

Change detection of trees in urban areas using multi-temporal airborne lidar point clouds

Wen Xiao^{*a,b}, Sudan Xu^a, Sander Oude Elberink^a, George Vosselman^a

^aFaculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands;

^bIGN, MATIS, 73 avenue de Paris, 94160 Saint-Mandé, France; Université Paris-Est

ABSTRACT

Light detection and ranging (lidar) provides a promising way of detecting changes of vegetation in three dimensions (3D) because the beam of laser may penetrate through the foliage of vegetation. This study aims at the detection of changes in trees in urban areas with a high level of automation using multi-temporal airborne lidar point clouds. Three datasets covering a part of Rotterdam, the Netherlands, have been classified into several classes including trees. A connected components algorithm was applied first to group the points of trees together. The attributes of components were utilized to differentiate tree components from misclassified non-tree components. A point based local maxima algorithm was implemented to distinguish single tree from multiple tree components. After that, the parameters of trees were derived through two independent ways: a point based method using 3D alpha shapes and convex hulls; and a model based method which fits a Pollock tree model to the points. Then the changes were detected by comparing the parameters of corresponding tree components which were matched by a tree to tree matching algorithm using the overlapping of bounding boxes and point to point distances. The results were visualized and statistically analyzed. The difference of parameters and the difference of changes derived from point based and model based methods were both lower than 10%. The comparison of these two methods illustrates the consistency and stability of the parameters. The detected changes show the potential to monitor the growth and pruning of trees.

Keywords: change detection, high vegetation, 3D modeling, airborne lidar, point cloud

1. INTRODUCTION

Change detection has become a major application of remote sensing techniques which provide viable data of repetitive coverage at short time and consistent quality. Especially, changes in vegetation covered areas are of great interest because they are crucial for ecosystem monitoring where digital change detection method is widely used¹. Vegetation in urban areas is a vital part of the living conditions. The proportion of vegetation covered areas is an essential factor for urban planning. Urbanization and industrialization will severely affect the growth of vegetation. Therefore the changes of vegetation should be monitored and estimated. As a relatively new remote sensing technology, airborne laser scanners provide a promising way of change detection of vegetation in three-dimensional (3D) perspective because it generates point clouds with accurate 3D coordinates. Using high density point clouds, the changes in both coverage and height can be detected². Vegetation changes in forestry at plot level, such as biomass or average height, have been studied³. Lidar data processing, e.g. feature extraction, classification, has been discussed a lot and some researches have taken vegetation into consideration. However, no research has focused on tree changes in 3D in urban areas.

In laser scanning data, vegetation is represented by irregularly distributed points. Rutzinger et al.⁴ detect high vegetation in urban areas using airborne lidar data. Parameters of individual tree are generated directly from the point clouds. Yu et al.⁵ develop an approach for extracting individual tree attributes, i.e. height, diameter at breast height (DBH) and stem volume. They also detect harvested trees and forest growth using airborne lidar data⁶. The estimation of height growth is accomplished by individual tree delineation and a tree to tree matching algorithm. Individual tree height growth is detected again by Yu et al.³, who present three change detection manners, i.e. differentiation between digital surface models (DSMs) and canopy height models (CHMs), canopy profile comparison and analysis of height histograms.

*wen.xiao@ign.fr; phone 33 (0)6 71 45 37 92; ign.fr

What is more, 3D tree modeling in lidar data recently becomes a hot topic. Models used in traditional remote sensing techniques and computer science are commonly utilized in lidar data. A fixed shape model or individual tree-wise models are applied in both mobile laser scanning (MLS) and airborne laser scanning (ALS) data. Rutzinger et al.⁷ utilize four different crown shapes having different diameters at three height levels using 2D enclosing circles. Then trees are modeled by an open source framework OpenAlea. Wang et al.⁸ analyze the vertical canopy structure of forest and also model trees in 3D. A voxel based method for individual trees delineation is implemented at different height levels. Then tree crowns are modeled and several crown parameters, e.g. tree height, crown height, crown diameter and volume are derived. Vosselman⁹ detects trees in airborne lidar data by computing the local maxima with a detection rate of 97%. The tree crown is modeled using a fixed shape whose diameter is adaptive to the height of the local maximum. Instead of regression models, Kato et al.¹⁰ develop a 'wrapped surface reconstruction' method. Tree parameters, e.g. tree height, crown diameter, crown base and volume are derived by the wrapped surfaces. And the results are validated by comparing with total station surveyed field measurements.

This paper proposes a method including a series of algorithms to detect the changes of trees semi-automatically. Section two describes the test site and data sets. Section three introduces the research methodology. The flowchart is depicted first and then each step is described in detail. Section four visualizes and statistically analyses the results of data processing. Discussions based on the analysis are followed. Conclusions are drawn in the last section.

