A method for linking TLS- and ALS-derived trees

Andreas Fritz¹, Holger Weinacker¹ & Barbara Koch¹

¹University of Freiburg, Department of Remote Sensing and Landscape Information Systems, andreas.fritz@felis.uni-freiburg.de

Abstract

Within the past decade progress towards automatic recognition of individual trees and their parameters was made in both TLS and ALS-data based algorithms. In this paper we present an approach to combine single trees derived from ALS and TLS-data in order to gain a higher level of information. Therefore, two data sets are used: 1. a set of 3D-stemfiles generated by the algorithm described in Bienert et al. 2007 and 2. a set of detected single trees for the corresponding area of the data set 1 based on the algorithm described in Gupta et al. 2010. The 3D-stemfiles include position, information regarding sweep and diameter in 10[cm] height intervals. The ALS-tree description covers the position, maximum crown diameter and length as well as tree top height. This information is used for a hierarchic approach of linking ALS and TLS-derived trees based on three different initial matching algorithms. The estimated position error is taken into account to generate an initial list of matching candidates. The 2D-distance based initial linking method linked 41% of the TLS-trees. It was found that 3D-estimation of the tree top based on sweep information of the TLS-trees led only to minimal more imputations than the 2D-approach. A possible reason is seen in the linear models chosen, which do not reflect the tree shape invariably. Future work focuses on the integration of species information and the quantification of false linkage, which could not be evaluated within this study.

Keywords: terrestrial laser scanning, airborne laser scanning, tree linking, 3D-line fitting

1. Introduction

In recent years, object recognition based on Lidar generated point clouds has been an important topic in both science and industry. Airborne laser scanner (ALS) data in different point density levels is available throughout major countries and updated in useful time intervals. In the field of forestry applications, research related to biomass estimation and single tree detection are among the most regarded publications. High potential has been demonstrated for the retrieval of relevant information regarding economic and ecological parameters of forests from ALS data. Næsset 2002 finds a good estimation of dominant height, mean height, mean diameter, basal area and volume of stands in Scandinavia. In Straub et al. 2009, methods for single tree based total stem volume predictions in different forest conditions in southern Germany are compared. In the latest publication of Gupta et al. 2010, singletree parameters predicted in the same stands as the previous mentioned study include tree top height and position, crown length and maximum crown diameter. A comparison of different algorithms regarding their performance under different forest conditions can be found in Vauhkonen et al. 2010. Particularly singletree delineation and the related description of quality parameters are of high importance for the diverse European wood industry. Precise economically relevant description of standing trees is among the top applications of terrestrial laser scanning in forests. Liang et al. 2008 detect tree trunks in a distance up to 60[m] with a single view scan. Bienert et al. 2007 present a commercialized system to detect trees in one scan mode. For each detected tree, a profile fitting in a height interval of 10[cm] was conducted, thereby quantifying diameter, volume and shape of the entire tree.

Combining the features extracted from both ALS and TLS data promises a precise and
comprehensive description of the ecosystem forest and its economic value. Few approaches are known to the author dealing with registration of ALS and TLS generated objects in wooden environments. Olófsson et al. 2008 present a method of linking field data to ALS derived trees. The linking is based on dbh, height and distances between ALS- and field-surveyed trees. Ground truth tree data generated in Stanford stemfiles during felling with harvesters is used in Holmgren et al. 2010. A 10[m]-radius sample plot size is used and the linking is based on ALS generated tree attributes as crown polygon and height.

In this paper, we present an approach of combining the TLS and ALS datasets on object (tree) level. We use the results of proven algorithms developed at the University Freiburg, Department of Remote Sensing and Landscape Information Systems (Gupta et al. 2010) and University Dresden, licensed by TreeMetrics Ltd. (Bienert et al. 2007).

2. Material

2.1 Study area

The study area is located in south-west Germany (49.0N 8.4E) in the municipal forest of Karlsruhe. The extent of the test site is 2.1x4.4[km²] and predominantly includes pine, beech and spruce. Within the area, 42 permanent inventory plots where selected to cover typical forest conditions for this site. The plots include scots pine (Pinus sylvestris), sessile oak (Quercus petrea), red oak (Quercus rubra) and European beech (Fagus sylvatica).

