

## ON OBJECT EXTRACTION USING AIRBORNE LASER SCANNER DATA AND DIGITAL IMAGES FOR 3D MODELLING

K. Nakano <sup>a,b,\*</sup>, H. Chikatsu <sup>c</sup>

<sup>a</sup> System Development and Research Center, AERO ASAHI CORPORATION,  
3-1-1, Minamidai, Kawagoe, Saitama, 350-1165, Japan - kazuya-nakano@aeroasahi.co.jp

<sup>b</sup> School of Science and Engineering, Tokyo Denki University,  
Ishizaka, Hatoyama, Saitama, 350-0394, Japan - 11hz035@ms.dendai.ac.jp

<sup>c</sup> Division of Architectural, Civil and Environmental Engineering, Tokyo Denki University,  
Ishizaka, Hatoyama, Saitama, 350-0394, Japan - chikatsu@g.dendai.ac.jp

### Commission III, WG III/1

**KEY WORDS:** LIDAR, Integration, Aerial Photogrammetry, Bundle Adjustment, Object Extraction, 3D Modelling

### ABSTRACT:

Airborne laser scanners are effective at extracting the micro topography or ground surface under trees, which cannot be detected by aerial photogrammetry, and are suitable for use in many applications, such as city modelling, DTM generation, monitoring electrical power lines, and detection of forest areas. The most remarkable aspect of these systems is their ability to acquire the 3D coordinates of huge object points in real-time. There are many studies on object extraction using point clouds from airborne laser scanner data, where the shape of an object depends on the density of a point. However, this is generally used for rough shapes or fitted geometric shapes. It is difficult to reconstruct detailed object shapes without many edge points, even if high-density point clouds are obtained. On the other hand, it is possible to acquire detailed object edges from digital camera images if the digital camera is equipped with an airborne laser scanner system. The procedures investigated in this paper for improving rough object shapes using airborne laser scanner data are as follows. Firstly, camera calibration is performed to integrate point clouds and digital images by simultaneous adjustment, such as by bundle adjustment with self-calibration using distance data taken directly from airborne laser scanner data. Secondly, the rough 3D object shape is extracted from the point cloud using normal vectors. Moreover, visualization of normal vectors is used for operator interpretation. Thirdly, the rough 3D object shape is converted into the image coordinates of multiple images by a collinearity condition. The 2D coordinates of detailed image shapes are acquired using characteristic image quantities from around the rough shape. Finally, the detailed 3D shape is computed using the spatial intersection of the 2D coordinates of detailed shapes and the orientation parameters. This paper describes fundamental studies for extracting object shapes for 3D modelling using airborne laser scanner data and digital images.

### 1. INTRODUCTION

Approximately 1200 GPS based control stations have been established by the Geospatial Information Authority of Japan (GSI). This infrastructure was utilized by various fields for the wide area crustal deformation caused by “The 2011 off the Pacific coast of Tohoku Earthquake” which occurred on March 11, 2011. Airborne laser scanner (ALS) systems are also used after disasters for topographic surveys using GPS based control stations. The ALS system has the advantage of acquiring detailed terrain data; however, objects such as buildings are generally represented as rough shapes by the discretely obtained point clouds. There are many studies using images to improve rough shapes. For example, Hu et al. (2004) demonstrated the extraction of buildings from LIDAR data and used edges extracted from high-resolution aerial images to refine laser data model accuracy. This approach is limited to buildings with primitive models. Chen et al. (2005) performed building reconstruction using aerial orthoimages and airborne laser scanner data; however, there were issues with the use and creation of orthoimages.

