論文

Extraction of Buildings from Airborne Laser Scanner Data and Aerial Photograph : A Wavelet-based Approach

航空機レーザースキャナーデータと航空写真を用いた建築物の抽出：ウェーブレットによるアプローチ法

T. Thuy VU, Mitsuharu TOKUNAGA** and Fumio YAMAZAKI****

Abstract : The extraction of objects from laser cloud points has become an interesting research topic with the emergence and implementation of several algorithms operating on high-density laser points. Without spectral information provided by the airborne laser scanner, these extraction methods suffer tremendous difficulties when low-density laser cloud points are adopted. To mitigate the difficulties in extraction, the size of object should be taken into account. This study proposed a multi-resolution clustering approach based on wavelet to extract the buildings in a dense urban area from the low-density airborne laser scanner data with the assistance of extractable information from the aerial photographs. The proposed approach was tested in Shinjuku-ku, Tokyo, Japan, and showed its efficiency.

和文概要：欧米では１平方メートルあたり５ポイント程度の高密度で観測する航空機レーザーデータを用いた建築物の抽出手法が提案されているが，日本で主に用いられている航空機レーザーデータは，１平方メートルあたり0.2ポイント程度の低密度データである。このデータからビルを抽出すると，観測点が疎であるためビルのエッジ抽出が十分に行えない。そのため，航空機画像を補助として，低密度の航空機レーザーデータから建築物を抽出する手法を提案した。

本手法はウェーブレットを用い点集合のレーザーデータを異なる解像度でクラスタリングし，航空機画像から得られたエッジ情報と組み合わせて建築物を抽出する。東京都新宿区の高層ビル地域に本手法を適用した結果，低密度の航空機レーザーデータからでも良好にビルを抽出することができた。

* Space Technology Applications and Research Program, Asian Institute of Technology, 12120, Thailand
** Environmental System Engineering, Kanazawa Institute of Technology Ishikawa 921-8501, Japan
***Institute of Industrial Science, The University of Tokyo Tokyo, Japan

1. INTRODUCTION

Airborne laser scanner data is acquired from an integrated system consisting of GPS, INS and laser scanner. It sends the laser pulse, receives the laser hit on the earth's surface, and has been considered as a highly accurate tool for topographic mapping (Wehr and Lohr, 1999). A variety of developed algorithms to extract the buildings from airborne laser scanner data mainly developed and implemented with the high density data such as Mass and Vosselman (1999) with approximately 5 points per square meter or Vosselman (1999) with approximately 7 points per square meter. However, such high-density laser points are not always available. Even though, high-density laser scanner data is available, the break lines are not explicitly presented in the data (Ackermann, 1999) (Axelsoon, 1999). Therefore, extraction method requires not only laser scanner data but also the additional data, such as aerial photograph. In this study, we have developed an algorithm to extract the buildings from the low-density laser data, which distribute about 0.2

points per square meter, with the assistance of extracted information from the aerial photograph.

The distribution of the objects in a dense urban area is very complicated. It creates extreme difficulties in extraction of the objects from both aerial photograph and airborne laser scanner data. Typical examples are the trees locating beside a house or the complex block of house including different-size components. Especially, when low-density airborne laser scanner data are adopted, less detail could be presented. Therefore, the sizes of the objects, here are the buildings, should be considered in mitigation of the difficulty in extraction. Analysis of the objects in the image or the cloud points at different resolutions has been proved as an excellent approach to detect and extract the target objects (Lega, et. al., 1995) (Starck and Murtagh, 1994). The key point is that the objects appear only at a certain range of scale, or resolution. Wavelet-based multi-resolution analysis has recently been attractive tool with the solid mathematical background in the 1980's to several researchers and engineers of different fields, even though the idea of wavelet originated in early 20th century. The fundamental idea of wavelet is to analyze the signal according to scale or resolution. In this study, wavelet-based multiresoltution analysis is adopted to detect the clusters distributed by laser points across the multiresolution space and consequently, extracts the buildings.

