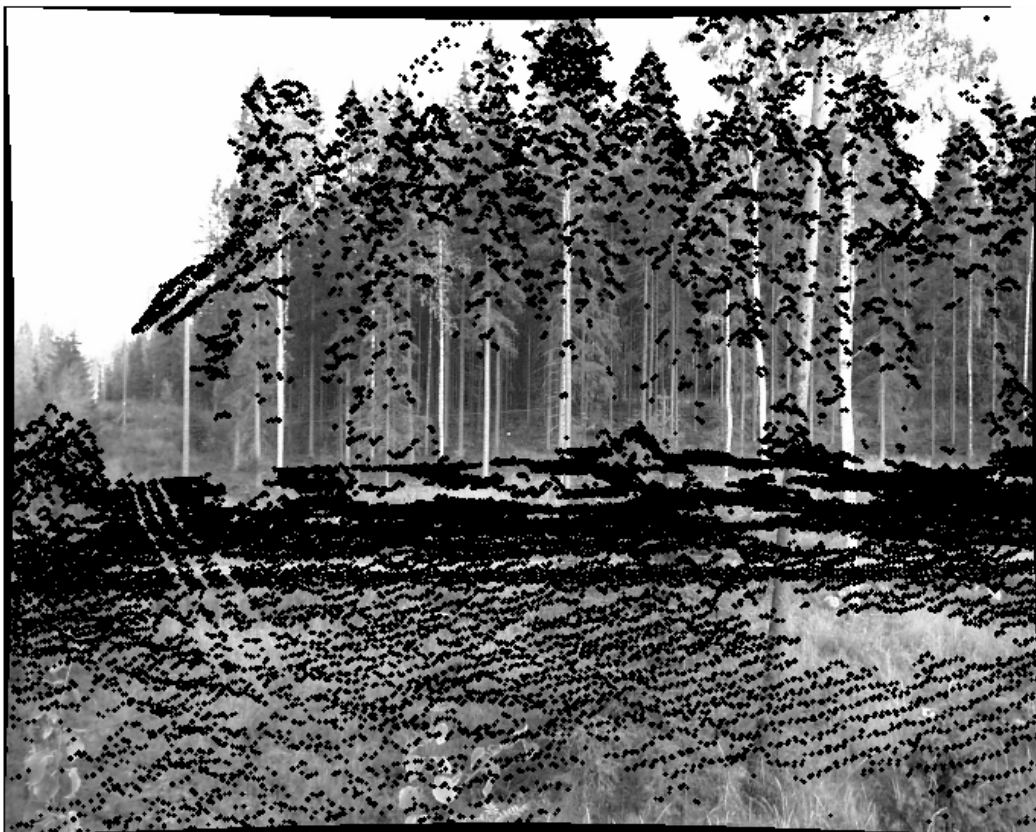


Figure 1. A terrestrial image efficiently visualizes where the backscattering of laser pulses actually occurs.

The materials used in this study include a pair of terrestrial images and airborne laser scanner data from a wooded area in southern Finland. Figure 1 shows one of the gray level terrestrial images of 1090 x 1363 pixels with image coordinates compensated for radial lens distortion. The exterior orientation of the image has been solved using control points the locations of which in a known reference coordinate system available in the test area have been measured with a tacheometer. A 3-D point cloud acquired using the Toposys-1 laser scanner of TopoSys GmbH,

Germany, from the altitude of 400 m with an average point density of 20 m^{-2} has been projected onto the image plane of the terrestrial image and rounded off to a binary image on integer pixels shown as black points. The exterior orientation of the terrestrial image has then been interactively adjusted by giving a small shift in the direction of the horizontal image axis so that the projected laser points fit the image better. The differences in original georeferencing between the laser scanner data and terrestrial images are believed to be mainly due to uncertainties in GPS positioning. There was also a time gap between the image acquisition and laser scanning during which changes may have occurred in the test area. There are power lines on the left side of the image which explain the laser points that look like hanging in the air.

4. METHODS

4.1 Extraction of trunks from a terrestrial image

Let us consider the case of Fig. 1 that shows a typical scene of a wooded area. The trunks are approximately vertical in the image and have a lighter color than the background forest under the light-colored sky. For other type of scenes, appropriate modifications to the following description of the method based on conventional feature extraction techniques may be needed.

