Preprocessing of Airborne Laser Scanner Data for Tree Extraction in Forest Area

Tatsuya Nawamura, Takahiro Endo, Masaki Teraoka, Yoshifumi Yasuoka
Yasuoka Laboratory, Institute of Industrial Science, University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo, 153-8505 JAPAN
nawamura@iis.u-tokyo.ac.jp

Masahiro Setojima
Kokusai Kogyo Co., Ltd.
5 Sanbancho, Chiyoda-ku, Tokyo, 102-0075 JAPAN,

Toru Katsura
Mitsubishi Paper Mills Co., Ltd.
3-4-2, Marunouchi, Chiyoda-ku, Tokyo, 100-0005 JAPAN

Abstract: As the Kyoto Protocol entered into force since February 2005, the reduction of carbon dioxide emission becomes a serious global issue involving developing and developed countries. As a consequence, it is necessary to estimate timber volume in a forest in order to grasp CO_2 absorption. The estimation is, however, not easy in extensive forest areas only with the conventional ground survey methods. Remote sensing is expected to be one of the effective tools for the estimation of CO_2 because of its quantitative and extensive measurement capability.

In this study, a preprocessing for extraction of individual trees by ALS (Airborne Laser Scanner) was investigated. In particular, a noise removal method in ALS point cloud data was developed to reduce tree extraction errors because there are many sudden drops (spiky noise) where laser beam penetrated into the tree canopy and they disturb the extraction of trees and the estimation of a timber volume. After the noise removal, quadratic surface fitting algorism was applied to extract individual trees from forest areas and to estimate their parameters including the number of trees, the crown height and the crown width. Results of the parameters from data with and without the noise removal process were compared. The results of quadratic surface fitting showed the effectiveness of our method with improved correlation coefficient and RMSE of 0.62 to 0.65 and 1.84 to 1.82, respectively. However, in terms of comparison of the physical parameters of the extraction rate of number of trees, the crown width and the crown height, the results changed from 0.88 to 0.75, 0.01 to 0.53 and 0.32 to 0.27 respectively, and some difficulties were pointed out.
Keywords: Japanese cedar, quadric surface parabola, noise removal, crown.

1. Introduction

Kyoto Protocol is an international treaty that stipulates targets for the developed countries' for GHG reduction [1]. In December 2003, the ninth meeting of the conference of the parties (COP9) has accepted afforestation and reforestation projects under the clean development mechanism (CDM), for the first commitment period (2008-2012) of the Kyoto protocol. Methodologies for monitoring GHG reduction certified by such projects at a regional or a local level is under active discussion with various approaches.

A surface area of canopy is one of the important parameters to estimate the CO_2 absorption by forests, and the remote sensing is an effective method, and techniques for monitoring the forest parameters at a regional level are broadly developed. Delineation methodologies of individual crowns with use of aerial images [2], high-resolution satellites images or LIDAR (Light detection and ranging) are major in such scale [3] [4]. Also in a regional scale, estimation of carbon biomass is proposed using LIDAR and aerial images [5].

In these previous studies, the pulse density of LIDAR is getting higher and higher in accordance with the technological development. However, in such high-density LIDAR data, some laser pulses returned from objects within a crown and they are recorded as sudden drops compared to surrounding points returned from the surface of a crown. Those sudden drops affect the estimation results. Usually, filters such as smoothing, median, and/or average are applied to remove such sudden drops. In additional to those existing methods, we proposed a new method for removing such noises.

This paper aims at development of methodology for removing sudden drops in a tree crown. Then, the quadric surface fitting was applied to study area, and validations of results are examined by comparing the outcomes from the raw data and the data after the removal process.

2. Data acquisition

1) Study Area

Our study area is part of a property of Mitsubishi Paper Mills Co., Ltd with total area of about 160 ha, and is located in Aomori prefecture, the northern part of Honshu Island (Fig. 1, 40°39' N, 141°5' E, 190 - 240 m above sea level). A study area, representing typical planted forests in Japan, was chosen from the study area regarding to the tree types, years from plantation, and thinning histories (Fig. 2). Dominant tree species is Japanese cedar (*Cryptomeria japonica*), forming cylinder cones in matured status.