2.2 TLS data

TLS data acquisition was conducted on the centre point of all plots by TreeMetrics Ltd. in November 2010. The scanning device was a Faro Imager 5006i. Data was processed with the algorithm described in Bienert et al. 2007. In a pre-processing step, all returns not within a radius of 15[m] of the plot centre were discarded. The settings were made for high precision and accuracy, which leads to a high omission error. Manual inspection did not reveal any false tree detection, but a high number of undetected stems can be observed. The log length of the detected trees is underestimated due to the strict parameter settings. The centre coordinates of the plots were measured by a surveying company and have a location error of < 5[cm]. The coordinates for each detected stem were calculated with distance and azimuth to the centre. The processing result is a list for each plot containing the detected trees in form of coordinates, diameter and volume in a height interval of 10[cm]. In 42 plots, 831 trees were detected (min=1, max=21, mean=8.9).

2.3 ALS data

In November 2009, ALS data was acquired with a Riegl LMS-Q560 Scanner. The point density is ca. 24[pt/m²]. The method used for detecting and delineating single trees is described in Gupta et al. 2010. In addition to the top tree position, two more parameters are calculated for each detected tree: maximum crown diameter and crown length. Within a 20[m] radius of the TLS-scanner position, 1277 trees were detected.

3. Methods

3.1 Data preparation

To reduce false imputation, data cleaning was performed on the TLS-tree set. Here, all small neighbouring trees that (a) are within 1.5[m] and (b) have a dbh difference > 10[cm] were removed. This is to dismiss trees that, although they are detected by TLS, most likely are not
detectable by ALS (i.e. because they are in the understory), as being linking candidates.

Figure 1 view through a TLS-scan with an understory tree covered by a dominant tree. After cleaning the smaller tree has been removed.

Figure 1 shows an example of two trees with a distance to each other of 0.7[m] and dbh-difference of 24.5[cm]. The smaller tree on the left was discarded according to the cleaning condition.

3.2 Initial linking

Firstly, the 2D Cartesian distance for each TLS and its neighbouring ALS tree is calculated when the condition

\[ H d_{TA} < |e| \]  

is fulfilled. Where \( H d_{TA} \) refers to the horizontal Cartesian distance between TLS and ALS-trees and \( e \) to the combined position error of both datasets.

\[ e = \sqrt{\varepsilon_{ALS}^2 + \varepsilon_{TLS}^2} \]  

\( \varepsilon_{ALS} \) and \( \varepsilon_{TLS} \) are the errors from the two detection algorithms. Each TLS-tree is then associated to potential ALS-trees. A n:m relation is possible, which potentially associates multiple TLS trees to multiple ALS-trees.

3.3 Advanced initial linking

On the test site, a distinct leaning can be observed, particularly pine trees. This implicates that the crown is often not perpendicular to the tree position at dbh. Horizontal distances up to 7[m] between tree top and dbh position were observed in a reference data set. Therefore, we considered additional information provided by the TLS-trees. Those trees have a certain length
which is not related to the actual tree height but more to branchiness, occlusion and parameter settings of the detection algorithm; however, the sweep of the stem is reflected very well. This sweep is used to estimate the potential position of the tree top by fitting a straight line through each 3D-coordinate tuple of the tree. For getting meaningful estimations, this approach was applied only to stems which are longer than 5[m]. The line fitting algorithm is explained in detail in Drixler 1993. The algorithm is based on quadrics and utilizes the concept of eigenvalues. Fitting the 3D-line follows four steps (Drixler 1993 p.47ff.):

1. The coordinates of the TLS-derived tree are reduced with their centre of gravity and the result is stored in matrix $A^*$
2. Three eigenvalues and eigenvectors of the matrix product of $A^* A^*$ are computed
3. The corresponding eigenvector of the maximum eigenvalue represents the direction vector of the fitted line.
4. The geometric gravity point of the TLS tree points is used as the origin point.
5. The variance-factor $\sigma$ is calculated with the two smaller eigenvalues as $\sigma = \sqrt{\frac{\lambda_1 + \lambda_2}{n-3}}$

This approach returns a 3D-straight line defined by 6 parameters $(x_{op}, y_{op}, z_{op}, x_{vect}, y_{vect}, z_{vect})$. The resolution of the $x$-$y$-changes along the stem is in the range of millimetres. This means the points along the stem may reflect small buckles and curvature at the stem bottom what can substantially influence the line fitting. To overcome this uncertainty, an iterative approach was developed that begins at the stem top and moves downwards. Initially, the topmost 10 points are used for an initial line fit. For each iteration, one point is added to the point set used by the fitting algorithm. The change of the variance-factor $\sigma$ per step is observed and when $\sigma$ becomes greater than 0.009, the process ceases. The threshold of 0.009 was found empirically.