A region of interest given by pixels located a threshold below an estimated outline of the forest against the sky is selected. The outline is extracted as an edge between the sky and forest as follows. The image is converted into a binary image using a brightness threshold and eroded with a structuring element consisting of points in the 3×3 neighborhood of the representative point. The eroded image is subtracted from the original one and for each column the edge pixel having the smallest row index provides an estimate of the outline (the rows increase downwards with respect to the scene). The part of the curve is removed where the actual outline is not within the image coverage and the remaining curve is smoothed out using local averaging.

Within the region of interest, vertical edges are detected using the Sobel operator and morphologically closed with the structuring element consisting of points in the 5×1 vertical neighborhood centered at the representative point. The edges are classified into left and right edges of trunks according to the sign of the partial derivative of the gray-level image function with respect to the horizontal image coordinate. The partial derivative is positive at left edges and negative at right ones. The regions between adjacent left and right edges are filled in for each row by starting from a right edge pixel and detecting the nearest left edge pixel located on the same row within a threshold from the right edge pixel. The resulting binary image consisting of ones at trunks and zeros elsewhere is segmented using a region growing algorithm presented in (Jokinen, 1997) and simplified to work with binary images. The center lines of trunks are estimated as straight lines through the pixels of each segment and the lines the parameters of which are close to each other are merged. Lines shorter than a threshold are rejected.

Attributive information is stored up for each trunk including the average gray level and width of the trunk segment and the species of the trunk. The trunks are classified into a pine, spruce, and birch according to the 95 percentile of the pixel gray levels in the segment. The trunk of a birch has a lot of light-colored pixels and a threshold for the percentile is set high. The trunk of a spruce is usually darkest while a pine is in between a birch and spruce. This classification is, however, subject to lighting conditions on the trunk. If color images were used, the color information might be helpful in classification and as an additional attribute for correspondence matching.

4.2 Extraction of trunks from laser scanner data

The method for the extraction of trunks from laser scanner data was developed in (Pyysalo, 2000) and it is described here briefly.

A digital tree height model is first generated from the laser scanner data as a difference of digital surface and terrain models. The height model is then segmented into single trees using a watershed algorithm based on detecting local maxima of the surface and analyzing the magnitude and direction of the surface gradient around the maxima. The trunk is defined as a straight vertical line from the top of the tree to the ground. The canopy points in each segment are divided into height layers and the location of the trunk is estimated as a weighted mean of the centers of the points in each layer giving more weight to the centers near the top of the tree where the branches are usually closer to the trunk. The trunks are classified into different species interactively according to a field survey performed.

4.3 Determination of corresponding trunks

Assume that we have a pair of convergent terrestrial images that have been georeferenced into the same coordinate system as the laser scanner data. The trunks are assumed to be approximately vertical in both of the images in accordance with Section 4.1. Concerning the relative orientation between the images, it is further assumed that the translation in the vertical direction and the rotations around the horizontal image axis and around the optical axis are small. A typical pair of terrestrial images is shown in Fig. 3 below. In case one of the exterior orientations or the relative orientation of the images differs from the assumptions, appropriate modifications to the following description may be needed.

For each trunk in the left image (left trunk), the corresponding trunk in the right image (right trunk) is searched with the help of 3-D trunks extracted from the laser scanner data (laser trunk). The search is performed for trunks tall enough and in a descending order according to the number of pixels in the segments of left trunks. The laser trunks are projected onto the left image and for a left trunk, the laser trunks are considered that satisfy the following conditions.

- The distance between the projected laser trunk and the left trunk is less than a threshold.
- The image row coordinate of at least one of the end points of the left trunk is between the image row coordinates of the end points of the projected laser trunk.
- The projected laser trunk is of the same species as the left trunk.

These candidates of laser trunks are projected onto the right image and for each of them the right trunks are possible matches that satisfy the following conditions.

- The distance between the right trunk and the projected laser trunk is less than a threshold.
- The image row coordinate of at least one of the end points of the right trunk is between the image row coordinates of the end points of the projected laser trunk.
- The attributive information of the right trunk regarding the gray level, width, and species is similar to the attributive information of the left trunk within appropriate tolerances.