2. METHODOLOGY

Figure 1 illustrates the complete processing for the extraction of the buildings. The step-by-step detailed processing is described in the following subsections.

2.1 Interpolation

To maintain the surface discontinuity typifying the urban area, the planar interpolation method on the triangulated irregular network (TIN) was preferred. The higher degree of interpolation will smooth, and therefore, loose the sharp edges of the buildings. It is noted that there might be the shadow, i.e. no reflected laser points, due to skyscrapers locating near the edge of the flight strip. This shadow causes the displacement of the detected edges of the buildings from the most potential position, i.e. no explicitly presented break lines in airborne laser scanner data. Based on the assumption that the surface is flat in the shadow, the adjustment could be carried out easily by adding the pseudo points as illustrated in Figure 2. As a result, the TIN was reconstructed and ready for the interpolation to

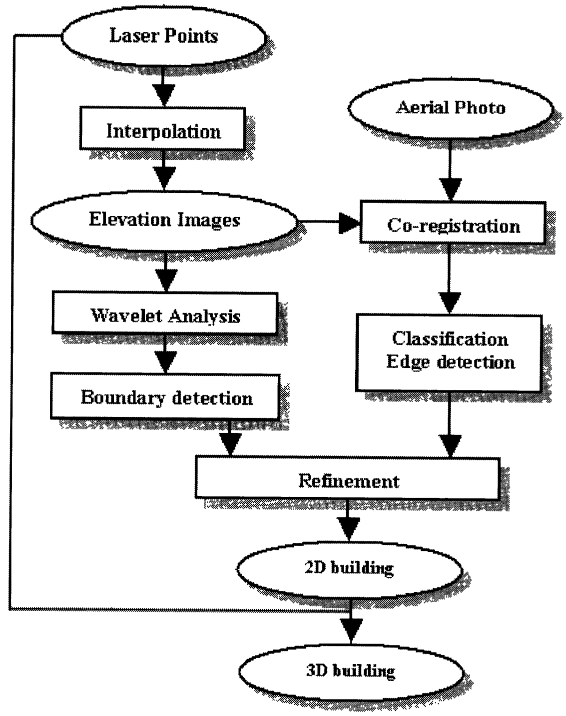


Figure 1 The processing flowchart for the extraction of the buildings.

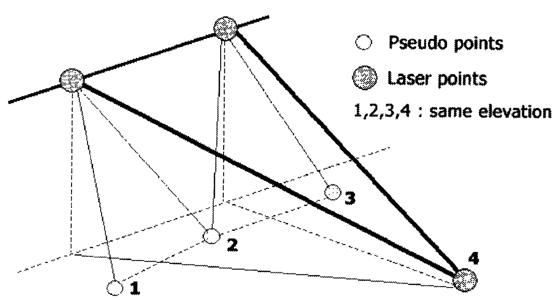


Figure 2 Improvement of planar interpolation on TIN.

grid format.

2.2 Wavelet analysis

A trous algorithm (Shensa, 1992) is applied to build up a multi-resolution framework. Let $\Phi(x)$ and $\Psi(x)$ are scaling and wavelet functions, respectively. The scaling function is chosen to satisfy the dilation equation as follows:

$$\frac{1}{2}\Phi\left(\frac{x}{2}\right)=\sum_{u}h(u)\Phi(x-u) \tag{1}$$

where h is a discrete low-pass filter associated with scaling function Φ and u is its size.

This equation shows the link between two consecutive resolutions, which are different by a factor of 2, by the low-pass filtering. The smoothed data $c_j(k)$ at a given resolution j and at position k can be obtained by the method of convolution:

$$c_j(k)=\sum_{u}h(u)c_{j-1}(k+2^{j-1}u) \tag{2}$$

The difference between two consecutive resolutions is calculated as

$$w_j(k)=c_{j-1}(k)-c_j(k) \tag{3}$$

The wavelet function $\Psi(x)$ is defined by

$$\frac{1}{2}\Psi\left(\frac{x}{2}\right)=\Phi(x)-\frac{1}{2}\Phi\left(\frac{x}{2}\right) \tag{4}$$