On the other hand, it is possible to acquire digital images if a digital camera is equipped with an ALS system. The digital camera is almost non-metric and needs camera calibration for accurate three-dimensional measurements. When a non-metric digital camera is used, its interior orientation parameters are generally computed beforehand using a test sheet or test target. However, if the digital camera is operated in severe conditions e.g. high-altitudes and low temperatures, camera calibration should be performed sequentially. The authors have been concentrating on developing a practical 3D measurement system for close range photogrammetry using consumer-grade digital cameras. The Image Based Integrated Measurement (IBIM) system is our photogrammetric system, which uses digital cameras and a hand-held laser distance meter (Nakano and Chikatsu, 2010). The orientation parameters of the triplet images are unknown and the pseudo-GCPs are simultaneously calculated by the collinearity condition, distance condition, and geometric constraint condition. It is possible to integrate point clouds and digital images by applying the concept of the IBIM system to ALS system camera calibration. With this motive, simultaneous adjustments, such as bundle adjustments with self-calibration, are proposed in this paper so that exterior

---

\* Corresponding author.

orientation parameters obtained from the GNSS/IMU system, distances and 3D object coordinates acquired from the laser scanner, and the interior orientation parameters are simultaneously adjusted. Combined block adjustment orientations were proposed in the late 1900's (Ackermann et al., 1972, EL-Hakim & Faig, 1981, Chikatsu et al., 1988). The proposed adjustment is widely expected to enable the utilization of the airborne laser scanner in generating large-scale maps and efficient aerial photogrammetry should be accomplished, except for geodetic data such as ground control points and aerial triangulation. Therefore, this paper uses calibration of non-metric digital cameras to integrate point clouds and digital images.

The object extraction procedures using ALS data and digital images are performed in three steps.

1) A rough 3D object shape is extracted using a normal vector map that is created from TIN by point clouds. Visualization of normal vectors is useful for operator interpretation.

2) The rough object shapes are converted into multiple image coordinates by a collinearity condition. The 2D shape coordinates of detailed images are acquired using image characteristics from around the rough shape.

3) The detailed 3D shape is computed using the spatial intersection of detailed 2D shape coordinates and orientation parameters.

A flowchart of the object extraction procedure is shown in Figure 1.

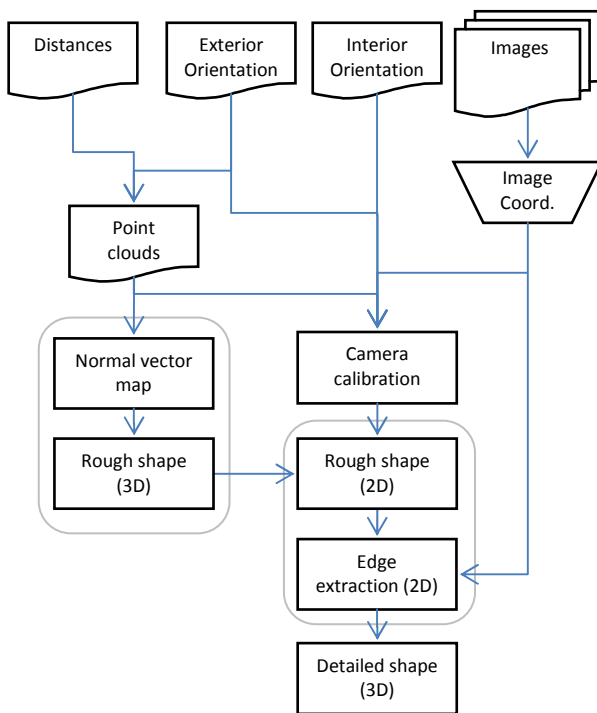


Figure 1. Object extraction flowchart

## 2. CAMERA CALIBRATION

The authors have been concentrating on developing a close range measurement system for consumer grade digital cameras using triplet images (Chikatsu et al., 2006). The measurement system was adopted into digital aerial photogrammetry in this paper because triplet images have following characteristics.

- Triplet images have advantages in generating stereo pairs.
- Triplet images have the flexibility for multiple images.
- Triplet images have the ability to increase geometric restriction.

Moreover, the IBIM system of the basic camera calibration concept has distance condition characteristics and also uses pseudo ground control points (GCPs), which are virtual points. Figure 2 shows the measurement concept used in this paper.