A piece of 3-D line is reconstructed in the object space by intersecting the planes defined by the left trunk and each candidate of right trunks. For each left trunk, the corresponding right trunk and the corresponding laser trunk are those for which the distance from the reconstructed piece of line to the laser trunk is shortest in case this distance does not exceed a given threshold.

5. TEST RESULTS

The methods for extraction of trunks from terrestrial images and laser scanner data and for determination of corresponding trunks were tested using the data described in Section 3. Appropriate values for the many thresholds and tolerances needed in the algorithms were found out experimentally. The results are illustrated in Figs. 2-4.

The center lines of trunks extracted from the terrestrial image of Fig. 1 are shown as red lines in Fig. 2. Most of the trunks found are located on the edge of the forest and there are no erroneous pieces of lines left although some trees have several pieces of lines stacked vertically. The classification of trunks into different species worked well and we have observed only one misclassification of trunks in Fig. 2. All the laser trunks estimated and classified into different species have also been projected onto the terrestrial image and shown as green lines in Fig. 2. There are 43 left trunks (sampled out of approximately 35 trees) and 29 laser trunks in Fig. 2. Quite a many of them are close to each other and it is hard to say which are the corresponding trunks although the exterior orientation of the terrestrial image is supposed to be rather accurate as determined by the GPS equipment and adjusted interactively.