2) LIDAR Data

LIDAR is a growing technology that can acquire height information of ground surface by measuring the time interval between the emission and reception of laser pulses. The data was acquired in 11 and 12 August 2004 from an airborne laser scanner with the cruising speed of 110 kt, the height of 1830 meter, the scan rate of 39.0 Hz, the pulse rate of 46.0 kHz, the scanning width of 647 meters, and the average footprint of 0.47 meter. The scanner was ALS50, a product of Leica Geosystems, which has the horizontal accuracy of ± 30 cm and the vertical accuracy of ± 15 cm. The average density of laser reflections was 14.65 pulses per square meter, and it could be equally expressed that every 0.10 square meter has at least one pulse in the average.

3) Field Survey

A field survey was conducted in the first week of August 2004. For each tree, the position, the height, the diameter at breast height (DBH), and the species were obtained as the ground truth. Precise positions of every tree were specified using the differential GPS, and the height were acquired with the laser measurement instrument. Information of DBH and the species were based on the actual measurement. Detail of the site is shown in table 1. The additional survey was carried out on 11 July 2005 for the measurement of crown width and height of crown. The crown width is measured from ground level using a survey tape and the average of N-S and E-W directions was calculated.

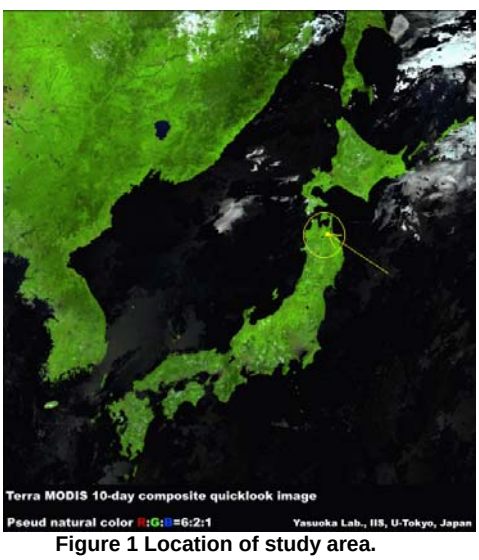

Figure 1 Location of study area.

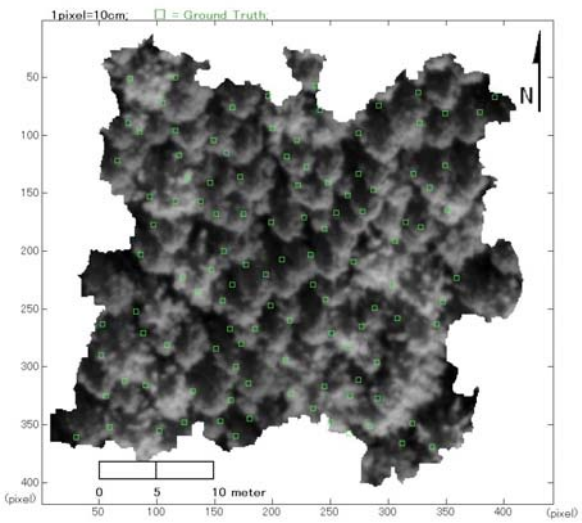

Figure 2 Aerial near-infrared photograph of study area.

Table 1. Summary of the field measurements in the study area.

Area [ha]	Number of trees	Number of trees per hectare	Year of plantation	Average height [m] (Standard deviation)	Average diameter at breast height [cm] (Standard deviation)
0.16	**118**	**732**	**1955**	**21.0 (3.0)**	**26.5 (5.8)**

3. Methodology

1) Removing Sudden Drops in LIDAR Data

LIDAR data from a foliage crown area contains some sudden drops compared to the surface of the crown because of the penetration of laser pulses through the crown. They are not error data since the laser pulses are returned from inner objects, but in the process of extraction of an individual tree crown and estimation of a timber volume, those sudden drops affect the results. In this study, a following algorithm was developed and applied to overcome such difficulty. Firstly, three-dimensional triangular irregular network (TIN) based on XY-plane is calculated from LIDAR data (Fig. 3). Then an index *d* is calculated for every LIDAR data in order to extract the sudden drops.

$$d = average\left(\frac{r_i}{average(r_i)} \cdot z_i\right) - z_0 \tag{1}$$

where z_i is the height of the LIDAR data around the target point, and z_0 is the height of the target points, r_i is the reciprocal of distance between the target and the surrounding points on XY-plane. The index *d* is the difference between z_0 and the average of z_i that is weighted by XY distance from *0* to each *i*. In this study, the threshold was defined empirically, and the points with the *d* value over 1.0 meter were removed from the LIDAR data.

