3D URBAN MAPPING BASED ON THE IMAGE SEGMENTATION USING TLS DATA

Katsuyuki NAKAMURA, Masafumi NAKAGAWA and Ryosuke SHIBASAKI
Center for Spatial Information Science,
Institute of Industrial Science, University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
Tel & Fax : (81)-3-5452-6417
E-mail : katsu@iis.u-tokyo.ac.jp

KEY WORDS: TLS (Three Line Scanner), 3D urban model, Semi automated technique, Stereo matching

ABSTRACT:
Recently, it is demanded to generate effectively 3D data of urban features such as roads and buildings. Automation of the 3D data generation with high accuracy should be achieved to meet this requirement. In many cases, edges and regions are extracted from stereoscopic images to construct 3D building models and road models with stereo matching techniques. However, process of edge extraction and region segmentation is so erroneous that manual correction or editing works by human edition cannot be avoided. However, manual editing work directly handling 3D objects in 3D space through stereoscopic observation requires great skill and a lot of time. In this study, the authors propose a semi automatic technique of 3D urban mapping that can allow novice operators to edit on 2D space of aerial images. In this method, an operator is only required to input the conjugate points of a building. Moreover, using the TLS (Three Line Scanner) images, an operator can edit only the nadir image. As a result, it turned out that semi automatic 3D mapping with good accuracy could be possible. Thus, precise 3D model construction can be possible without conducting a stereoscopic observation.

1 INTRODUCTION

1.1 Background

Recently, 3D spatial data has been widely used in some fields such as car navigation system, for games, flood/landscape and other simulations and so on. Therefore, it is demanded to generate effectively 3D data of urban features such as roads and buildings. Automation of the 3D data generation with high accuracy should be achieved to meet this requirement. However, it is difficult to automate a generating procedure and there are no photogrammetric systems that can completely automate a process of spatial information acquisition.

1.2 Objectives

In many cases, edges and regions are extracted from stereoscopic images to construct 3D building models and road models with stereo matching techniques. However, the process of edge extraction and region segmentation is so erroneous that manual correction or editing works by human edition cannot be avoided. However, manual editing work directly handling 3D objects in 3D space through stereoscopic observation requires great skill and a lot of time.

In this study, the authors propose a semi automatic technique of 3D urban mapping that allows a novice operator to edit on 2D space of aerial images. In this method, an operator is only required to input the conjugate points of a building. Moreover, using the TLS (Three Line Scanner) images, operators can edit only the nadir image and other process are automated. It is tried to generate exact 3D urban model by semi automatic technique.

2 THEORY AND METHODOLOGY

2.1 TLS (Three Line Scanner)

TLS (Three Line Scanner) is an optical sensor for aerial survey. TLS is composed of three linear CCD arranged in parallel and can acquire three images of each direction (forward, nadir and backward) at the same time. Orienting it on an aircraft perpendicularly to flight direction, and scanning a ground place, triple stereo images of a ground object can be acquired (See figure 1). As a result, occlusion area can be reduced greatly. It is also possible to get 3D coordinates by stereo matching using two images of three. One of the advantages of a linear CCD sensor is that

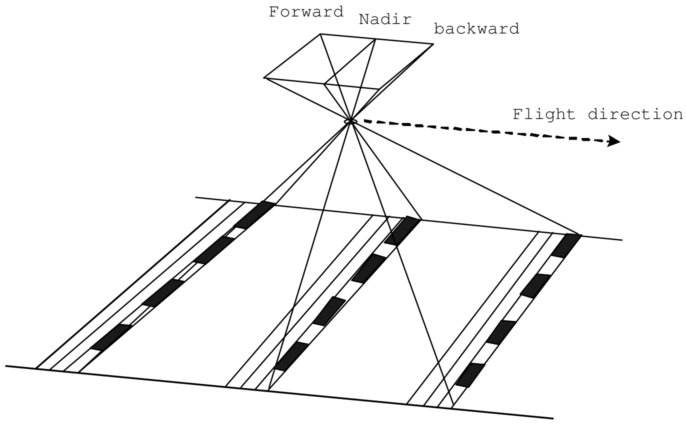


Figure 1: The concept of survey using TLS (Three Line Scanner)

more pixels can be arranged in a single sensor. This can achieve a resolution compatible with that of an aerial photo. Though a linear CCD sensor can acquire data only one line at a time (The ground resolution of TLS data in this research is 3cm approximately). Moreover, the image is greatly influenced by fluctuation of the position and altitude of an airplane. The fluctuation influences can be reduced by setting up a stabilizer between the airplane and TLS.