2. DATASETS AND STUDY SITE

Three datasets were obtained on behalf of municipality of Rotterdam, two of them were obtained in 2008 (March and November) and the third one in April 2010. They are all under the Dutch coordinate system. Point density of the data in March of 2008 is around 10 to 15 pts/m² (points per square meter), while the other two are about 30 to 50 pts/m². A part of the small island (Noordereiland) along the river in Rotterdam (Figure 1) was selected as study area where plenty of trees vary in size and shape.

The datasets have been classified into several predefined classes in which vegetation is one of them (green in Figure 1(a)). But the vegetation classification results have to be improved and misclassifications are to be dealt with. Normally, misclassification is categorized as commission and omission errors. By visual inspection, several kinds of commission errors can be recognized in the data set. Small segments like walls, roofs of complex shape on buildings as well as cars, poles and even some ground points are classified as vegetation. On the other hand, some vegetation points are classified as other classes like buildings and ground, which is omission error.

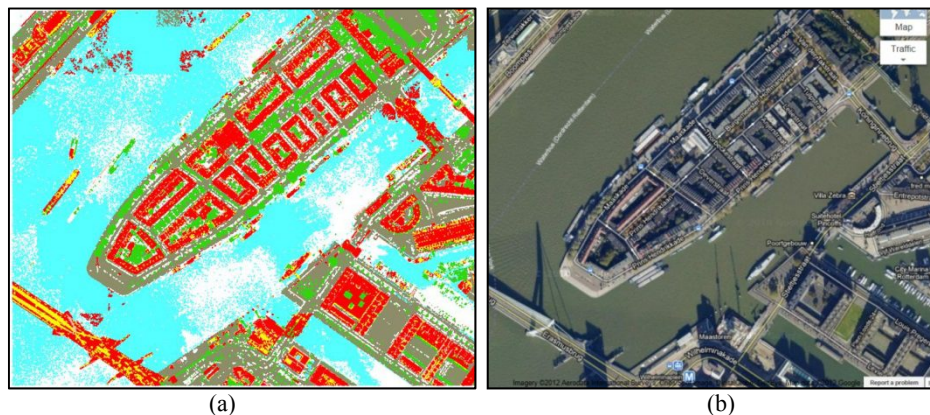


Figure 1. Study area in Rotterdam, (a) Lidar points (vegetation: green), (b) Google Map satellite image.

The overall accuracy of classification reached to 98.1%, while the completeness is 96.9% and the correctness is 97.8% in the study area¹¹. The omission errors are a few misclassified points which are minority with regard to the point amount of a tree. The error will hardly affect the parameter derivation result. So this paper focuses on commission error.

3. METHODS

Point clouds are processed by a sequence of algorithms. Firstly, the vegetation data are pre-processed by connected components algorithm, local maxima identification and trunk points' removal. Then two independent methods, point

based and model based, are implemented to derive tree crown parameters which are then assigned to each component for comparison. 3D alpha shapes algorithm is applied to thin and unified the point density. Corresponding trees are matched by the overlapping of bounding boxes and point to point distances. In the end visual inspection, comparison of methods and generic knowledge are proposed to assess the quality of the change detection results. The flowchart of the method is depicted as Figure 2.

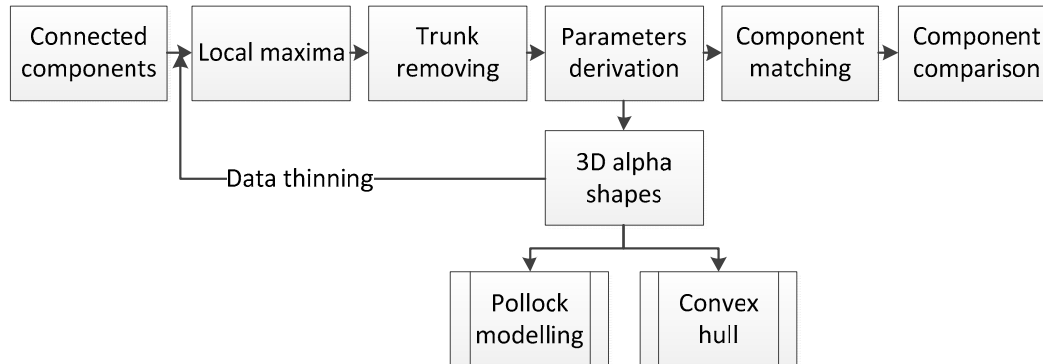


Figure 2. Flowchart of data processing.