The cubic B-spline with the properties of compact support, symmetry, differentiability and one zero-crossing was chosen to be a scaling function. The implementation for 1D data is the convolution with the mask $\left[\frac{1}{16}\ \frac{4}{16}\ \frac{6}{16}\ \frac{4}{16}\ \frac{1}{16}\right]$. The *a trous* algorithm is easily extensible to the two-dimensional space. This leads to a convolution with a mask of 5×5 pixels for the wavelet connected to the cubic B-spline scale function. The coefficients of the mask with all the elements scaled up to 256 are:

$$\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

It is noted that wavelet is one kind of the linear multi-resolution (or multi-scale) analysis, which suffers from the distortion of the objects at a coarser resolution. To mitigate this problem, Median filter is applied prior to wavelet filtering. The wavelet analysis, therefore, is outlined in terms of pseudo codes as follows:

- Input a parameter: the number of the resolutions to be analyzed, e.g. $k$max.
- Init $k=1$, e.g. scale equals 1
- Assign the original image to *im_in*
- For each k, k is increased by 1 until $k=k$max
  - ○ Median filter *im_in* with the kernel size equals to 2^k+1, obtain *im_med*
  - ○ Detect the strong signatures by differencing *im_in* and *im_med* and thresholding with a threshold of 3σ, where σ is the standard deviation of the difference.
  - ○ Assign *im_in* to *im_tmp*
  - ○ Replace the values of the strong signatures in *im_tmp* by the ones in *im_med*
  - ○ Wavelet filter *im_tmp* (*a trous* algorithm as presented above), obtain *im_wave*, which is the wavelet-smoothed image at scale k.
  - ○ The wavelet coefficients or detailed image is the difference between *im_wave* and *im_in*.
  - ○ Assign *im_wave* to *im_in* for the next loop.

The wavelet coefficient images are used to detect the boundaries of the multi-resolution clusters, which depict the existence of the objects based on their size.

2.3 Cluster boundary detection

Tracking the chain in the detailed images, the boundaries of the clusters were found across the multi-resolutions space. Subsequently, the detected chain-codes of the cluster boundaries were converted to vector format. Selection of the appropriate resolutions was made through interactive processing. The decision was dictated by the distribution of the objects in the study area. As it is mentioned, due to the effect of interpolation and the gap between the laser points, the edges of the objects could not appear vertically. It required the careful checking of the laser points along the edges of the objects in the further processing.

2.4 Processing the aerial photograph

The ortho-rectified aerial photograph was used as the additional data in detection of the buildings. Firstly, the aerial photograph was co-registered with the interpolated elevation image. Subsequently, the spectral and structural information were extracted from this aerial photograph by the traditional techniques such as unsupervised classification and local edge detector. The selection of the suitable method for processing is discussed later in a specific test site. As it is known, each test site has its unique distribution of the objects. The proposed framework of this study preferred to open the capability for the end users in selecting the appropriate methods, which are available in the commercial software, for their test sites.

2.5 Refinement of the detected edges

The refinement of the detected edges from airborne laser scanner data through the extracted information from the aerial photograph was conducted in vector format. Firstly, the segments of the lines in from both scenes, i.e. airborne laser scanner data and aerial photograph, were intersected together to eliminate the redundant information. These were the blob elevations induce from the bushes or trees, or the concrete objects on the street. Secondly, the corresponding segments that were detected from both data sources were detected based on slope and distance. Subsequently, the same label marked the corresponding segments. Lastly, the visual refinement was carried out to finalize the detected boundaries of the buildings.