2) Template Matching by Quadric Surface Parabola

A crown form is represented as a quadric surface parabola, generalized ellipsoids following Sheng [6], where the space occupied by the foliage within an individual tree crown is determined by six parameters: 2-D coordinates of the crown top (X, Y), the tree height (cz), the crown width (cw), the crown height (ch), and the crown curvature (cc). The surface of a tree crown is then given by the following mathematical expression: (2)

$$\frac{(cz+ch)^{cc}}{ch^{cc}} + \frac{(X^2+Y^2)^{cc/2}}{cw^{cc}} = 1 \tag{2}$$

A template parabola is defined with following parameters: cw=2.3m, ch=2.3m, cc=1.0, and the correlation coefficient (R) is calculated by each XY pixels with respect to Z-axis. Fig. 4 shows the R-value of each XY pixels (R-map).

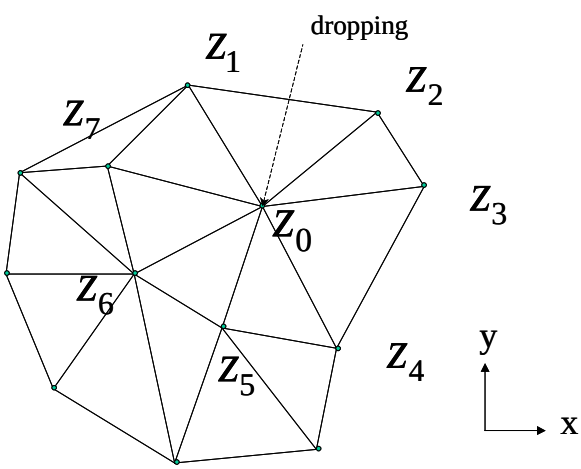

Figure 3 Removing a sudden dropping point by using TIN and height data.

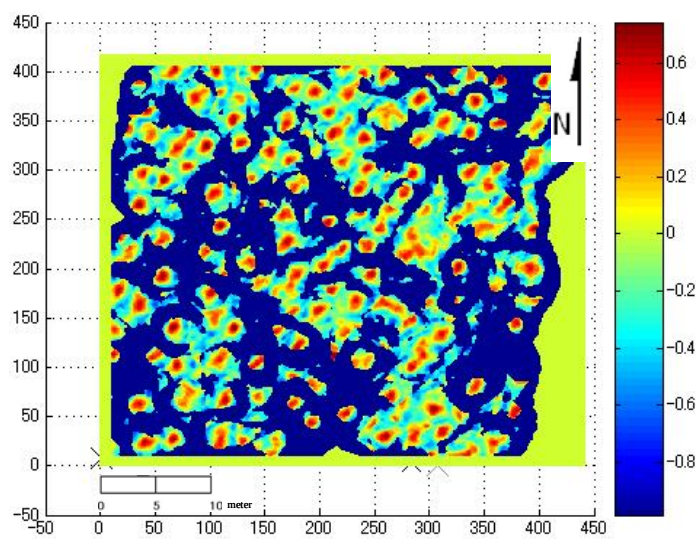

Figure 4 Correlation coefficient (R) between template and LIDAR data (R-map).

3) Fitting of Quadric Surface Parabola

After the template matching, fitting operation is carried out from XY-positions with the highest R-value. Then, parameters are examined with the combination of cw of 0.9-4.1 m by every 0.4 m, ch of 1.5-5.5 m every by 0.5m, cc of 0.8-2.0 by every 0.4, respectively to produce candidates for the specific position. A final candidate is defined by the highest R-value. And then, a projected area of a final candidate in R-map is blanked and one crown model (parabola) is extracted. This is one cycle to make parabola model and this process continues till there are no R over a certain value in the map. In this case, 0.3 was set as the threshold. Through these processes, a series of parabolas are extracted, which represent the state of the forest surface (Fig. 5).

