Available online at www.sciencedirect.com
SCIENCE DIRECT®

Advances in Space Research 33 (2004) 317–322

ADVANCES IN SPACE RESEARCH
(a COSPAR publication)

www.elsevier.com/locate/asr

Automatic building extraction from laser scanning data: an input tool for disaster management

J. Dash [a,*], E. Steinle [b], R.P. Singh [c], H.P. Bähr [b]

[a] School of Geography, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[b] Institute for Photogrammetry and Remote Sensing, University of Karlsruhe, Karlsruhe, Germany
[c] Department of Civil Engineering, Indian Institute of Technology, Kanpur 208 016, India

Abstract

Estimation of damages caused by a disaster is a major task in the post disaster mitigation process. To enhance the relief and rescue operation in the affected area it is required to get a near real time damage model. For this purpose a fast method of data acquisition with suitable methods for extracting the man-made objects is required. Laser scanning data provide the height of the ground objects, which can be used for developing models to extract the man-made features in a complex urban environment. Using the height variation along the periphery of objects present in the data, a method based on standard deviation was developed to distinguish between tree and building.

Keywords: Laser scanning; Disaster management; Building extraction

1. Introduction

Natural disaster always leaves a devastating impact on human life especially in the developing countries, where its effect coupled with existing problems of poverty, lack of infrastructures and communication hinders the overall growth and development process. Although it is almost impossible to control the occurrence of natural disaster but its prediction and preparedness along with an effective post disaster management programme can minimise the risk (Jayaraman et al., 1997). Remote sensing techniques either by space-borne or air-borne sensors were used in last decades to detect, identify and monitor the impact and effect of natural disasters like earthquakes (Massonnet, 1995), landslides (Mantovani et al., 1996) and floods (Cobby et al., 2001). It can be further used to provide some useful information to continue the relief, rescue and rehabilitation operation.

Airborne laser scanning (ALS) is now gaining importance in research and commercial use particularly in the field of mapping and monitoring environmental and ecological changes (Drake et al., 2002; Irish and Lillycrop, 1999), urban planning especially development of 3D city models (Haala and Brenner, 1999; Maas and Vosselman, 1999), disaster management (Cobby et al., 2001; Steinle et al., 2001). Facilities for acquiring data both in day and night including high planimetric resolution (up to 1 m), with high accuracy in position and partly weather independency are some of the features which enable users to use them in a wide range of applications (Ackermann, 1999; Wehr and Lohr, 1999).

Natural disaster engulfs a large number of people which depends on factor like population density, type and extent of damage caused by the disaster, time of the day, etc. ALS data can be used to find out the type and extent of damage by extracting the buildings and other man-made objects and comparing their status in pre and post disaster data, therefore it is helpful to produce a near real time damage model (Steinle et al., 2001). ALS data can be used for two purposes: First to generate a model depicting the buildings of the area by using the post disaster data and comparing with the pre disaster database to find the damage model. The model should convey information about the type, shape and other related information of the building. It would be helpful in accessing the damage. Second to find the shortest

* Corresponding author.
E-mail address: jadu@soton.ac.uk (J. Dash).

0273-1177/$30
doi:10.1016/S0273-1177(03)00482-4

possible path for the relief and rescue teams to reach the needy as many of the roads get blocked by rubble and debris. An automatic method for extracting buildings was therefore generated using only ALS data with the intension to aid the above process in tackling disaster management.

2. Laser scanning system and data used

LASER stands for "*Light amplification by stimulated emission of radiation*". In this process an atom is simulated by an external atom, as a result it gains some energy and move to an exited stage. While coming to its normal stage, the resulting energy in turn produces light of certain wavelengths (Maiman, 1960). ALS systems are active remote sensing systems using a laser pulse as the sensing carrier. It involves two optical pulses, the emitted and the received, which are used to locate an object. The basic idea behind the measuring process involves the measurement of the time difference between emitted and received pulses by the sensor and converting the travel time into distance. A typical ALS system consists of a laser ranging unit, an opto-mechanical scanner and a position and orientation unit (Wehr and Lohr, 1999).