3.1 Pre-processing

Connected components To eliminate the misclassified vegetation points and to group the points of a tree together as a component, connected components algorithm was implemented. The attributes of component were used to distinguish vegetation components from others. After classification, most of the misclassified components were small fragments. Therefore, the size of a component (number of points inside a component) was used to differentiate trees from fragments. The second attribute used was the height span of a component (the distance from the lowest point to the highest point). Components that had greater height span than an upper threshold or smaller height span than a lower threshold could be removed from the dataset. In order to remove components from high buildings or points hanging in the air, the minimum height of a component was utilized. Besides, normal distribution of the components was used to remove fragments that of regular shapes. In addition to the geometry attributes, spectral information, e.g. reflectance, intensity and true color were also used since, for example, the reflectance of trees is different from other objects like building roofs.

Local maxima After implementing connected components, some trees were connected together. It is important to identify the number of trees in each component, or at least label the component as single tree component or multiple tree component, namely the component contains one or more than one tree. For each point in a component, distances from other points were calculated. If the distance was less than a threshold range, which was adaptive to the component height, the heights were compared. If there was no higher neighbor point found, this point was the local maximum. So the number of highest points within a certain range was regarded as the number of trees in that component.

Trunk removing Because of the occlusion by the leaves and the scanning geometry, some of the trees have sufficient points on the trunks but others do not. This inconsistency will affect the process of parameter derivation and the result of change detection. So it is necessary to remove the points on the trunk before parameter derivation. For single tree components, the component was cut into slices by a certain length from the bottom upwards. For each slice, together with the previous slice below, a 2D bounding box was computed. If the hypotenuse of the bounding box was smaller than a predefined threshold, the points within the slice were considered as trunk points. The bounding box will keep moving upwards until the hypotenuse is greater than the threshold. In the end, a reference height will be obtained. The points above the height belong to the crown and the points below belong to the trunk. As for multiple tree components, the multiple trunks within the slice were separated using connected components algorithm. Then the bounding box was generated for each trunk component.

3.2 Tree parameter derivation

Alpha shapes Alpha shapes algorithm is famous for shape reconstruction from a dense unorganized set of points. Indeed, an alpha shape is a linear approximation of the original shape¹². The definition of alpha shapes is based on an underlying triangulation. As for 2D alpha shapes, circles with a certain radius (alpha) will approach the data points until they touch points on the edges of the triangles. As shown in Figure 3, the edges touched with circles describe an

approximate shape of the original points. For 3D alpha shapes, a triangulation is calculated first, and then spheres, instead of circles in 2D, will pass through the triangles. So the triangles that are touching spheres will represent the original shape of the data points. The points belonging to the vertices of the triangles are sufficient to describe the shape of a tree, so the points inside the alpha shape can be eliminated and the data will be reduced significantly. The area and volume of points can be calculated directly through the alpha shape.

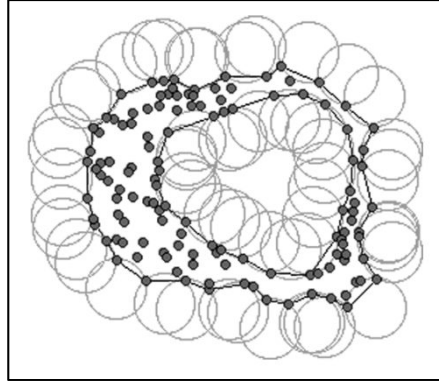


Figure 3. 2D alpha shapes¹³.

For multiple tree components, to reduce the effect of the connection of trees, the alpha values were optimized according to the shape of each component. The optimized alpha value was defined as a smallest alpha value such that the alpha complex has one solid component. So the gaps between trees were minimized. Then the areas and volumes of grouped tree components were extracted from the alpha shapes. For single tree components, a bigger alpha value was necessary because most of these components were small trees and optimized alpha shapes of small trees still had points inside the crown. In this case, the volume of the alpha shapes was actually much smaller than the real crown volume. Also if the alpha values are quite different from each other of the same tree from two epoch datasets, the change detection result will also be affected. To avoid the problems stated above, a consistent and big enough alpha value $(10m)^7$ was set for single tree components. If the alpha value is positive infinity $(\alpha \rightarrow \infty)$, the alpha shapes are the same as convex hulls. So the areas and volumes of single tree components were derived through convex hulls.

3D tree modeling 3D alpha shapes and convex hull are both point based parameter derivation method which has both advantages and disadvantages. Point based method is straight forward, simple to process but sensitive to noises and outliers. On the contrary, model based method is more stable. These two methods can be compared so as to assess the quality of the results. In this paper, single tree components were modeled and the results were compared with point based method.