2.6 Detection of 3D building

Based on the detected edges from the previous processing, the elevation of a building was obtained by the elevation of the laser points falling within the boundary of the building. As a result, the laser points were classified into the bare earth points and the overlying object points. As it is stated earlier, the bare earth points might be misclassified as the overlying object points due to the interpolation and the gap between the laser points. It is obvious that the points belonging to the objects and located at the edges of the objects have a sharp leap in elevation when compared to the elevation of neighbors. The elevation threshold in the Delaunay neighbors could clarify this misclassification. Let OP is the object point set that has been detected and P as the remainders set of points. The wrongly classified point can be detected by the following equation.

$$WCP=\left\{\begin{array}{l} OP_k : (OP_k \backslash in\ OP) \text{ and} \\ (P_i \backslash in\ P) \text{and} \\ (OP_k \backslash in\ N_i) \text{ and} \\ |Z(OP_k)-AveZ(N_i)| \leq StdP \end{array}\right\} \quad (5)$$

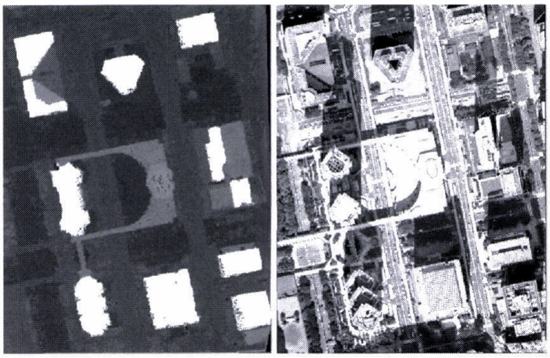
where WCP is the wrongly classified point set, $StdP$ is the given threshold, N_i is the Delaunay neighbor point set of terrain point P_i, $Z(P_i)$ is the elevation of the point P_i, $AveZ(N_i)$ is the average of elevation in N_i, $\backslash in$ denotes the "is-element-of" symbol.

After removal of wrong classified laser points, the elevation of a building was calculated by the average of all laser points falling within its detected boundary. Subsequently, the attribute of the buildings is imported to build up the 3D database of the buildings and 3D visualization.

3. TESTING

3.1 Study area

Tokyo locates at the latitude of 35°41′N and longitude of 139°41′E. The area of Tokyo is about 2168 squared kilometers with the population of 12 million. A typical urban area of 512m×512m in Shinjuku-ku, Tokyo, Japan was selected to test the competence of the proposed algorithm (see Figure 3). There are lots of buildings along with the crowded human activities in this area. The narrow streets appear in the tiny spaces between the very complex structures of the buildings. In addition, there exist numerous moving objects on the streets, trees aligned along the streets and buildings. The objects with different sizes, which are interspersed each other, typified the area. It is obvious to speculate the very difficult situation in filtering laser points and



a) b)

Figure 3 Data employed: a) Interpolated image; b) Aerial photograph.

detecting the buildings.

The parameters of the surveying flight with airborne laser scanner, which was conducted by Kokusai Kogyo Co., Ltd, over our testing area is given in Table 1. The approximate laser point density of acquired data is 0.2 point/m². It is quite low point density for such an application in urban areas. Table 2 describes the parameters of the aerial photograph that was provided and ortho-rectified by Nakanihon Air Service.

Table 1 The parameters of airborne laser scanner data acquired over Shinjuku, Tokyo.

Operation Altitude	2700m
Scan Swath Width	720m
FOV	16°
Scan Rate	19.5Hz
Pulse Rate	15KHz
Cross Track Spacing	1.93m
Along Track Spacing	2.83m
X, Y Positional Accuracy	0.3m RMSE absolute
Z Positional Accuracy	0.15m RMSE absolute

Table 2 The parameters of the aerial photograph.

Operation Altitude	3000m
Focal length	300mm
Image Size (meter)	500m×500m
Image Size (pixel)	2000pixels×2000pixels
Resolution	25cm×25cm
Scale	1/10000

3.2 Results

During pre-processing stage, the ortho-rectified aerial photograph was co-registered and resampled in the same frame with the laser scanner data. The interpolated image (Figure 3) shows the clusters formed by laser points at the chosen finest resolution. Applying wavelet-based multiresolution analysis, the cluster boundaries were detected from the airborne laser scanner data and converted in vector format (Figure 4). The multiresolution analysis could classify the objects based on their size. Figure 4 illustrates the existence of the cluster at two consecutive resolutions. Generally, the buildings can be considerably larger than other objects in an urban area. Across the multiresolution space, we could determine the existence of the large clusters, which probably belong to buildings.