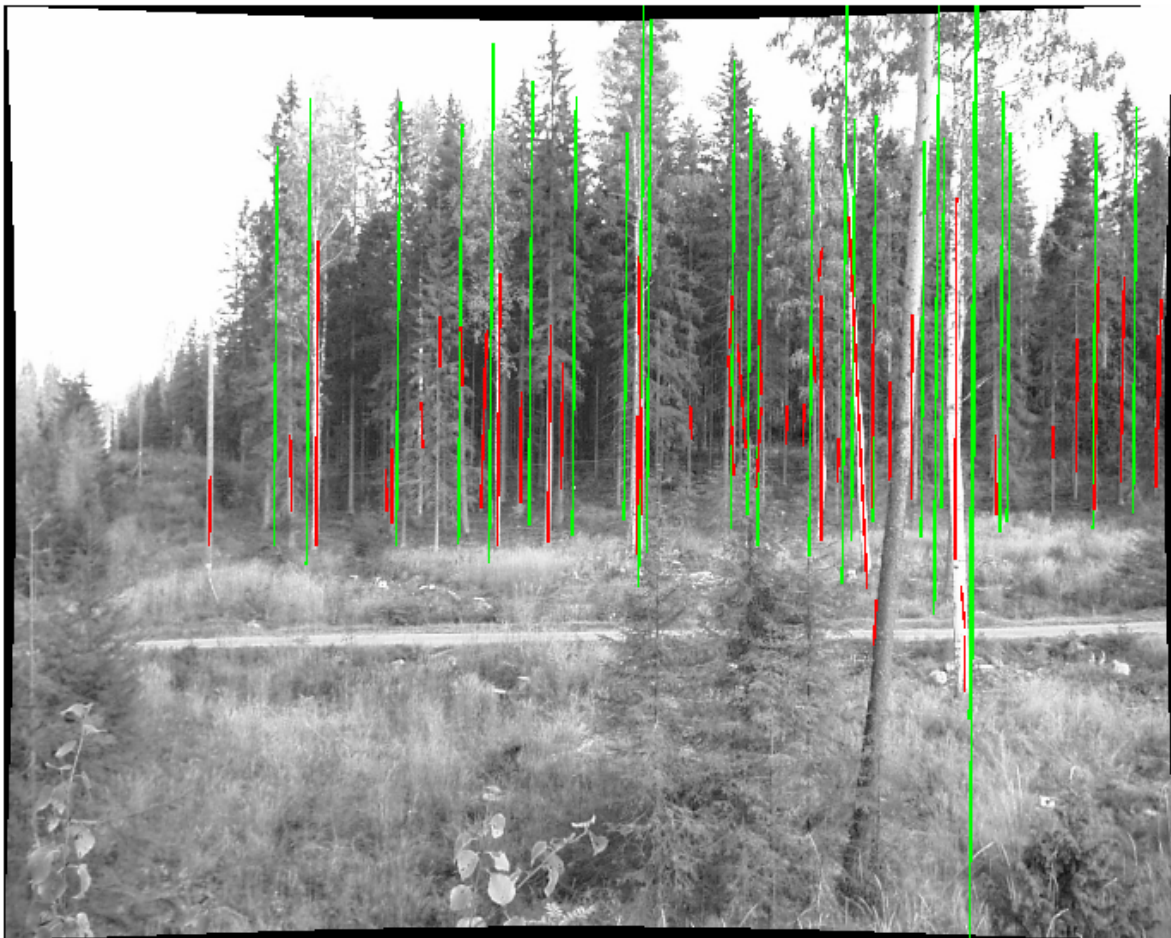


Figure 2. Trunks extracted from the terrestrial image are shown as red lines and trunks extracted from the laser scanner data and projected onto the image as green lines.

Figure 3 shows a pair of convergent terrestrial images. The corresponding left, right, and laser trunks established are shown as red and green lines, respectively. Ten correspondences were found out of the total of approximately 15-20 matches existing between the trunks extracted. Comparing the left and right images, it may be realized that there are several trees that are in shadow of other trees in either of the images. In order to verify that the corresponding laser trunk is the correct one, the result is plotted on a digital surface model generated from the point cloud of the laser scanner data via Delaunay triangulation and linear interpolation into a regular grid of 0.5 m x 0.5 m ground resolution as shown in Fig. 4. The trees are located at light-colored spots and the green squares show the locations of the corresponding laser trunks and the red squares are the center points of the pieces of lines reconstructed from the corresponding left and right trunks. The angular forms illustrate the viewing areas of the left and right cameras located at the vertices of the angles. The RMS distance between the laser trunks and the corresponding center points of pieces of lines is 2.1 m.