A parametric model developed by Pollock¹⁴ was used to model trees of different shapes. The shape of the crown can be adjusted by a parameter from an ellipsoid to a cone:

$$\frac{z^n}{a^n} + \frac{(x^2 + y^2)^{\frac{n}{2}}}{b^n} = 1 \quad (1)$$

The origin is the center of the circle of the crown, and z-axis points vertically upwards. In the equation a is the radius of the intersection of surface with the z-axis, b is the radius of the circle of the crown and n is a positive real number that determines the shape of the crown surface. When n=2 the surface is an ellipsoid, and as n decreases to 1 the surface becomes a cone. To make the model more realistic as a tree crown, the base was changed to an ellipse instead of a circle, and then a rotation in x-y plane and the shift of coordinates were added as following.

$$\frac{z^n}{c^n} + \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} \right)^{\frac{n}{2}} = 1 \quad (2)$$

$$x = (X - X_0) \cos \beta - (Y - Y_0) \sin \beta \quad (3)$$

$$y = (X - X_0) \sin \beta + (Y - Y_0) \cos \beta \quad (4)$$

In which, a , b are the two semi axes of the crown base ellipse, c is the semi axis in z direction, n is the real number that determine the crown shape, X_0 , Y_0 , Z_0 are coordinates of the origin in the global coordinate system and β is the rotation angle.

The adjusted Pollock model was implemented using nonlinear least square fitting in three steps: 2D crown base fitting, upper crown fitting and lower crown fitting (Figure 4). The crown was divided into upper crown and lower crown because they have different shapes. The crown base (a , b , β) was estimated first in 2D, then a and b were used as constraints for the 3D Pollock model fitting.

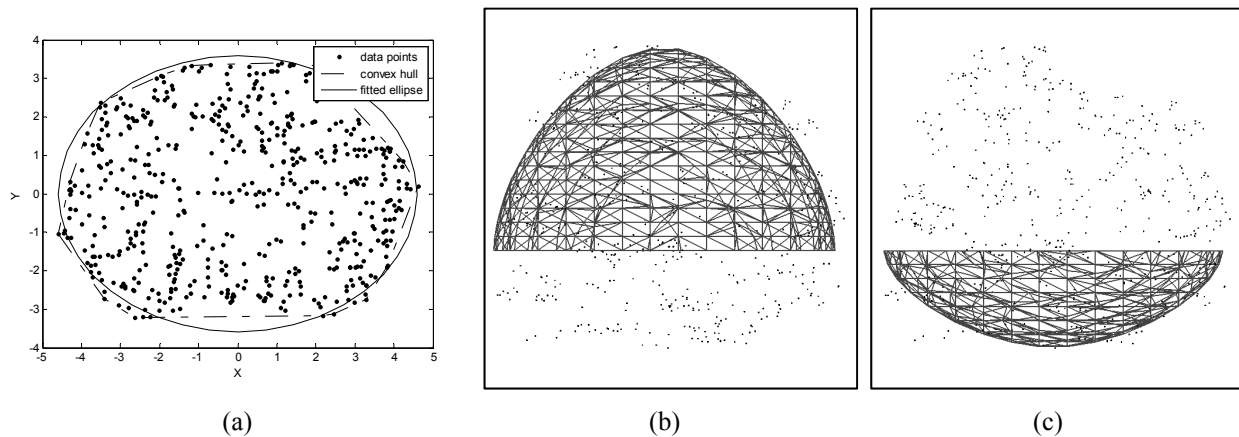


Figure 4. Adjusted Pollock model fitting, (a) 2D crown base fitting, (b) upper crown fitting, (c) lower crown fitting.

After the 3D model fitting, the height of the tree was the crown height plus the vertical axis ($Z_0 + c$). The crown area was the area of the fitted ellipse (πab) and the crown volume was the sum of the volumes of the upper and lower crown.

3.3 Tree oriented change detection

After deriving the parameters of the trees in each data set, the corresponding trees were identified based on the locations of trees. This point based tree to tree matching was accomplished by calculating the overlapping of bounding boxes and point to point distances. First of all, for each of the components in dataset 1, a bounding box was derived. Then the overlapping bounding box in dataset 2 was searched. To further check whether these two bounding boxes were the corresponding components, the distances from points in dataset 1 to points in dataset 2 were calculated. If the number of distances that were smaller than 1m was greater than a certain percentage of the smaller size of the two components, these two components were matched. Because it happens that several components in dataset 1 correspond to one component in dataset 2. So the number of nearby points is compared with the smaller one between the two components. The corresponding components were given the same label in both dataset. If there were no corresponding components found in the other dataset, they were not labeled.