On the other hand, lens distortion is the most important interior orientation parameter, and many distortion models have been proposed (Brown, 1971, Murai, Matsuoka, Okuda, 1984). This paper uses Brown's 1971 model, which takes the 7th degree of the radial polynomial equation and the tangential distortion into account, and has been widely used in close range photogrammetric fields.

$$\begin{aligned} x &= x' + \frac{x'}{r} (K_1 r^3 + K_2 r^5 + K_3 r^7) + P_1 (r^2 + 2x'^2) + 2P_2 x'y' \\ y &= y' + \frac{y'}{r} (K_1 r^3 + K_2 r^5 + K_3 r^7) + 2P_1 x'y' + P_2 (r^2 + 2y'^2) \end{aligned} \quad (1)$$

where  $r = \sqrt{x'^2 + y'^2}$  = the radial distance from the principal points

$x, y$  = corrected image coordinates

$x', y'$  = image coordinates

$K_1, K_2, K_3$  = radial distortion coefficients

$P_1, P_2$  = tangential distortion coefficients

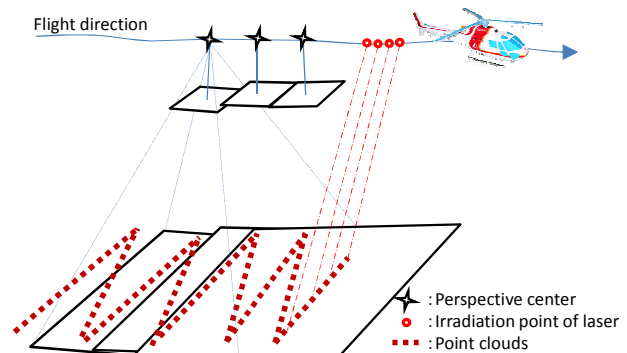


Figure 2. Measurement concept

The exterior parameters ( $X_0, Y_0, Z_0, \omega, \phi, \kappa$ ) and the interior parameters ( $f$  [focal length],  $u_0, v_0$  [principal points],  $a, b$  [scale factor, shear factor],  $K_1, K_2, K_3, P_1, P_2$  [lens distortion]) are unknown parameters of the multiple images and the pseudo-GCPs ( $X_i, Y_i, Z_i$ ), respectively. These unknown parameters are simultaneously calculated by the collinearity condition, distance condition, and geometric constraint condition under the local coordinate system. Here, the collinearity condition is shown as Equation (2) and the distance condition is shown as Equation (3).

$$\begin{aligned} x &= -f \frac{m_{11}(X - X_0) + m_{12}(Y - Y_0) + m_{13}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)} \\ y &= -f \frac{m_{21}(X - X_0) + m_{22}(Y - Y_0) + m_{23}(Z - Z_0)}{m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0)} \end{aligned} \quad (2)$$

where  $x, y$  = corrected image coordinates

$f$  = focal length

$X, Y, Z$  = pseudo-GCP object coordinates

$X_0, Y_0, Z_0$  = perspective center

$m_{ij}$  = rotation matrix elements

$$D = \sqrt{(X - X_0 - \Delta X)^2 + (Y - Y_0 - \Delta Y)^2 + (Z - Z_0 - \Delta Z)^2} \quad (3)$$

where  $D$  = distance from exposure station to pseudo-GCP

$X, Y, Z$  = pseudo-GCP object coordinates

$X_0, Y_0, Z_0$  = perspective center

$\Delta X, \Delta Y, \Delta Z$  = differences between the laser scanner irradiation point and perspective center

The trifocal tensor is a geometric relation of three images containing the same objects from different perspectives (Hartley, 1993). The trifocal tensor is expressed by three square matrixes ( $3 \times 3$ ), which are  $T_1, T_2$ , and  $T_3$ , the components of these matrixes are  $t_{1ij}, t_{2ij}$ , and  $t_{3ij}$ , and the image coordinates of matched points for these three images are  $(x_1, y_1, z_1), (x_2, y_2, z_2)$ , and  $(x_3, y_3, z_3)$ . Thus, the following equations are obtained by the geometric relation.