2.2 Searching for conjugate points in the object space

In this study, the searching for conjugate points of features is conducted as follows. First, a feature boundary polygon (e.g. building boundary polygon) is inputted manually and their image coordinates are converted to ground coordinates. Basic equations of this conversion are shown (1), (2) and (3).

$$X = (Z - Z_0)\frac{a_{11} + a_{21}y - a_{31}c}{a_{13} + a_{23}y - a_{33}c} + X_0 \tag{1}$$

$$Y = (Z - Z_0)\frac{a_{12} + a_{22}y - a_{32}c}{a_{13} + a_{23}y - a_{33}c} + Y_0 \tag{2}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix} \begin{pmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

Where, X, Y and Z are ground coordinates of the object, X_0, Y_0 and Z_0 are coordinates of center of projection, c is focal length, x,y are photographic coordinates, a_{ij} is element of rotation matrix, ω,φ and κ are rotation angles. However, at this moment, correct Z value is not known. By using the results of triangulations of forward and backward images, the feature boundary polygons represented with ground coordinates are back-projected onto the forward and the backward images. If the Z values are correct enough, the projected boundary will match to that of a conjugate feature in the forward and backward images. The searching for the conjugate boundary polygon is conducted while changing the Z value and projecting the boundary polygon of the nadir image to the forward and backward image respectively. In this case, 10m is given as initial value of ground height and the searching for conjugate boundary polygon is conducted up to 70m. In this sense, searching is conducted based on the object space. The concept of searching is shown figure 2 and the details are described below.

2.3 Stereo Matching Methods

It is important to narrow down the search range in stereo matching not only because of reducing the amount of calculation, but also to reduce miss matching. In this study, two methods are applied to perform the stereo matching.

In the first phase, rough matching is performed in order to narrow down the search range. In this phase, statistical characteristics (average and standard deviation) were used. In the second phase, matching by correlation value between nadir and fore/back images is applied to specify the conjugate points. In this phase, the range limited by the 1st matching was searched as detailed matching. Flow of this process is shown as follows.

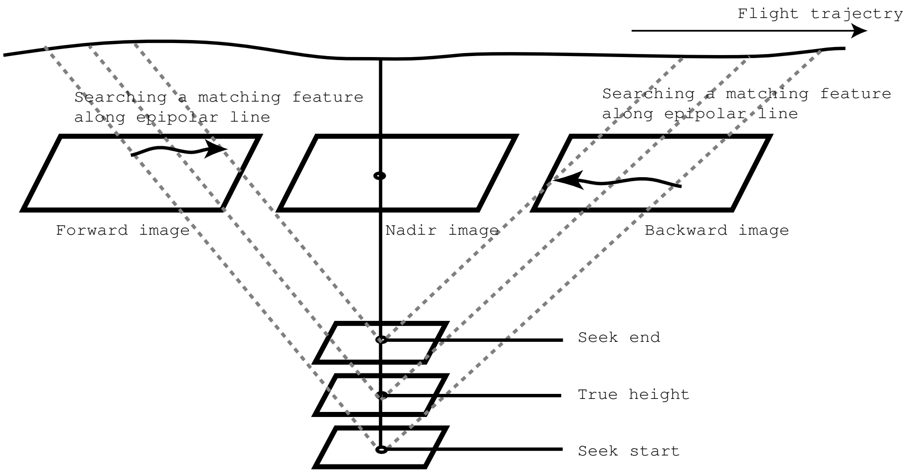


Figure 2: Searching for conjugate points in the Object space

(1) Start of measuring the building
(2) Inputting the feature points (Loop of optional n time)
(3) Projecting to the object space by collinearity condition equation (1),(2) and (3)
(4) Back projecting to fore/back images (Initial Z value is given as 10m to 70m)
(5) Generating the epipolar line (Where, unit width of matching is set to m, number of matching points is set to n)
(6) [1st matching] Stereo matching using statistical characteristics (Search of i time (i : 0, 10, n))
(7) Calculating the matching point (1st matching point is set to k)
(8) Interpolating the fore/back templates by the Nearest neighbor method
(9) [2nd matching] Stereo matching using correlation value (Search of j time (j : k-10, 1, k+10))
(10) Defining the matching point
(11) Determining the ground height Z.
(12) Return (1) or end of process