Four categories were introduced to analyze the changes: cut, newly planted, area change and volume change. The latter two were further divided into increase and decrease. As explained above, the corresponding relation between the components in two datasets may be many-to-one. In this case, the parameters of components were added up and then compared with the component in the other dataset. Since it is meaningless to compare the change of height of multiple tree components, the height of tree is not compared.

4. RESULTS AND ANALYSIS

4.1 Pre-processing results

After connected components, each component had its own component number which is illustrated by color as Figure 5(a). Figure 5(b) shows the result of non-vegetation components removing. Majority of the remaining components were trees. Because even after components removing, some non-vegetation or bush components still remained in the data, the tree components along the roads were selected manually since they are important for end users like urban planners. Figure

5(c) illustrates the subset and the results of local maxima. The threshold of the range for local maxima was set to a quarter of the absolute height of each component according to experimental tests.

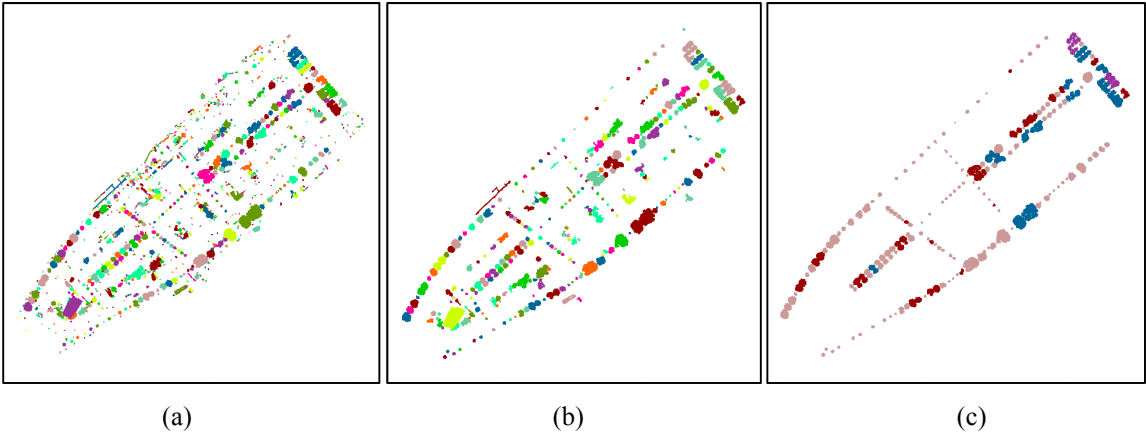


Figure 5. Preprocessing results, (a) connected components (component number illustrated by a random color), (b) components removing, (c) local maxima (reddish yellow: 1 component; dark red: 2; blue: 3; purple: 5 components).

Non-vegetation components and small bush components are to be removed because only high vegetation (trees) components are the objects of this research. However some small trees are removed in the meantime and some non-tree components are not. So completeness and correctness¹⁵ are used to assess the quality of this step. True positive (TP) is the real tree component after removing; false positive (FP) is the remaining incorrectly classified tree component and false negative (FN) is the real tree component that have been removed. The completeness and correctness of the result are assessed by visual inspection.

$$\text{Completeness} = TP / (TP + FN) \tag{5}$$

$$\text{Correctness} = TP / (TP + FP) \tag{6}$$

$$\text{Quality} = TP / (TP + FN + FP) \tag{7}$$

The number of components of the three datasets before and after removing and the accuracy results are shown as Table 1. The first dataset had the lowest overall quality; especially the correctness was quite low. That was mainly affected by classification because the classification result of the first dataset was not as good as the other two.

Table 1. Number of components and the accuracy analysis results.

Dataset	2008.03	2008.11	2010.04
Before removing	1451	1169	2118
After removing	306	229	275
Completeness	93%	97%	98%
Correctness	86%	97%	93%
Overall Quality	81%	94%	91%

4.2 Tree models and derived parameters

Because the upper crown normally has more points and is bigger than the lower crown, the crown position is fixed using the upper crown so that the whole model has a continuous smooth surface which can be directly used for 3D virtual city visualization. Figure 6 shows the modeling results of different shapes and sizes.

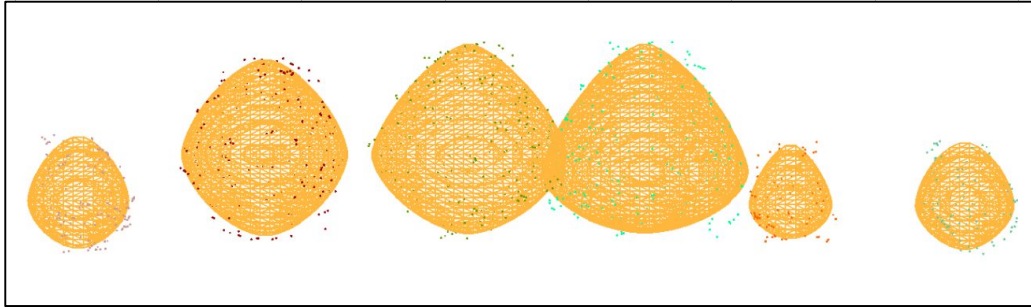


Figure 6. 3D modeling results.