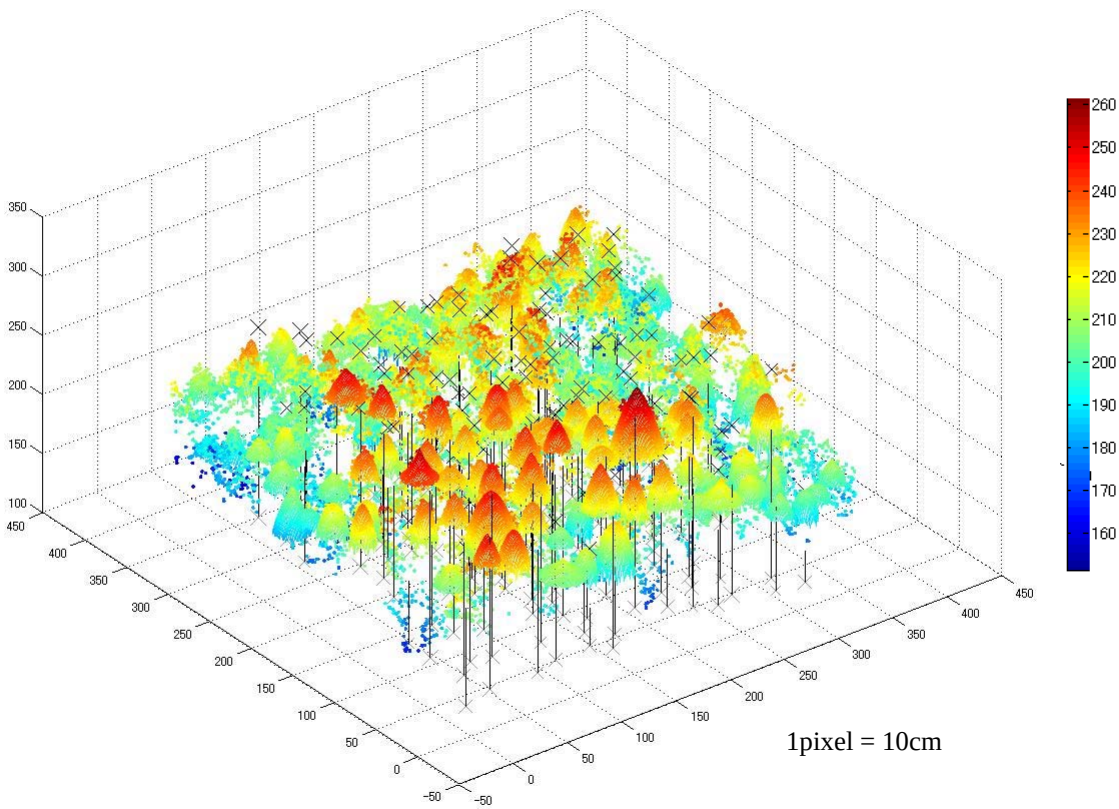

Figure 5 Three-dimensional perspective view of the estimation results.

4. Results and Discussion

1) Comparison of the parabola fitting results between the raw data and the processed data

The results of parabola fitting between the raw data and the processed data are compared by examining the R and the root mean square error (RMSE) both in three-dimensional space. Figure 6 shows the line plot of the R and the RMSE. Comparing the raw data and processed data, the average value of the R and the RMSE has improved from 0.62 to 0.65 and 1.84 to 1.82 respectively. This shows the removal of sudden drops by our method contributes to the parabola fitting.

2) Comparison of the number, the width and the height of crown between ground truths and estimation results

Figure 7 shows the extracted parabolas and the ground truth tree positions by GPS measurement on the XY-planes. The numbers of extracted parabola of the raw data and processed data were 104 and 89 respectively, and the extraction rates were 0.88 and 0.75 respectively. The decrease of extraction rate after applying our method was due to the decrease of LIDAR points as a result of the removal process. In our quadratic surface parabola fitting process, a skip of the fitting is implemented with the threshold with the number of LIDAR points less than 7.33. The threshold is the half of the average density of LIDAR points of the raw data (14.65 points / m^2), and the same threshold is applied for the both fitting processes of the raw and the processed data. In some areas, the number of LIDAR points decreased by our removal process and a skip of the fitting process could occur. The quadric surface fitting methodology could be developed depending on the number of the points for both the raw and the processed data.

Physical parameters of trees, namely crown width and crown height, are then compared between ground truth and the estimation results by identifying corresponding ground truth trees by the visual inspection. The parabola width (pw) and the ground truth tree crown width (cw), and the vertical parabola height (ph) and the ground truth tree crown height (ch) are scattered in Fig. 8, for each raw and processed data.

The numbers of the R between pw-cw and ph-ch were changed from 0.01 to 0.53 and 0.32 to 0.27 respectively by applying our method. The result of the width shows that the R after the estimation is improved, however, not drastically as we expected. On the other hand the result of the height did not improve and the value of R stayed in the low correlation coefficient.

In the actual forest, a surface area of a crown that is substantially exposed to the sunlight (sun-lighted area) is heavily relying on the neighborhood trees by overlapping to each other crown. In our study, an extracted parabola is considered to be estimating the sun-lighted areas. The validation with cw and ch, measured from the ground level, might be not enough to explain such sun-lighted areas.