The laser ranging unit is responsible for generating the laser pulse with some solid laser material (e.g., Nd: YAG) in the wavelength between 800 and 1600 nm. It also has an electro-optical receiving system for receiving the reflected laser pulse. As the laser pulse from the aircraft reaches the ground, it diverges to a circle called as *'laser footprint'*. It can be defined as the total area from which the measurement is undertaken and gives an idea about the basic accuracy. Scanning mechanism in a typical ALS varies from system to system. There are various types of scanning mechanisms (oscillating mirror type, rotating mirror type, fiber optic type) and each uses different opto-mechanical systems and produces different types of track on the ground (Petzold, 1999). A fiber optic type scanner acquired the data used in the study, where an array of fiber lines is mounted on the focal plane of the transmitting and the receiving optics. For detail descriptions refer to Wehr and Lohr (1999). To obtain accurate range measurement in a given coordinate system, the ALS system must be supported by a position and orientation system (POS), which consists of DGPS (differential GPS) and IMS (inertial measurement system). DGPS consist of two GPS, one fixed and the other roving in the near vicinity, so they receive the signal with almost identical errors. Subtraction of both the signals gives the accurate measurement of relative position of the two receivers. If one of the receivers is placed on a point of known coordinate, then the absolute coordinate of the second receiver can be determined within accuracy of 'cm' level. The IMS on the other hand is installed on the aircraft and records information about the attitude (pitch, roll, heading) of the aircraft, in order to acquire the precise location of the laser receiver in three-dimensional space. A complete system of ALS data acquisition system is shown in Fig. 1.

The data used in the present study are acquired by TopoSys sensor-I (TopoSys, 2002) and covers a part of the city of Karlsruhe, Baden-Württemberg, Germany. The specification of the sensor used to acquire the data is given in Table 1. Since the main aim of the study is to define an automatic method for building extraction using only the ALS data, which implies distinguishing the building and vegetation using only height data. So focus was on the examination of those parts of the city where trees and buildings co-occur in a significant number. The data contain first and last returns acquired at different time. The first return data were acquired in the summer 1997, while the last return in the beginning of 1998 (Steinle et al., 2001).

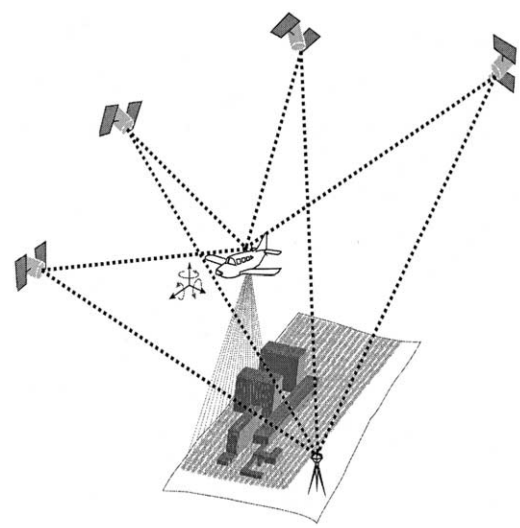

Fig. 1. Different components of an ALS system.

Table 1
Sensor characteristic of Toposys sensors used for data acquisition

Sensor type	Pulsed fiber scanner
Range	<1000 m
Wavelength	1.54 μm
Pulse length	5 ns
Scan frequency	650 Hz
Pulse repetition rate	83,000 Hz
Resolution of distance	–
Measurement	0.06 m
FOV	14°(±7°)
Swarth (at maximum range)	220 m
Av. measurement density (at maximum range)	5 measurements/m^2
Measurement possibilities	First or last pulse

3. Building extraction method

In the urban environment, the complex man-made structures may not have similarities amongst them in respect of shape, size, height and other criteria. In the process of automation, the primary task is to distinguish individual buildings and other structures using the laser data. The other challenge being to discriminate the trees in the periphery of man-made structures. The problem becomes more complex when the heights of the building and nearby tree are almost same as the process involves height values made available from laser data. The steps followed are:

- Preparation of normalised digital elevation model (nDEM).
- Segmentation.
- Extraction of the object border polygons.
- Application of modified standard deviation (MSD) algorithm.