The parameters were derived through both point based and model based methods. To find out the consistency of these two methods, the difference between the two results with respect to the Pollock model results in the dataset March 2008 was computed. The standard deviation of the difference of area was 0.0721, and standard deviation of the difference of volume was 0.0764.

Moreover, the differences of change (model change minus convex hull change) with respect to the Pollock model results (normalized by model change) between March 2008 and April 2010 in which there were 78 corresponding single tree components were calculated. The standard deviation of the differences in area was 0.0607 and 1.5 times the half of the interquartile was 0.0518. The standard deviation of the differences in volume was 0.0985 and 1.5 times the half of the interquartile was 0.0789. Both of the values were under 10% which illustrated the consistency of these two methods.

4.3 Detected changes and analysis

The extracted parameters were then assigned back to each component as its features so that the components can be compared and visualized.

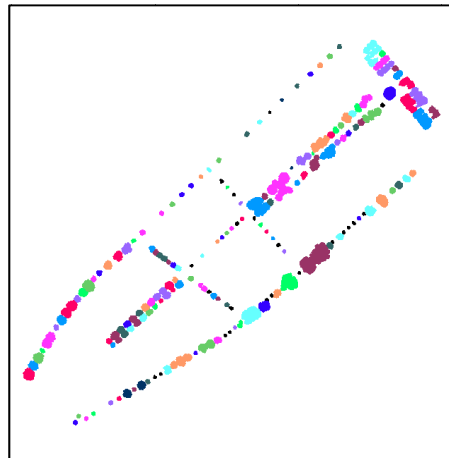


Figure 7. Components matching result (cut trees-dark blue; planted trees-black; matched components in other random colors).

Two datasets of different epochs were merged together. Corresponding components have the same label and different matches have various labels shown by colors (Figure 7). Components that have been cut are in dark blue color whereas newly planted components are in black.