$$-z_2 z_3 g_{22} + z_2 y_3 g_{23} + y_2 z_3 g_{32} - y_2 y_3 g_{33} = 0 \quad (4)$$

$$z_2 z_3 g_{21} - z_2 x_3 g_{23} - y_2 z_3 g_{31} + y_2 x_3 g_{33} = 0$$

$$z_2 z_3 g_{12} - z_2 y_3 g_{13} - x_2 z_3 g_{32} + x_2 y_3 g_{33} = 0$$

$$-z_2 z_3 g_{11} + z_2 x_3 g_{13} + x_2 z_3 g_{31} - x_2 x_3 g_{33} = 0$$

where  $g_{ij} = x_1 t_{1ij} + y_1 t_{2ij} + z_1 t_{3ij}$

It is understood that image coordinate  $(x_2, y_2)$  is the spatial intersection of two epipolar lines on the second image (centre image) and is calculated by Equation (5), derived from Equation (4). Therefore, the geometric constraint condition uses Equation (5).

$$x_2 = z_2 \frac{z_3(g_{11} + g_{12}) - g_{13}(x_3 + y_3)}{z_3(g_{31} + g_{32}) - g_{33}(x_3 + y_3)} \quad (5)$$

$$y_2 = z_2 \frac{-z_3(g_{21} + g_{22}) + g_{23}(x_3 + y_3)}{-z_3(g_{31} + g_{32}) + g_{33}(x_3 + y_3)}$$

Camera calibration is performed by calculating these unknown parameters, which can be calculated as values by minimizing the following function,  $H$  (Equation (6)), under the least squares method.

$$H = \sum_{i=1}^m \sum_{j=1}^n \left\{ \begin{array}{l} p_1 (\Delta x_{ij}^2 + \Delta y_{ij}^2) + p_2 (\Delta D_j^2) \\ + p_3 (\Delta x_c^2 + \Delta y_c^2) \\ + p_4 (\Delta X_i^2 + \Delta Y_i^2 + \Delta Z_i^2) \end{array} \right\} \Rightarrow \min \quad (6)$$

where  $\Delta x_{ij}, \Delta y_{ij}$  = residuals for image coordinates

$\Delta D_j$  = residuals for distance

$\Delta x_c, \Delta y_c$  = residuals for the centre image of the geometric constraint condition

$\Delta X_i, \Delta Y_i, \Delta Z_i$  = residuals for pseudo-GCP object coordinates

$m$  = number of pseudo-GCPs

$n$  = number of images

$p_1, p_2, p_3, p_4$  = weight for each condition

### 3. OBJECT EXTRACTION PROCEDURES

Object extraction was performed in the ALS data phase and image phase. Detail procedures are as follows.

#### 3.1 Visualization of normal vectors

There have been many approaches to building extractions from ALS data since the late 1990s, such as using height data and normalized DSMs, subtracting DTM from DSM, region growing, slope maps, and normal vectors. In particular, normal vectors are used for terrain or road surface information, rooftops, trees, and so on. Normal vectors are calculated from a TIN

generated from random point clouds. A normal vector is managed by the grid for efficiency. Gridding is normalized using a whole value by combining the normal vectors in the grid range. A normal vector map was produced in the X, Y, Z direction assigned to the R, G, B channels. The normal vector map is shown in Figure 3. It seems that normal vector maps indicate the shapes of houses as well as the slope of the roof by color gradation. The X, Y, and Z directions indicate each characteristic, e.g., the X and Y directions show North and East lighting on the map (Figure 4) and the Z direction shows object shapes (Figure 5).