3.1. Preparation of normalised digital elevation model (nDEM)

The term digital elevation model generally refers to a digital representation of topographic surface where height values of the terrain are given. According to USGS definition "A digital elevation model (DEM) is a digital cartographic/geographic data set of elevations in *xyz* co-ordinates. The terrain elevations for ground positions are sampled at regularly spaced horizontal intervals" (USGS, 2002). Hence, ALS data can be used directly as a DEM because it has *xyz* coordinates for each pixel. As it contains the heights of the upper surface of the terrain and the objects there on, so it is referred as digital surface model (DSM). In the DSM, the absolute height of the object is given with reference to MSL (mean sea level). To deduce the actual height of the object, height of the terrain should be subtracted from the DSM. It is necessary to create a digital terrain model (DTM), where the terrain is represented without the overlying objects. The subtraction of DTM from DEM results in the absolute height values of the objects and the model representing such heights is called as normalised digital elevation model (nDEM). The requirement at this stage is to produce a DTM from the original (ALS) data, a convex/concave hull algorithm proposed by von Hansen and Vögtle (1999) was used for this purpose. According to this the end and beginning of a convex or concave hull are searched along a line, i.e., from where there is a constant increase/decrease in the height values. An imaginary line then connects these end and beginning points and is utilised for calculating the height of each component of the convex or concave hull. Points having height less than that calculated by considering the concavity and distance between the end points were selected. The minimum of these selected points replaces the previous end points of the line, which were then classified as terrain points. The process stops when there is no terrain point left in the data set. For more detail refer to von Hansen and Vögtle (1999). The DTM generated was then subtracted from the DSM obtained from original ALS data to get the nDEM.

3.2. Segmentation

As it was decided to study the height variations along the boundary pixels of object, it is necessary to find the contour lines of the objects. For this, the objects present in the nDEM have to be segmented into various areas. The textural variation in height values of the laser data was used to separate them into various segments. The segmentation is also found suitable for eliminating trees having low height. According to the segmentation algorithm (Steinle et al., 2001), all the pixels of a particular window size should be checked if they are greater than a particular threshold or not, if so, these should be considered as the seed pixels. The threshold was determined by observing the height value of shrubs and small trees. In the next step, the height differences of the adjacent pixels should be checked against a second threshold. If this height difference is below this threshold then they should be considered as part of the preliminary seed. This iteration continues till none of the adjoining pixels have the height difference below the second threshold or all the pixels are checked. The result of the segmented data thus obtained also segregates some low height objects (mostly trees) from it.

3.3. Extraction of the object border polygons

For analysing the height values along the boundary of the segmented object, it is required to find the borderline of each individual segment. This can be achieved by triangulation, which involves creation from sample points a set of non-overlapping triangularly bounded facets in which the vertices of the triangles are the input sample points. A process based on Delaunay triangulation, which applies the principle of Voronoi diagram (Okabe et al., 1992) was used for this purpose. As a result, polygons corresponding to each segment were extracted from the segmented data.

3.4. Application of modified standard deviation (MSD) algorithm

As it is more likely that the height values of a building are uniform along its periphery or, if there is a change, then quiet uniform, in contrast to vegetation where this may normally not be true. Fig. 2(a), which represents a profile along the boundary of a building segment, shows uniformity in the data values. On the other hand Fig. 2(b), which represents a profile along the boundary

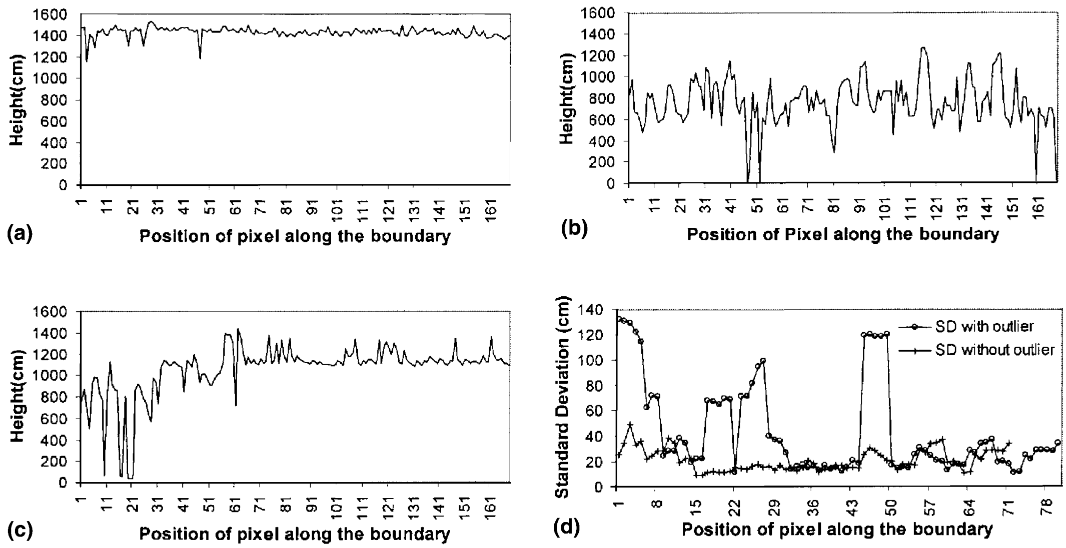

Fig. 2. Profile along the boundary of: (a) building, (b) tree, (c) mixed segment, (d) variation SD with and without outlier.

of a tree, shows more variability in the height values. In Fig. 2(c), which represents a profile along the boundary of a mixed segment (where both tree and building are classified under one segment), data are more variable wherever it is contributed due to the tree part of the mixed segment.