The fitted line is then elongated to the height of the highest detected tree in the ALS-data. The 3D minimum distance is measured between the fitted line and the ALS trees and if distance values fulfill the criteria described in formula (1), the pair is put on the list of potential matching candidates.

A third list is generated with the help of 3D-boxes. The estimated tree-top and the geometric gravity point, both derived from the line fit approach specify a 3D-box. The parameters maximum crown diameter and tree top height define another 3D-box. When the two shapes intersect the trees are put on the list of possible matching candidates.

### 3.3 Final linking procedure

The list of potential matches contains all pairs of TLS- and ALS-trees in which, for example, one TLS tree is linked to several ALS-trees and vice versa. Figure 2 outlines the different steps of the linking procedure. The algorithm begins with a random TLS-tree and analyses its neighbourhood situation. It checks how many neighbouring ALS- and TLS-trees are present. The associated neighbours are again checked for their neighbours and so on. This process stops if there are no more partners on the list. The network is stored in a temporary list, which is the starting point for further analyses.
3.3.1 Analysing tree networks

In case of a real 1:1 association, the algorithm stops and declares the ALS-TLS-tree pair as successfully linked. A real 1:1 situation exists when the ALS- and TLS-tree do not have further potential partners within the distance of $e$.

For situations where the initially randomly picked TLS-tree has a network of partners, the horizontal distance, the dbh, the maximum crown diameter of the neighbouring trees are of relevance. The linking is realised according to a decision tree (Figure 2). The algorithm starts with another 1:1 check, since through the following steps new 1:1 situations can arise and goes on with distance comparison.

The distance comparison is defined as those trees will be linked with the tree that is the shortest distance away and if the difference to the second shortest is less than the standard deviation of all distances in the initial list.

When no significant distance difference can be observed, the maximum crown diameter is taken into account. If a 1:1 situation is achieved by checking whether a TLS-tree is exclusively within a buffer defined by the maximum crown-diameter, the ALS- and TLS-tree pair is declared as
successfully linked. If there is still no match possible, a general estimation of the dbh of the ALS tree is made by a logarithmic regression equation based on site specific parameters and literature review. Species information is not available and, therefore, not taken into account. The procedure compares each pair concerning the dbh estimated and measured. The standard deviation of all TLS measured dbh values is taken as the threshold for imputation. When the difference between the measured and estimated dbh is greater than the standard deviation, the pair is discarded from the list. In the case that only one pair meets the condition, it is matched. If more than one combination is possible, then no match is performed.

### 4. Results

The distance threshold $e$ was calculated with the following values:

1. Since $e_{ALS}$ was not specified in Gupta et al. 2010, the author checked the position accuracy manually and concluded that the horizontal distance between the reference tree top and found tree top was up to $\pm3.2[m]$.
2. In Bienert et al. 2007, the position error is not mentioned, but analyses regarding the error of the dbh estimation were performed. Seeing as the diameter is through the stem centre associated with the tree position those result were used as input for $e_{TLS}$. Bienert et al. 2007 found a maximum error for the tree diameter of $\pm19.6[cm]$ for Sitka spruce.

A test series with $e = 1.5 ... 7[m]$ applied to the 2D-distance based candidate list showed a strong correlation ($R^2 = 0.98$) between $e$ and the number of matched pairs.

#### 4.1 2D-distance based candidate list

The 2D-distance based initial linking leads to an initial list of 880 possible combinations. After the cleaning of understory trees, 858 combinations remain. Clearing the 1:1 situations leads to matches of 185 TLS-trees to one ALS tree and 673 combinations need further investigations. 6 matches were made through the second 1:1-check. The query regarding the distance difference matched 106 TLS-trees. The maximum crown diameter was responsible for 34 matches and the dbh comparison led to 24 imputations. In total 355 (41%) out of 831 TLS-trees were linked to an ALS-tree.