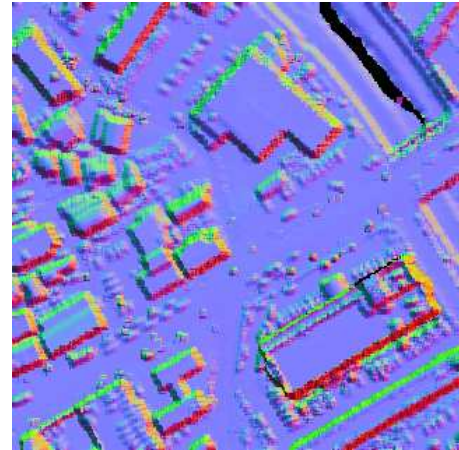


Figure 3. Normal vector map

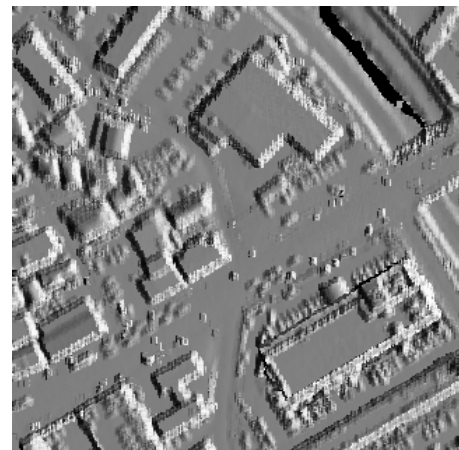


Figure 4. The X direction of the normal vector map

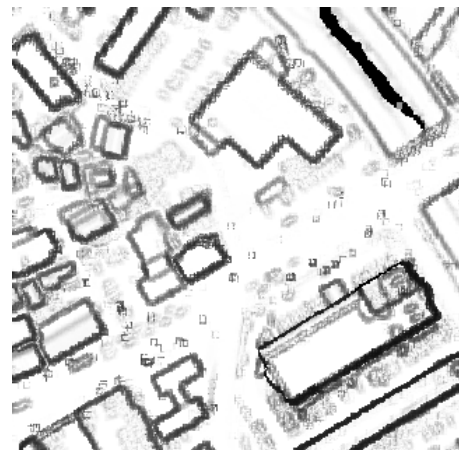


Figure 5. The Z direction of the normal vector map

### 3.2 Rough shape extraction

The 3D rough object shape is extracted from the normal vector map via image processing. The image processing procedures are as follows.

- Binarizing using Otsu's threshold method (Otsu, 1980)
- Noise reduction for small objects such as telegraph poles and cars
- Thinning and line tracing

Figure 6 shows the resulting 3D rough shape using image processing.

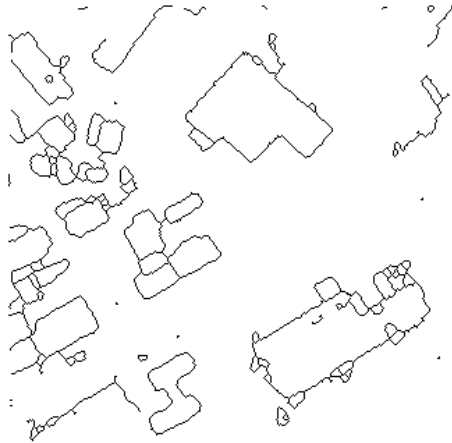


Figure 6. 3D rough shape

### 3.3 Image shape extraction

The rough object shapes are converted into multiple image coordinates by the collinearity conditions. Images around the object are clipped in order to limit edge extraction processing (Figure 7, Figure 8 (a, b)). The flight direction is from west to east and clipped images are rotated 90 degrees, as shown in Figure 7, 8. The Canny operator is used to extract object edges from clipped digital images. The Canny operator result is shown in Figure 9 (a). Please note that the Canny operator result is not binarized at this stage in order to use the threshold, depending on the situation. The edge potential map (Figure 9(b)) was created according to the distance from the point of the edge that is considered to exist when a point around the laser has been converted. Edge candidates (Figure 9(c)) that were computed using the edge potential map and the Canny operator result using Otsu's threshold method were thinned. The thinned edge candidates are reliable; however, more than one object is disconnected. All edges are extracted by minimum threshold value in order to connect the edges (Figure 9(d)). This will be the connected endpoint of two edge candidates that are included in the complete edge.