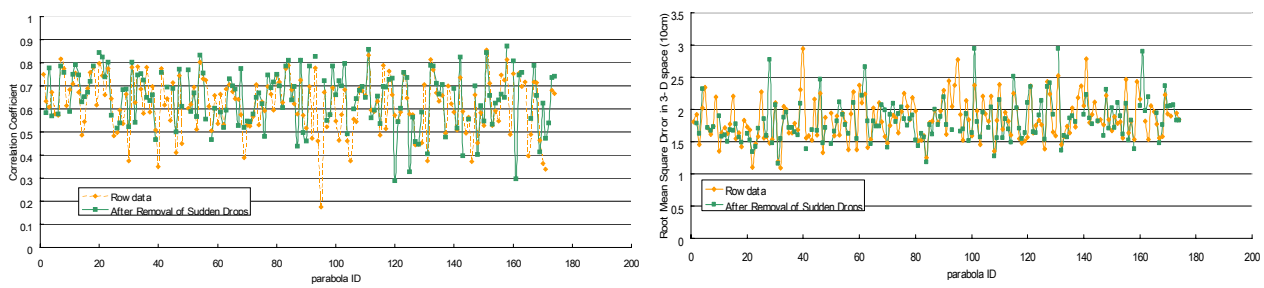

Figure 6 Line plot of correlation coefficient (left) and RMSE distance (right) with respect to nth of parabola fitting.

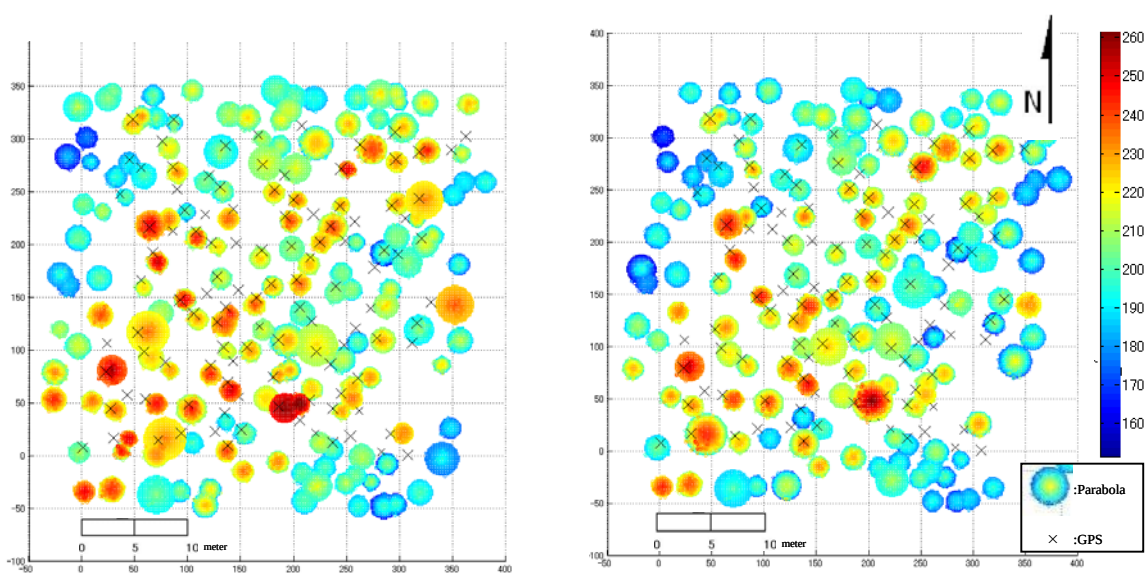

Figure 7 Extracted parabolas and the ground truth position of trees by GPS on XY-plane using raw data (left) and using processed data (right).