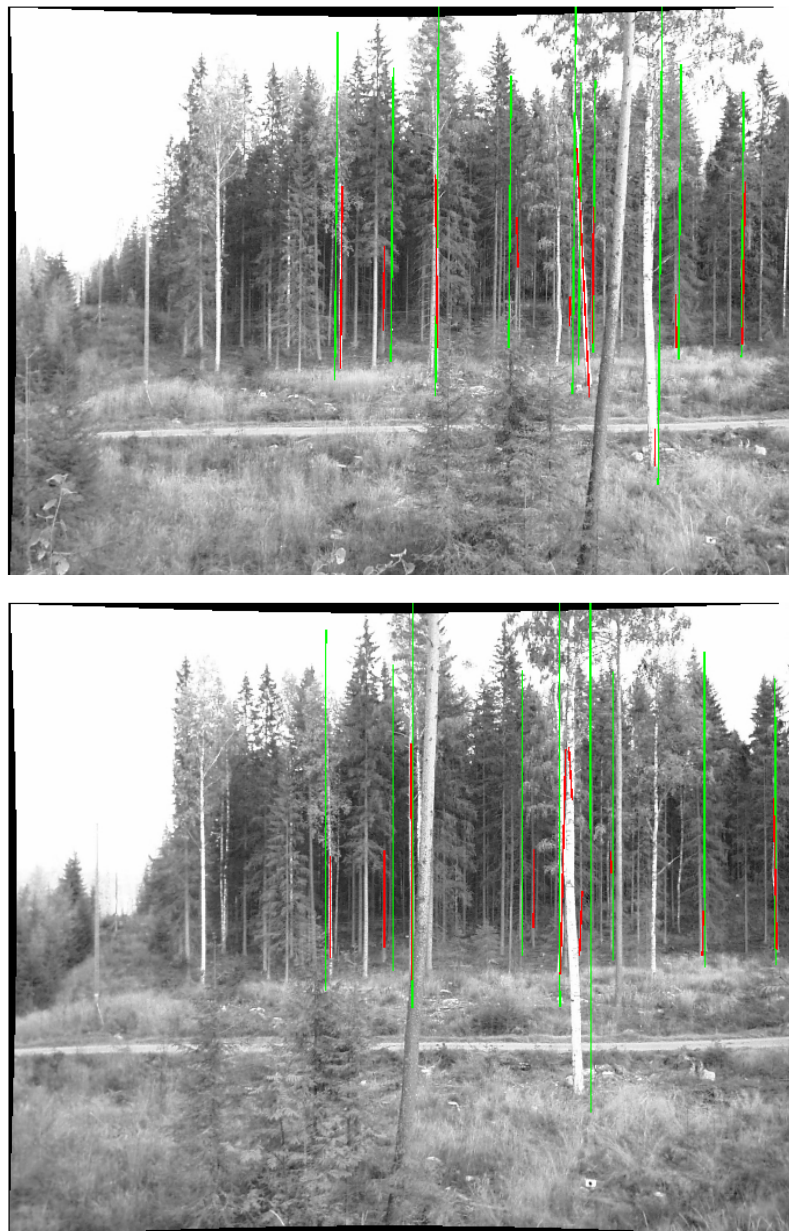


Figure 3. The corresponding trunks found in the left and right images are shown as red lines and the corresponding laser trunks projected onto the images as green lines.

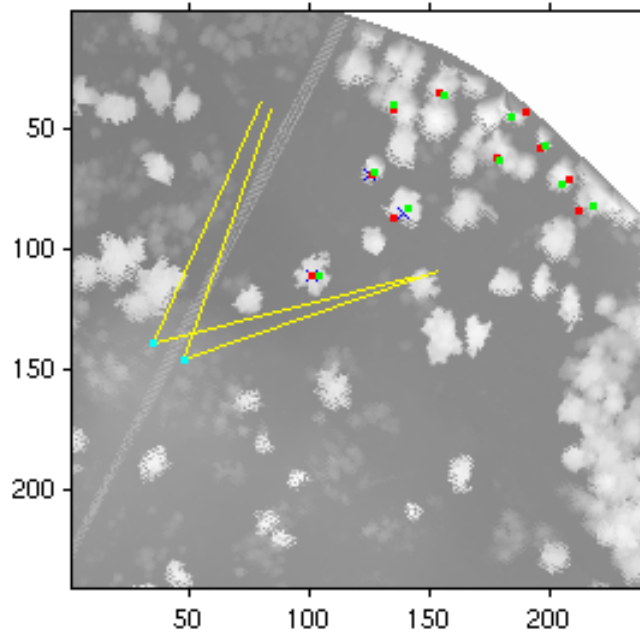


Figure 4. The positions of the corresponding trunks and the viewing areas of the cameras visualized on a digital surface model of 0.5 m x 0.5 m ground resolution (green squares denote laser trunks, red squares denote center points of pieces of lines reconstructed from left and right trunks, and blue crosses denote tacheometer measurements).

The locations of some of the trunks had also been measured by a tacheometer and compared to the laser scanner measurements in (Pyysalo and Hyyppä, 2002). The accuracy of the locations derived from the laser scanner data was reported to be better than one meter. It appeared that three of the trunks measured by the tacheometer coincided with the corresponding trunks established between the left and right images and laser scanner data. The locations of these tacheometer measurements are shown as crosses in Fig. 4. The RMS distances between the locations of trunks measured by the tacheometer and the locations of corresponding trunks determined from the images and laser scanner data are 1.1 m and 1.5 m, respectively. The precision of locations of 3-D trunks determined from the images is affected by uncertainties in the left and right trunks and uncertainties in the exterior orientations of the images. An error is introduced when trunks are approximated as straight lines in the images while true trunks (and especially trunks of birches) are not really straight but often curved or twisted. For the three trunks mentioned above, the end points of the left and right trunks are within a distance from the true center lines of the trunks, the standard deviation of which is 5 pixels in the image column coordinate. Error propagation yields then a root mean trace of 0.5 m for the covariance matrices of the ground coordinates of center points of pieces of lines reconstructed from the left and right trunks. The standard deviations of the exterior orientation parameters were 0.03 degrees in the rotation angles and 0.02 m in the translations of the left and right cameras on the average. When the covariance matrices of the exterior orientation parameters are included, the root mean trace of the covariance matrices of the ground coordinates of center points of reconstructed pieces of lines increases less than 0.001 m.