An object is extracted from triplet images that are taken as multiple images in this paper. Occlusion changes the shapes of objects in images with different perspective centers. Therefore, the intersection image was calculated using triplet images in order to reduce the edge mismatch caused by occlusion. The intersection image is created by affine transformed images that are transformed using converted rough image shape coordinates. The intersection image is shown in Figure 10. A blunder mask is computed for each image using the blue part, which indicates the occlusion location and changes in the intersection image. Finally, a 3D model is created using the extracted edges and

feature points, which are calculated using the collinearity conditions.

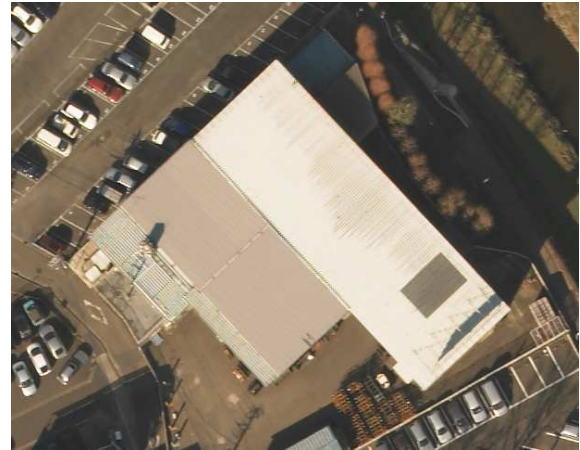
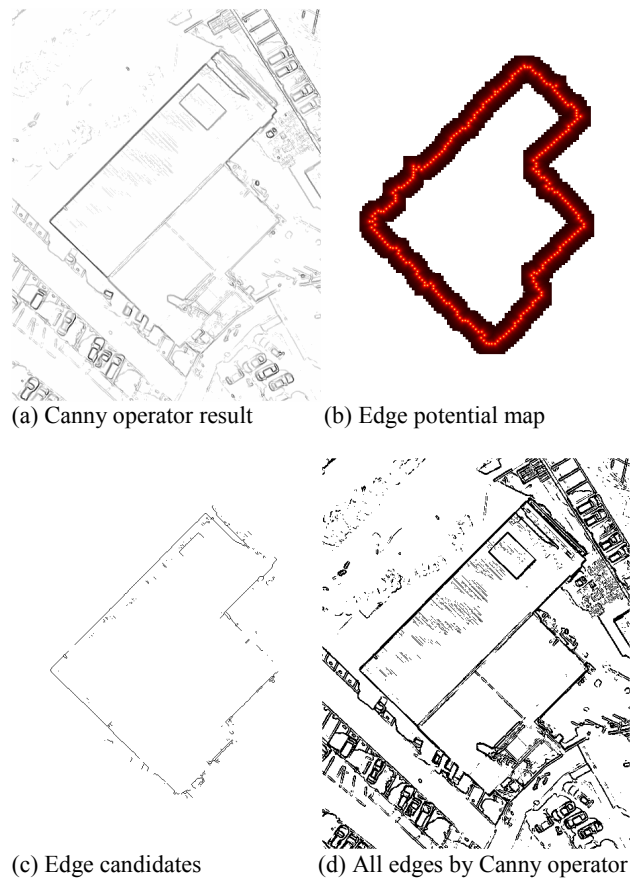


Figure 7 Clipped digital camera center images



(a) left image (b) right image  
Figure 8. Clipped digital camera images



(a) Canny operator result (b) Edge potential map  
(c) Edge candidates (d) All edges by Canny operator