Standard deviation is an indicator, computed in a special way, which gives the variation of the normally distributed data and is used to measure the spread about the mean. The algorithm used in the study determines the standard deviation in a local domain, i.e., considering the height of a small number of boundary pixels (pixels falling in the polygon corresponding to a segment) in one go instead of considering the whole data and is called as the modified standard deviation method (MSD) because of the following two reasons: First, all consecutive pixels are not considered for calculating the value, as it was found that the presence of outlier (unusual variation in the data set) will affect the overall result. Fig. 2(d) shows how the presence of outlier in the data can affect the result. In buildings these outliers are generally present due to special type of construction like chimney, terrace, etc. Second, determination of standard deviation with respect to a best fit line rather than a mean line, as in general, the standard deviation is determined with respect to the mean value, which in case of height data it is a horizontal line if represented in 2D. However, there are some buildings where the height along boundary constantly changes because of its typical shape as shown in Fig. 3(a). In this case if the standard deviation with respect to mean is considered then it will be found against the horizontal line passing through the mean height as shown in Fig. 3(b) and will give a higher value of SD. This effect can be minimised by finding the standard deviation with respect to the best fit line passing through all the height values as indicated in Fig. 3(b). For this, the formula used is given by Eq. (1)

$$\mathrm{MSD} = \sqrt{\frac{\sum_{i=1}^{n}(h_i)^2}{n-1}}, \tag{1}$$

where h_i is the distance of the point with respect to the best fit line and n the number of data values used in the computation.

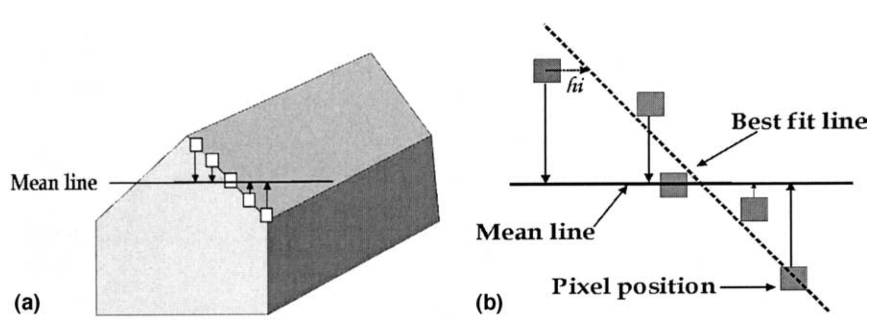

Fig. 3. (a) Building where the border is constantly changing will give higher value of SD with respect to a mean line. (b) The principle of MSD method.

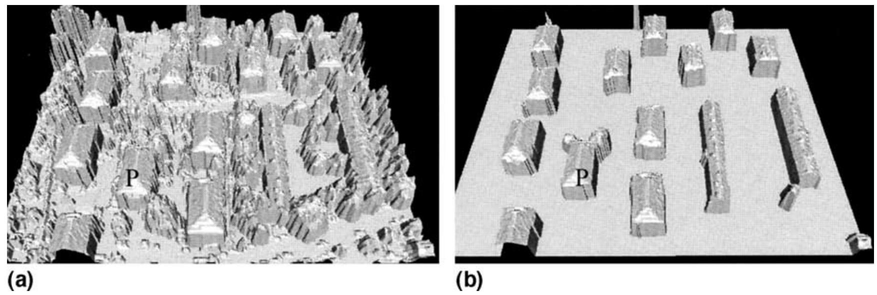

Fig. 4. (a) 3D visualisation of the nDEM data using Cosmo software (CAI, 2003). (b) 3D visualisation of the resulted data directly from MSD method.