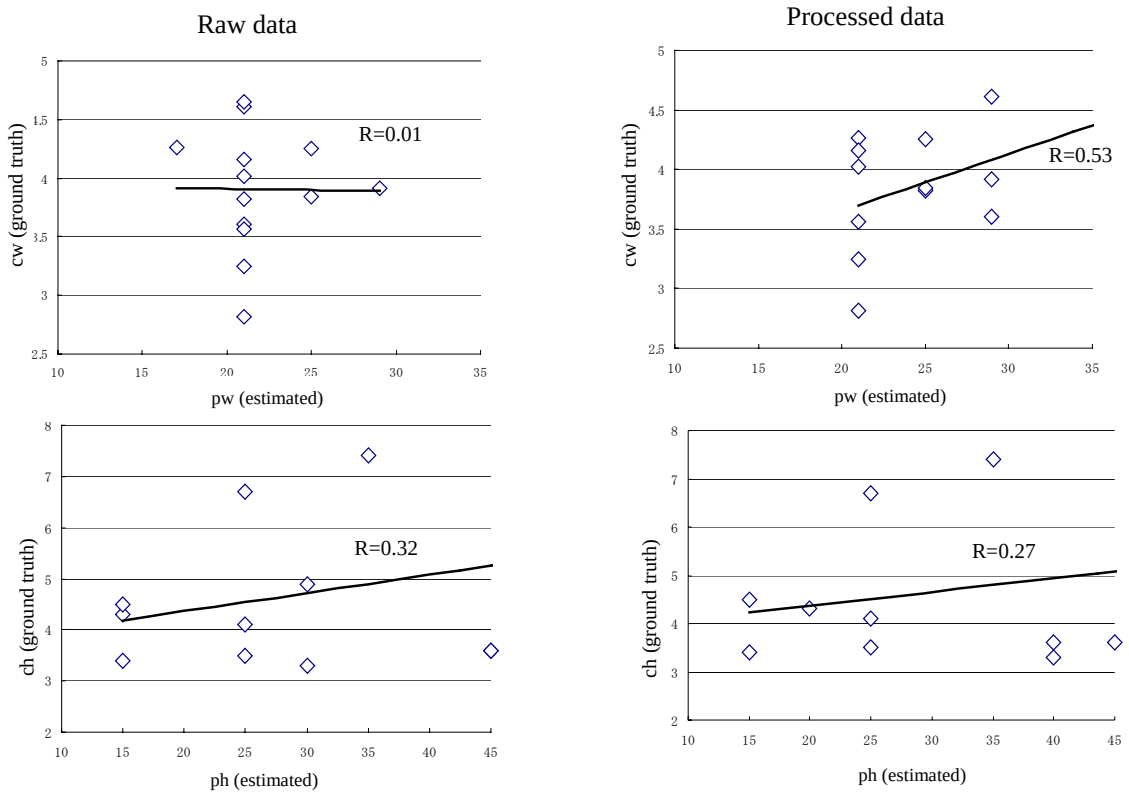

Figure 8 Scatter plots of pw-cw (upper), ph-ch (bottom) using raw data (left) and processed data (right).

5. Conclusions

We developed a methodology for the removal of sudden drops in a crown, and showed the effectiveness of our method with the validation of improved R and RMSE of 0.62 to 0.65 and 1.84 to 1.82, respectively. However, in terms of comparison of the physical parameters of the extraction rate of number of trees, the crown width and the crown height, the results changed from 0.88 to 0.75, 0.01 to 0.53 and 0.32 to 0.27 respectively. The comparison results did not improved as drastically as we expected. Future works remain in the improvement of parabola fitting process regarding the decrease in the number of LIDAR data points after the proposed method. Also the validation between the ground truths and the estimated results with a consideration of substantial sun-lighted crown areas should be developed.

Acknowledgement

Author would like to thank co-workers of Kokusai Kogyo Co., Ltd, and Mitsubishi CO., Ltd, and the all members of Yasuoka laboratory for participating in the field survey.

References

[1] **URL**: United Nations Framework Convention on Climate Change: Kyoto Protocol, 1997. Available at: http://unfccc.int/.
[2] Gougeon, F., 1995. A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. *Canadian J. of Remote Sensing, 21(3)*. pp. 274-284.
[3] Andersen, H.E., Reutebuch, S.E. & Schreuder, G.F., 2002. Bayesian Object Recognition for the Analysis of Complex Forest Scenes in Airborne Laser Scanner Data. *ISPRS Commission III, Symposium 2002*, pp. A-035 ff.
[4] Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J., Dawson, T.P., 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, 36, pp. 368-380.
[5] Teraoka, M., Endo, T., Setojima, M., Katsura, T., Yasuoka, Y., 2004. Estimation of Physical Parameters of Individual Trees by Data Fusion of Aerial Images and Laser Scanning Data, *Proceedings of the 25th Asian Conference on Remote Sensing*, Chiangmai THAILAND, CD-ROM, 2004
[6] Sheng, Y., Gong, P., and Biging, G., 2001. Model-based conifer crown surface reconstruction from high-resolution aerial images. *Photogrammetric Engineering and Remote Sensing*, 67(8), pp.957-965.