Figure 9. Edge extraction process images



Figure 10. Intersection image

#### 4. EXPERIMENT

##### 4.1 Data acquisition and camera calibration evaluation

Data acquisition was performed in our test field in order to confirm the validity of the proposed simultaneous adjustment. The test field had 117 GCPs, and seven GCPs could be utilized in the triplet area. The GCPs were obtained by static observations using GPS that were set on the edges of road paints. Table 1 shows the data components used in this investigation and Figure 11 shows the centre image used in this investigation. It was taken at about 820 m and the image scale is about 1/13,700. Therefore, the GSD (Ground sampling distance) was about 12 cm.

Domain	Item	Contents
Test field	All GCPs	117 points
Platform	Helicopter	AS-350BA
Flight	Speed	120 km/h
	Altitude	820 m
GPS	Rate	1 sec
IMU	Rate	200 Hz
Laser	Scan rate	50 kHz
	Scan angle	30 degree
	Pulse	4 pulse
Digital Camera	Focal length	60 mm
	Sensor size	9 x 9 μm
	Image size	5440 x 4080
	Color depth	16bit

Table 1. Data components

Table 2 shows the root mean square errors for seven check points and permissible errors means restrictions for check points that are established by GSI when generating each scale map. RA (Relative Accuracy) means the value that was computed from equation (7). Standard error was computed under the

assumption that image-coordinate pointing is accomplished with one-pixel accuracy.



Figure 11. Whole centre image

Type	$\sigma_x$ [m]	$\sigma_y$ [m]	$\sigma_z$ [m]	RA
Proposed method	0.209	0.269	0.242	1/1963
Standard error	0.123	0.123	0.482	1/1601
Permissible error				
1/500	0.150	0.150	0.200	1/2813
1/1000	0.300	0.300	0.300	1/1578
1/2500	0.750	0.750	0.500	1/699

H = 820m in RA

Table 2. RMSE for check points

$$RA = \frac{1}{\sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{H}}} \quad (7)$$

$$\sigma_x = \sqrt{\frac{\sum \sigma_{xi}^2}{n_x}}, \quad \sigma_y = \sqrt{\frac{\sum \sigma_{yi}^2}{n_y}}, \quad \sigma_z = \sqrt{\frac{\sum \sigma_{zi}^2}{n_z}}$$

where  $H$  = Altitude

$\sigma_x, \sigma_y, \sigma_z$  = RSM error of X, Y, Z

$\sigma_{xi}, \sigma_{yi}, \sigma_{zi}$  = differences in X, Y, Z coordinates

$n_x, n_y, n_z$  = numbers of check points

$$\sigma_{x0} = \sigma_{y0} = \frac{H}{f} \sigma_p \quad (8)$$

$$\sigma_{z0} = \sqrt{2} \frac{H}{f} \frac{H}{B} \sigma_p$$

where  $\sigma_{x0}, \sigma_{y0}, \sigma_{z0}$  = standard error

$H$  = altitude

$f$  = focal length

$B$  = base line

$\sigma_p$  = pointing accuracy

It can be seen that the vertical coordinate value, Z, is better than the horizontal coordinate values because the vertical coordinates of the pseudo-GCPs are constrained by laser distances as a characteristic of the proposed method. On the other hand, it can be found that the relative calibration accuracy is less than 1/500 and more than 1/1000 in comparison to the permissible error. It is consequently concluded that camera calibration using pseudo GCPs is practical.

## 4.2 Object extraction

Object extraction procedures using experimental data are shown in Chapter 3. It can be said that object shape extraction using the conversion between object space and image space for camera calibration was successful. Figure 12 shows 3D modelling of the object extraction result. Blue line shape is obtained by manual plotting, red dots are rough shape by laser, and green line is object extraction result. Object extraction procedures can create 3D models; however, it produces strange shapes owing to mismatches.