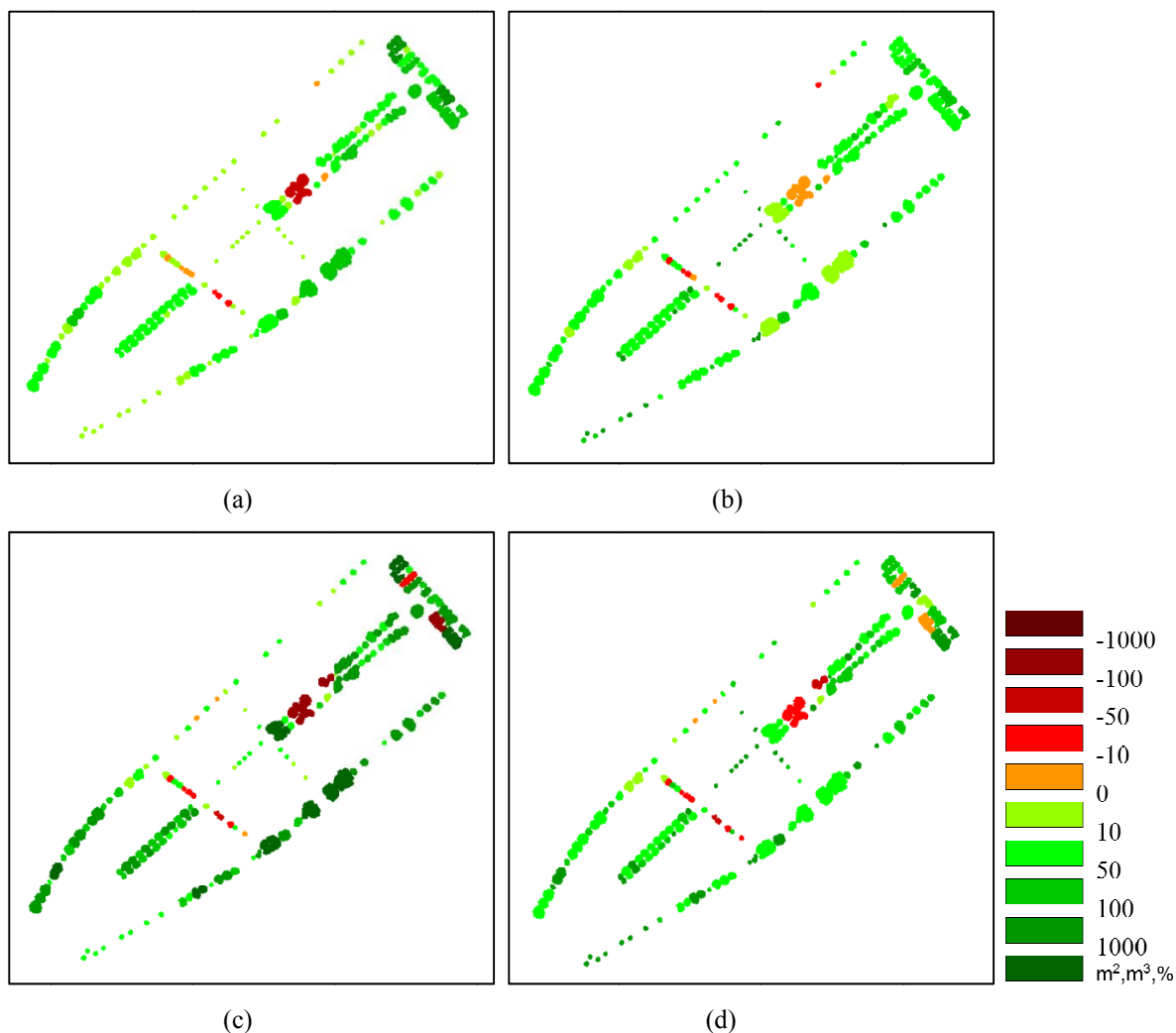


Figure 8. Change detection results (2008.03-2010.04), (a) area change (m^2), (b) area change ratio (%), (c) volume change (m^3), (d) volume change ratio (%).

Each dataset was compared with the other two among the three datasets. Figure 8 (a) and (c) show the changes of areas and volumes between March 2008 and April 2010, in which a sequence of colors illustrate different changes. Figure 8 (b) and (d) show the change ratios (normalized change) for area and volume. Most of the components increase in area and volume. Similar size components have similar changes. Smaller components have equal or bigger change ratios. Table 2 shows the detected changes among three datasets.

Table 2. Change detection results.

Categories	Cut	Planted	Area change		Volume change	
Change	Only in data1	Only in data2	Area \uparrow	Area \downarrow	Volume \uparrow	Volume \downarrow
08.03-10.04	4	22	132	9	128	13
08.03-08.11	3	20	131	1	129	3
08.11-10.04	3	4	44	102	40	106

According to the change detection results, many tree components are detected as planted in the latter two datasets compared with the first one. One reason is that some of the trees are indeed planted, but more likely, that is caused by omission error. Some real tree components in the first dataset are incorrectly removed because the component size and point density are too small even though the threshold of component size is proportional to the density.

For the first comparison, 12 components' change ratios out of 141 are smaller than 10 per cent (positive and negative), which actually cannot be assured by the parameter derivation methods. Considering the errors of pre-processing and modeling, every tree pair will be slightly different even if no changes happened. So a very small proportion of change should be taken care of. 13 components' volumes have decreased while 9 components' areas have decreased which means 4 components' volumes have decreased but the areas have increased. It can be true but also can be caused by the errors of parameters since most of the parameters of these components are close to zero.

Completeness and correctness errors will end up with changes as cut or planted, thus these two categories are verified through the original vegetation points. The change result between March 2008 and April 2010 is visually inspected. 3 out of 4 cut trees and only 2 out of 22 planted trees are confirmed. The results of cut or planted are severely affected by the accuracy of classification and quality of pre-processing.

Multiple tree components have greater change per tree than single tree components, even though 3D alpha shapes minimize the gaps between connected trees. The gaps will be bigger when trees grow. So the changes of trees actually include the changes of gaps. In other cases, the matching relations are many-to-one, i.e. two single tree components in the first dataset to one big component in the second dataset. Then the change detection results will be bigger than real because first two components are computed individually however the second one takes the gap into consideration.

In general, small trees have less changes in area and volume compared with bigger trees and trees of similar sizes have the same behavior of changes, which have been reflected by the change results. Moreover, the change ratios of small trees are the same and even greater than bigger trees because in principle smaller trees grow faster. Thus even without ground truth, these generic knowledge assess the results and suggest the feasibility of the method.