Aerial photograph was employed to assist the construction of 3D buildings through its spectral information. Five classes such as conretel, concrete2, concrete3, shadow and vegetation were defined for the running of ISODATA unsupervised classification. Mathematical morphology, such as erosion and dilation, were applied to eliminate the isolated pixels and merge the small clusters to the near by large clusters. Without ground truth data, the unsupervised classification produced an acceptable classified result for visual assessment. Subsequently, the mask was applied to select the concrete classes, which had high probability to be part of

a) b)

Figure 4 Detected cluster boundaries: a) Finer resolution; b) Coarser resolution.

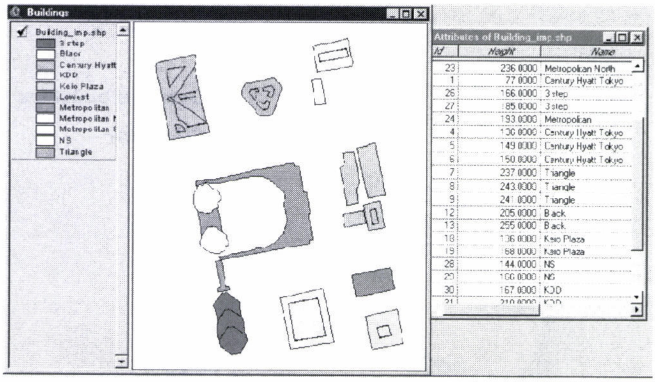
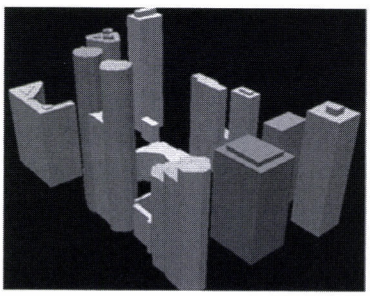
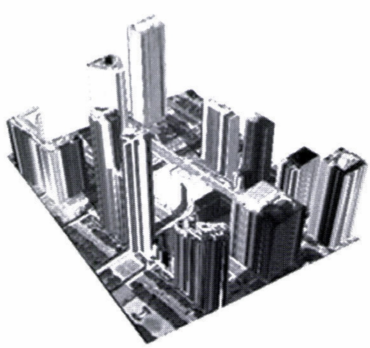


Figure 5 The final detected buildings.



a)



b)

Figure 6 3D visualization : a) Detected buildings ; b) Texture mapping.

buildings. The result showed many fragments of detected concrete classes due to the shadows of the skyscrapers. Anyhow, this information was utilized as much as possible to refine the information detected from laser point's data. Following the scheme described in previous section, the detected buildings and its primary database are shown in Figure 5. Additional data about buildings like name of owner, types of services, number of rooms, etc, can be imported easily into the prepared primary database.

Another product of this study was the 3D visualization. While Figure 6a illustrates simply the perspective view of detected buildings, Figure 6b depicted the texture mapping of the detected buildings by the aerial photograph.

4. ACCURACY ASSESSMENT

The detected buildings were compared with the 2D vector data. Visually (Figure 7), the buildings were detected completely, except the missing of the details along the edges and on the top of the buildings. The problem of missing details get worse at the building at the top right of the scene, named 'black' building. In the aerial photograph, the missed buildings showed the spectral information that resemble to vegetation. On another hand, in airborne laser scanner data, they were eliminated in the coarse resolution due to their small size.

Quantitative comparison was made for nine individual buildings between the results of the detection and 2D vector data. Table 3 shows the results of comparison both in absolute value and relative value. An average of 14% differences was observed in area. The difference was induced mainly from the lower parts and along the edges of the buildings like in the case of 'Keio Plaza' building and 'Triangle' building.