4. Results and discussion

The above-described method was applied to the segmented ALS data to extract the building and it was observed that majority of the tree part from the image were removed as shown in Fig. 4. Results thus obtained can be summarised as follows:

- The buildings which were clearly represented in the original dataset, i.e., no part of the building is missing, were classified properly irrespective of their shape and size.
- Majority of tree segments were fully removed from data, despite their having shape similar to building.
- In the case of mixed segment, MSD method was able to remove the tree part leaving behind the buildings only. In some cases, although it was not able to remove the total tree, but was successful partly. The building marked 'P' in Fig. 4 is an example of mixed segment where a major part of the tree was removed from a mixed segment.

5. Conclusion

MSD method works efficiently with high resolution ALS data to detect the buildings automatically for a particular scene. This method can be applied for both pre and post disaster datasets as an input for preparing a near real time damage model. In developing countries like India, which was severely affected by many natural disasters in last decades, this technology is in its infancy, so it is required to use the available technology to its full brim. Although the initial cost of the data acquisition system is high, but its usability can compensate the cost. Finally, it can be inferred that the airborne laser scanning can go a long way as disaster mitigation tool.

Acknowledgements

The authors would like to thank Deutsche Forschungsgemeinschaft (http://www.dfg.de/english/index.html) for funding parts of the work through the collaborative research center 461 (Strong earthquakes: A challenge for Geosciences and Civil Engineering) in the project part CS (Image Analysis in Geosciences and Civil Engineering). One of the authors (JD) is grateful to DAAD (German Academic Exchange Service) for granting a scholarship under IIT-DAAD M.Tech sandwich program to work in Germany. The authors are also thankful to Mr. Ajay Mathur, Department of Geography, University of Southampton for his useful comments during the preparation of this paper.

References

Ackermann, F. Airborne laser scanning present status and future expectation. ISPRS J. Photogramm. Remote Sens. 54, 64–67, 1999.

CAI. http://www.cai.com/cosmo/, 2003.

Cobby, D.M., Mason, D.C., Davenport, I.J. Image processing of airborne scanning laser altimetry data for improved river flood mapping. ISPRS J. Photogramm. Remote Sens. 56, 121–138, 2001.

Drake, J.B., Dubayah, R.O., Clark, D.B., et al. Estimation of tropical forest structural characteristics using large-foot print lidar. Remote Sens. Environ. 79, 305–319, 2002.

Haala, N., Brenner, C. Extraction of building and trees in urban environments. ISPRS J. Photogramm. Remote Sens. 54, 130–137, 1999.

Irish, J.L., Lillycrop, W.J. Scanning laser mapping of the coastal zone: the SHOALS system. ISPRS J. Photogramm. Remote Sens. 54, 123–129, 1999.

Jayaraman, V., Chandrasekhar, M.G., Rao, U.R. Managing the natural disasters from space technology inputs. Acta Astronautica 40 (2–8), 235–291, 1997.

Maiman, T.H. Stimulated optical radiation in ruby. Nature 187 (4736), 493–494, 1960.

Mantovani, F., Soeters, R., Western, C.J.V. Remote sensing techniques for land slide studies and hazard zonation in Europe. Geomorphology 15, 213–225, 1996.

Maas, H., Vosselman, G. Two algorithms for extracting building models from raw laser altimetry data. ISPRS J. Photogramm. Remote Sens. 54, 153–163, 1999.

Massonnet, D. Application of remote sensing data in earthquake monitoring. Adv. Space Res. 15 (11), 37–44, 1995.

Okabe, A., Boots, B., Sugihara, K. Spatial Tessellation – Concept and Applications of Voronoi Diagrams. Wiley, England, 1992.

Petzold, B. DTM determination by laser scanning: an efficient alternative, in: Proceedings of the OEEPE Workshop on Automation in Digital Photogrammetric Production, 1999.

Steinle, E., Kiema, J., Leebemann, J., et al. Laserscanning for analysis of damages caused by earthquake hazards, in: Proceedings of the OEEPE – Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models, Stockholm, 1–3 March, 2001.

TopoSys. http://www.toposys.com, 2002.

USGS. http://www.usgs.gov/, 2002.

von Hansen, W., Vögtle, T. Extraktion der Geländeoberfläche aus flugzeuggetragenen Laserscanner – Aufnahmen. Photogrammetrie – Fernerkundung – Geoinformation 4, 229–236, 1999 (in German).

Wehr, A., Lohr, U. Airborne laser scanning – an introduction and overview. ISPRS J. Photogramm. Remote Sens. 54, 68–82, 1999.