#### 4.2 3D-line fit based candidate list

After clearing small trees likely covered by a big tree, the initial list of potential pairs within of $3.2[m]$ minimum 3D-distance of the fitted line to an ALS-tree-top contains 619 combinations (previously 597). In total, 284 (48%) TLS-trees were matched with an ALS-tree. The initial 1:1 situation was dominating the imputation with 179 cases. The distance comparison led to 56 matches and the crown diameter was responsible for 25 links. Analysing the (estimated) dbh of the two tree sets linked 21 TLS-trees. Only 3 pairs were matched through a recheck for 1:1 situations.

#### 4.3 3D-box intersection based candidate list

The 3D-box intersection returns approximately double the amount of potential pairs (1224) compared to the 3D-line-fit based method. After clearing understory trees, 1202 remain. Compared to the other approaches, the number of initially matched pairs is relatively low (57) whereas matching through distance difference (153) dominates. Imputation through crown inclusion and dbh-comparison was responsible for 45 respective 22 pairs. In total 294 (24%) TLS-trees were linked to an ALS-tree.
4.4. Matching reliability

The three different results were analysed in respect of differences in linked tree pairs. 181 pairs were matched with all three initial linking methods. With both the 2D-distance based and the 3D-line fitting based candidate list, 227 pairs were matched. 3D-boxes and 2D-distance based initial list generation led to 216 pairs which appear after the algorithm in both final imputations lists. When comparing the two 3D-approaches, 221 pairs appear in both candidate lists. In total, 454 different TLS-trees could be matched when running all three initial list generators separately and combining the results.

5. Discussion & Conclusion

The results show that depending on the initial linking method approximately one third of the TLS-trees could be linked to an ALS-tree. One has to consider the fact that in our case detection algorithms for terrestrial laser scanners include trees with a dbh of 7[cm] and higher. This implies a substantial number of detected trees that do not take part in the crown layer and are thus barely detectable in airborne laser scanner data. We exclude potentially fully covered TLS-trees that are close to another TLS-tree to reduce that effect, however this does not attack the problem of isolated standing trees that are not dominant enough to appear in the crown layer visible to ALS-scanners. To solve this issue, a more reliable dbh-estimation is essential. In general, the fact that the species information is missing thus far is seen as a major obstacle for a better linking. With known species in either TLS or ALS-tree data, the dbh-estimation could be of a higher accuracy. Recent work described in Heinzel and Koch 2011 show promising results for tree species detection in mixed stands. It is planned to include these findings in the future. The three approaches of making an initial candidate list do not necessarily produce the same pairs. Only 181 matches were identical in all three approaches. A possible reason for that is the consideration of the third dimension. It was observed that, even with a low variance factor, the fitted line did not always reflect the growth direction of the stem reliably. In these cases, the projected tree top position is very likely incorrect and therefore a wrong ALS-tree gets linked. This problem does affect the 3D-boxes approach as well. The fitting approach returns an
over-parameterized definition of a line since a straight line in 3D is defined by four independent parameters \((\alpha, \beta, \gamma, d)\). The method can be questioned when taking this into consideration. However, we believe that the negative effect of the fitting algorithm itself is not causal for wrong tree top estimation. The high resolution of 2-dimensional direction changes is seen as the major reason for the failing of the fitting. Since the underlying model is a straight line, this leads to an insufficient result in cases where the tree has a distinct sweep along the complete stem. For future work, better indicators need to be found to decide whether or not a line fitting is appropriate or not.

A major problem is to quantify false linkage, since there is no information that can prove correct linking. A solution to that problematic could be to use the TLS-point cloud to assess the accuracy of the tree top position and maximum crown diameter. First visual inspection showed that with good registration, useful accuracy estimation with respect to linking is possible. A cut through the point TLS-point cloud with the detected ALS-trees on top is shown in Figure 3.

The strong linear correlation between the numbers of matched pairs and \(e\) is not surprising since the number of potential pairs increases dramatically. The quality of the result is certainly problematic because more unrealistic matches are made when the maximum distance is increased.

In future research the effect of the linking on biomass and forest parameters estimation is investigated.

**Acknowledgements**

The research leading to these results has received funding from the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement n° 245136. I would particularly like to thank Garret Mullooly for scanning and processing the TLS-data.

**References**