2.3.1 Approximate matching by statistical characteristics

Matching by statistical characteristics was selected as an approximate matching. The average and standard deviation of templates are applied to search for a matching point. Rough matching are conducted as first searching in order to reduce the amount of calculation using large unit width. In this case, the authors used following evaluation functions to define the matching point. Here, evaluation function $E(p)$ at searching point p is defined by next equation.

$$E(p) = \left|\mu\left(T^p_{\text{nadir}}\right) - \mu\left(T^p_{\text{fore/back}}\right)\right| + \left|\sigma\left(T^p_{\text{nadir}}\right) - \sigma\left(T^p_{\text{fore/back}}\right)\right| \tag{4}$$

Where, T^p_{nadir} is nadir template obtained on the bisis of searching point p, $T^p_{\text{fore/back}}$ is the fore/back template, $\mu(T)$ is average concentration of template T, $\sigma(T)$ is standard deviation of the concentration. Then, matching point p^* satisfies the following equation.

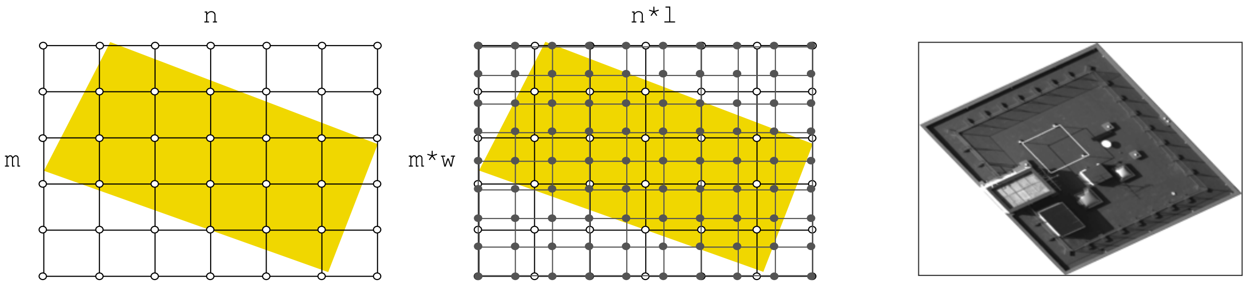
$$E(p^*) = \min_p E(p) \tag{5}$$

2.3.2 Detailed matching by correlation value

Correlation matching was selected to search a matching feature. Template matching in this phase was conducted with minimum unit width after the 1st matching. It is needed to interpolate a fore/back template as size of nadir template when calculate a correlation value. In this study, nearest neighbor method was selected as interpolation method. The advantages that the authors select a nearest neighbor as an interpolation method are stated below.

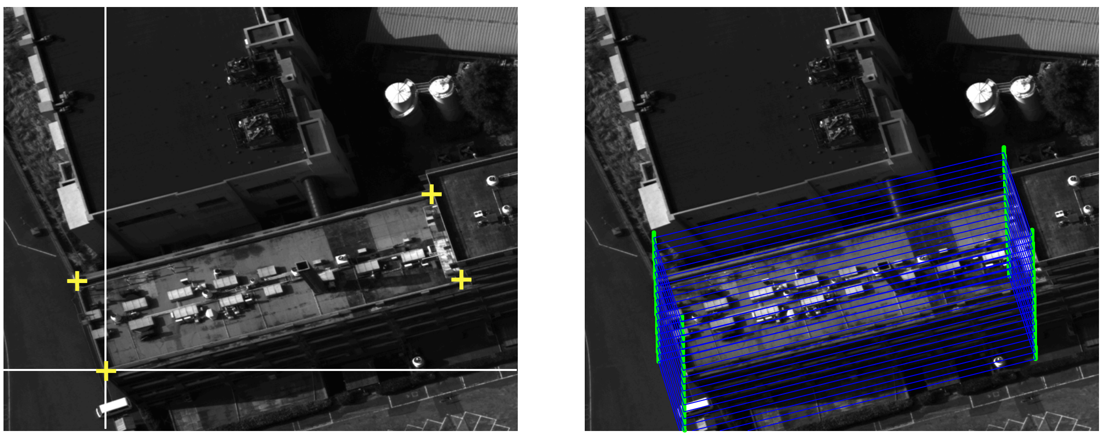
(1) Not to break the original pixel value
(2) To keep the interpolation speed fast