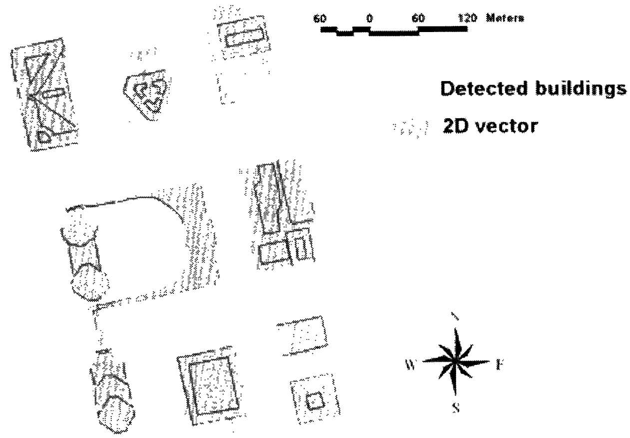


Figure 7 Comparison between detected buildings and 2D vector data.

Table 3 Quantitatively compared results.

ID	Building name	Difference	
		Absolute (m²)	Relative (%)
1	Century Hyatt Tokyo	772	8.71
2	Triangle	1127	28.97
3	Black	106	3.90
4	Metropolitan	33	0.28
5	Keio Plaza	4112	39.58
6	3 step	215	5.54
7	NS	418	6.44
8	Lowest	833	26.39
9	KDD	189	6.20
Average		867.22	14.00

5. CONCLUSION

This study proposed a new algorithm based on wavelet analysis to combined extract the buildings from airborne laser scanner data and aerial photograph acquired over a dense urban area. The multi-resolution analysis has proved a suitable approach to adopt the low-density airborne laser scanner for urban application. Most buildings in the study area were detected successfully and compared to 2D vector data. The difference in area between detected results and 2D vector data was 14%, due mainly to the missing lower parts of the skyscrapers. To obtain higher accuracy in detection, the algorithm was not a fully automatic processing. Some further studies should be carried out to improve the level of automation, which is capable to process larger area, with the consideration of the computation time and the obtained accuracy.

ACKNOWLEDGEMENT

Airborne laser scanner data for this study was provided by Kokusai Kogyo Co., Ltd, Japan.

(受付日2003.1.31，受理日2003.9.10)

REFERENCES

Ackermann, F., Airborne laser scanning-present status and future expectation. ISPRS Journal of Photogrammetry & Remote Sensing. Vol. 54 : pp. 64-67, 1999.

Axelsson, P., Processing of laser scanner data-algorithms and applications. ISPRS Journal of Photogrammetry & Remote Sensing. Vol. 54 : pp. 138-147, 1999.

Haala, N., Detection of buildings by fusion of range and image data. International Archives of Photogrammetry and Remote Sensing. Vol. 30, No. 3 : pp.341-346, 1994.

Haala, N., and Brenner, C., Extraction of buildings and trees in urban environment. ISPRS Journal of Photogrammetry & Remote Sensing. Vol. 54 : pp.

130-137, 1999.

Lega, E., Scholl, H., Alimi, J.M., Bijaoui, A., and Bury, P., A parallel algorithm for structure detection based on wavelet and segmentation analysis. Parallel Computing. Vol. 21 : pp.265-285, 1995.

Maas, H.G., and Vosselman, G., Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry & Remote Sensing. Vol. 54 : pp.153-163, 1999.

Mallat, S., A wavelet tour of signal processing. Academic Press, London, 1999.

Shensa M.J., Discrete Wavelet Transforms : Wedding the a trous and Mallat Algorithms. IEEE Transaction on Signal Processing, Vol. 40, 10 : pp.2464-2482, 1992.

Starck, J.L., and Murtagh, F., Image restoration with noise suppression using wavelet transform. Astronomy and Astrophysics. Vol. 288 : pp.342-348, 1994.

Vosselman, G., Building reconstruction using planar faces in very high density height data. ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, Munich, pp.8-10 September, 1999.

Wehr, A., and Lohr, U., Airborne Laser Scanning-an introduction and overview. ISPRS Journal of Photogrammetry & Remote Sensing. Vol. 54 : pp. 68-82, 1999.