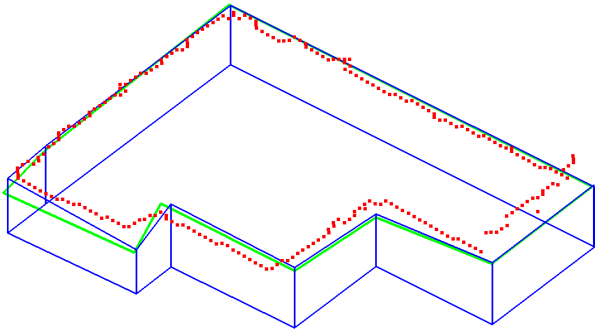


Figure 12. 3D modelling

## 5. CONCLUSIONS

A camera calibration technique and object extraction procedures were developed in this study in order to achieve 3D modeling using ALS data and digital camera images.

It is confirmed that camera calibration using pseudo GCPs and simultaneous adjustment shows GSI restrictions of less than 1/500 and more than 1/1000 in generating each scale map. Therefore, it is concluded that simultaneous adjustment using pseudo GCPs, distance conditions, and geometric constraint conditions is practical because the simultaneous adjustments perform interior and exterior orientations without any GCPs or aerial triangulation.

The object extraction procedure was established using ALS data and digital images. The normal vector map is a useful tool for operator interpretation and rough object shape extraction. Moreover, it was effective at extracting object shapes by image processing using ALS data.

However, there are some issues requiring further investigation. These problems include accuracy improvement and automatic generation of pseudo GCPs for camera calibration and object shape extraction.

## References

- Abdel-Aziz, Y. I., 1982. Accuracy of the normal case of close-range photogrammetry. *Photogrammetric Engineering Remote Sensing*, 48, 207-213.
- Ackermann, F., Ebner, H., Klein, H., 1972, Combined block adjustment of APR Data and Independent Photogrammetric Models, *The Canadian Surveyor*, Vol. 26, pp.384-396.
- Canny, J., 1986. A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp.679-698.

- Chen, L., Teo, T., Rau, J., Liu, J., Hsu, W., 2005. Building Reconstruction from LiDAR Data and Aerial Imagery. In *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Vol.4, pp.2846-2849.

- Chikatsu, H., Kasugaya, N., Murai, S., 1988, An Adjustment of Photogrammetry Combined with the Geodetic Data and GPS, *16th International Society for Photogrammetry and Remote Sensing*, Vol.26, pp.110-121.

- Chikatsu, H., Odake, T., 2006, Ubiquitous Digital Photogrammetry by Consumer Grade Digital Camera, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVI, PART 5 (CD-Rom), ISSN 1682-1750.

- EL-Hakim S.F. and Faig W., 1981, A combined Adjustment of Geodetic and Photogrammetric Observations, *Photogrammetric Engineering and Remote sensing*, vol.47, No.1, pp.93-99.

- Hartley, R.I. 1993. Camera calibration using line correspondences. In *Proc. DARPA Image Understanding Workshop*, pp. 361-366.

- Hu, J., You, S., Neumann, U., Park, K., 2004. Building Modeling from LiDAR and Aerial Imagery. In *Proceedings of ASPRS 2004 Annual Conference [CD-ROM]*.

- Murai, S., Matsuoka, R. and Okuda, T., 1984, A Study on Analytical Calibration for Non-Metric Cameras and Accuracy of Three Dimensional Measurement, *International Archives of Photogrammetry*, vol.25, issue 5, pp.570-579.

- Nakano, K. and Chikatsu, H., 2011. Camera-Variant Calibration and Sensor Modeling for Practical Photogrammetry in Archeological Sites, *Remote Sensing*, vol. 3, issue 3, pp. 554-569.

- Otsu, N., 1980. An Automatic Threshold Selection Method Based on Discriminant and Least Squares Criteria, *The transactions of the institute of Electronics and Communication Engineers of Japan*, IECE J63-D(4), pp.349-356.