Flow of template matching and matching results are shown in figure 3, 4, 5, 6 and table 1

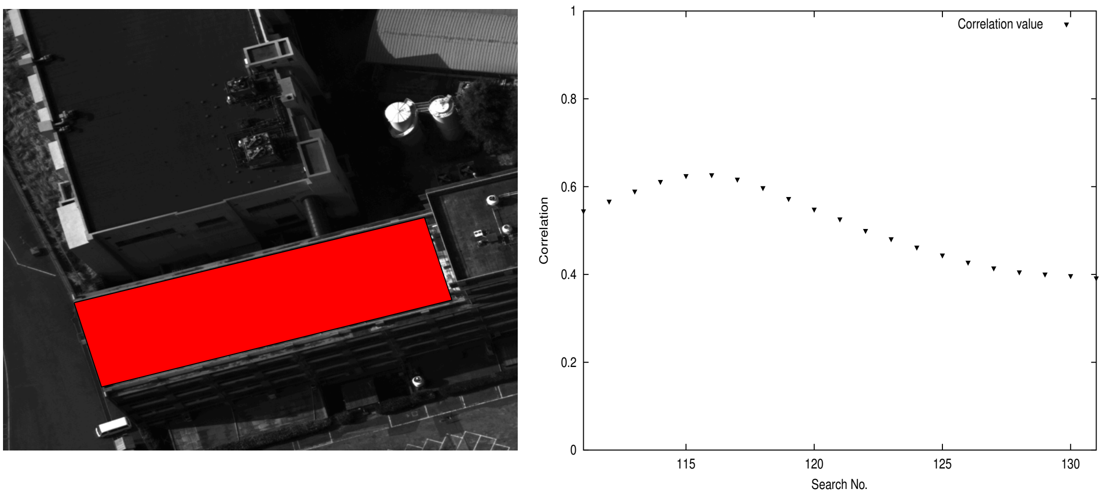


(a)Nadir template and it's MBR (b)Interpolation of Fore/Back template (c)Interpolated template

Figure 3: Interpolation of building templates by the Nearest neighbor method for calculating the correlation value



(a) Inputting the feature points (b) Searching for the matching feature



(c) Matching result (d) Transition of correlation

Figure 4: Flow of the template matching

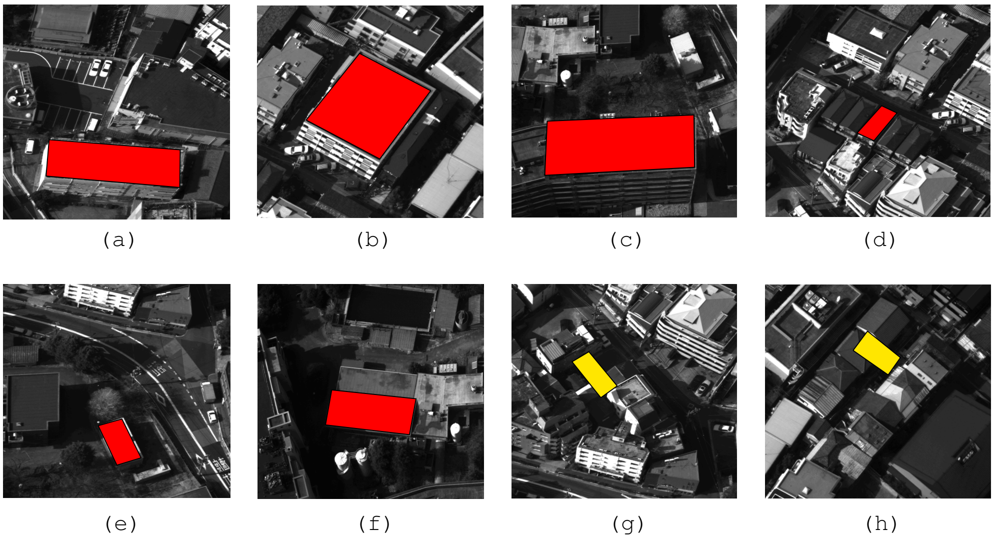


Figure 5: Examples of matching results (Red:Success, Yellow:Error)