6. CONCLUSIONS

This short article has presented a method to find automatically for each trunk in the left image in a pair of terrestrial images, the corresponding trunk in the right image and the corresponding trunk in the airborne laser scanner data. The results indicate that the method works in rather complex wooded areas although there are quite a many thresholds that need to be carefully given

case by case and some preprocessing steps are currently performed interactively. 50-67% of the corresponding trunks were found in the test case. It is crucial for the performance of the method that the laser scanner data and the images have been accurately georeferenced into the same coordinate system as the method is based on the projection of laser trunks onto the images and the intersection of left and right trunks in 3-D. Since the viewing directions of the terrestrial and airborne sensors are different and the data they provide of different type, the trunks extracted from the terrestrial and airborne data are only approximately the same and this evidently affects the performance of the method, too. The information about corresponding trees is useful for subsequent scene modeling and better understanding of laser reflections with the help of complementary information from terrestrial images. Further studies are required to find if the correspondences could be used for refining the exterior orientations of the images.

7. REFERENCES

Böhm, J. and Haala, N., 2005. Efficient integration of aerial and terrestrial laser data for virtual city modeling using lasermaps, ISPRS WG III/3, III/4, V/3 Workshop on Laser scanning 2005, Enschede, The Netherlands, pp. 192-197.

Coors, V., Huch, T., and Kretschmer, U., 2000. Matching buildings: pose estimation in an urban environment, Proc. IEEE and ACM International Symposium on Augmented Reality, Munich, Germany, pp. 89-92.

Duong, V.H., 2003. Real-time building identification system development using aerial imagery, Master's Thesis No. SR-03-03, Asian Institute of Technology, Thailand.

Haala, N. and Brenner, C., 1999. Extraction of buildings and trees in urban environments, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 54, No. 2-3, pp. 130-137.

Haala, N. and Böhm, J., 2003. A multi-sensor system for positioning in urban environments, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, No. 1-2, pp. 31-42.

Habib A. and Kelley, D., 2001. Single-photo resection using the modified Hough transform, Photogrammetric Engineering and Remote Sensing, Vol. 67, No. 8, pp. 909-914.

Jaynes, C. and Partington, M., 1999. Pose calibration using approximately planar urban structure, Proc. Asian Conference on Computer Vision, <http://www.metaverselab.org/pub/paper/accv-jaynes-00.pdf> (4.7.2005).

Jokinen, O., 1997. Building 3-D city models from multiple unregistered profile maps, Proc. International Conference on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa, Canada, pp. 242-249.

Karner, K., Bauer, J., Klaus, A., and Schindler, K., 2002. Metropogis: A city information system, Proc. IEEE International Conference on Image Processing, Rochester, NY, United States, Vol. 3, pp. 533-536.

McIntosh, K. and Krupnik, A., 2002. Integration of laser-derived DSMs and matched image edges for generating an accurate surface model, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 56, No. 3, pp. 167-176.

Pyysalo, U., 2000. A method to create a three-dimensional forest model from laser scanner data, *The Photogrammetric Journal of Finland*, Vol. 17, No. 1, pp. 34-42.

Pyysalo, U. and Hyypä, H., 2002. Reconstructing tree crowns from laser scanner data for feature extraction, *Photogrammetric Computer Vision*, Graz, Austria, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIV, Part 3 B, pp. 218-221.

Rönnholm, P., Hyypä, H., Pöntinen, P., and Haggrén, H., 2003. Interactive relative orientation between terrestrial images and airborne laser scanning data, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Dresden, Germany, Vol. XXXIV, Part 3/W13, pp. 130-135.

Schmid, C. and Zisserman, A., 1997. Automatic line matching across views, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, PR, United States, pp. 666-671.

Stamos, I. and Allen, P., 2001. Automatic registration of 2-D with 3-D imagery in urban environments, *Proc. International Conference on Computer Vision*, Vancouver, BC, Canada, Vol. 2, pp. 731-736.

Zhao, W., Nister, D., and Hsu, S., 2004. Alignment of continuous video onto 3D point clouds, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, DC, United States, Vol. 2, pp. 964-971.

ACKNOWLEDGMENTS

This research was financially supported by Helsinki University of Technology and the Academy of Finland.