5. CONCLUSION

Components features are feasible for non-tree components removing especially when considering both geometrical and spectral features. 3D alpha shapes thin the datasets and more importantly reserve the points on the vertex which can be directly used for 3D tree modeling and other parameter derivation method. Also the area and volume can be calculated through the alpha shape. Datasets of different point densities are unified using a same alpha value. The Adjusted Pollock model shows an obvious advantage for 3D tree modeling by providing the crown shape parameter n . So every single tree has its own crown shape, which is more realistic. The separation of upper crown fitting and lower crown fitting is proven feasible and the results have high linear correlation with convex hull which is based directly on the points. The standard deviation of the differences between Pollock model and convex hull is about 8%. Also the differences of changes between these two methods are under 10%. These suggest that adjusted Pollock model have great potential of modeling trees accurately and vividly for a 3D virtual city. The proposed method provides a guideline for change detection of trees in multi-temporal airborne lidar point clouds. The growth and pruning of trees are successfully detected. Moreover, the program processes the datasets semi-automatically.

Datasets from different scan strips can help to further verify the parameters and results of changes. Following research may focus on accurate individual tree delineation from multi-tree components. Other types' datasets are helpful to improve the research result. The imagery obtained simultaneously with the point clouds is useful to identify the commission and omission errors at the level of trees. Mobile mapping system (MMS) is capable of capturing the details of the trunk and crown of trees. So the combination of mobile and airborne laser scanning data will facilitate the modeling of trees with accurate trunk information.

ACKNOWLEDGEMENTS

The authors wish to thank Fugro Aerial Mapping for acquiring and the Municipality of Rotterdam, the Netherlands for providing the lidar data.

REFERENCES

- [1] Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B. and Lambin, E., "Digital change detection methods in ecosystem monitoring: a review," *Int J Remote Sens*, Papers 25, 1565-1596 (2004).
- [2] Houldcroft, C. J., Campbell, C. L., Davenport, I. J., Gurney, R. J. and Holden, N., "Measurement of canopy geometry characteristics using LiDAR laser altimetry: a feasibility study," *Geoscience and Remote Sensing, IEEE Transactions on*, Papers 43, 2270-2282 (2005).
- [3] Yu, X., Hyypä, J., Kukko, A., Maltamo, M. and Kaartinen, H., "Change detection techniques for canopy height growth measurements using airborne laser scanner data," *Photogramm Eng Rem S*, Papers 72, 1330-1348 (2006).
- [4] Rutzinger, M., Höfle, B. and Pfeifer, N., "Detection of high urban vegetation with airborne laser scanning data," *Proc. ForestSAT07, OnCDROM* (2007).
- [5] Yu, X., Hyypä, J., Vastaranta, M., Holopainen, M. and Viitala, R., "Predicting individual tree attributes from airborne laser point clouds based on the random forests technique," *Isprs J Photogramm*, Papers 66, 28-37 (2011).
- [6] Yu, X., Hyypä, J., Kaartinen, H. and Maltamo, M., "Automatic detection of harvested trees and determination of forest growth using airborne laser scanning," *Remote Sens Environ*, Papers 90, 451-462 (2004).
- [7] Rutzinger, M., Pratihast, A. K., Oude Elberink, S. J. and Vosselman, G., "Tree modelling from mobile laser scanning data-sets," *The Photogrammetric Record*, Papers 26, 361-372 (2011).
- [8] Wang, Y., Weinacker, H. and Koch, B., "A lidar point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest," *Sensors*, Papers 8, 3938-3951 (2008).
- [9] Vosselman, G., "3d reconstruction of roads and trees for city modelling," *International Archives of Photogrammetry and Remote Sensing*, Papers 34, 231-236 (2003).
- [10] Kato, A., Moskal, L. M., Schiess, P., Swanson, M. E., Calhoun, D. and Stuetzle, W., "Capturing tree crown formation through implicit surface reconstruction using airborne lidar data," *Remote Sens Environ*, Papers 113, 1148-1162 (2009).
- [11] Xu, S., Oude Elberink, S. J. and Vosselman, G., "Entities and features for classification of airborne laser scanning data in urban area," *Proc. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (Accepted in 2012).
- [12] Edelsbrunner, H., Kirkpatrick, D. and Seidel, R., "On the shape of a set of points in the plane," *IEEE Transactions on Information Theory*, Papers 29, 551-559 (1983).
- [13] Da, T. K. F., Loriot, S. and Yvinec, M., "CGAL User and Reference Manual," 18th October 2011, <http://www.cgal.org/Manual/latest/>
- [14] Pollock, R. J., "Model-based approach to automatically locating tree crowns in high spatial resolution images," *Proc. SPIE*, 526-537 (1994).
- [15] Heipke, C., Mayer, H., Wiedemann, C. and Jamet, O., "Evaluation of automatic road extraction," *International Archives of Photogrammetry and Remote Sensing*, Papers 32, 151-160 (1997).