Table 1: Examples of calculation results

Building ID	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Accuracy	Good	Good	Good	Good	Good	Good	Error	Error
Estimated height (Z)	39.20	38.80	56.90	36.30	32.10	32.80	48.30	69.10
Correlation (*r*)	0.49	0.45	0.72	0.21	0.80	0.75	-0.10	-0.16
Average (nadir)	97.64	39.26	76.21	18.75	102.71	85.07	20.28	19.35
Average (fore/back)	66.83	45.66	37.38	31.30	55.09	62.38	14.15	21.23
St.dev. (nadir)	73.46	45.90	67.43	23.61	104.19	59.98	50.34	23.66
St.dev. (fore/back)	57.23	54.50	31.60	44.67	81.96	45.87	23.16	29.30

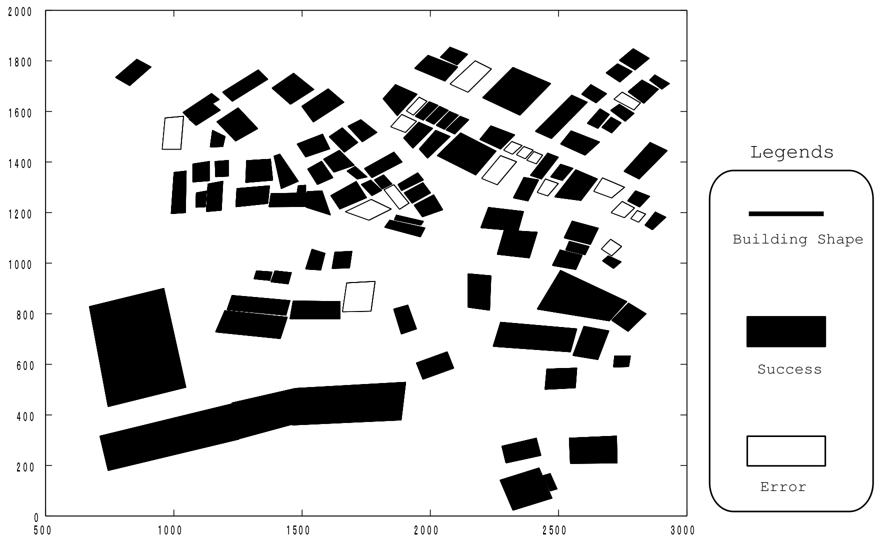


Figure 6: Success-Error distribution map

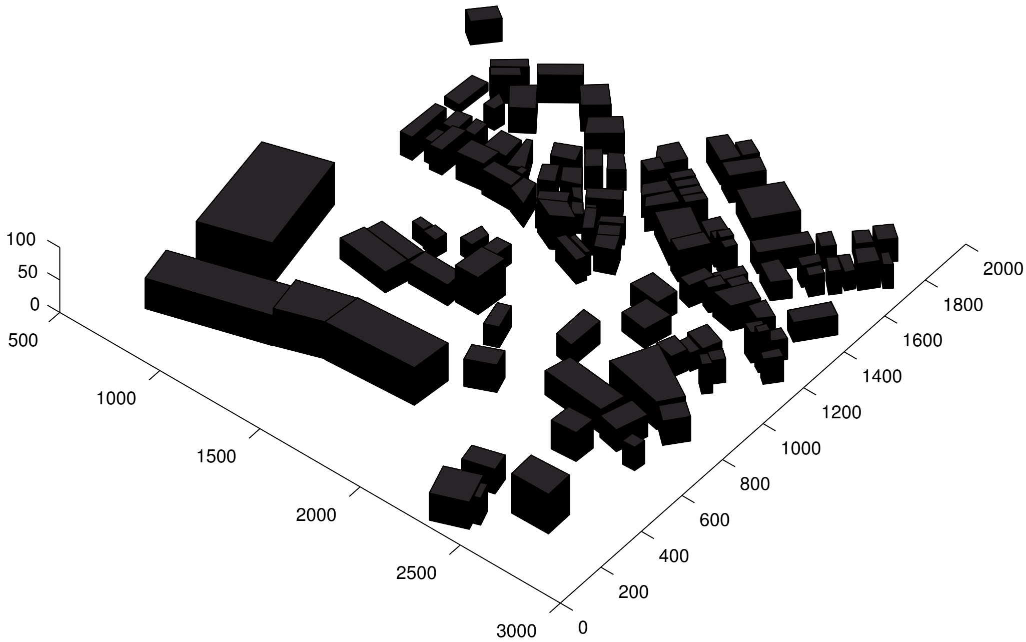


Figure 7: Constructed 3D urban model (Roppongi, Tokyo, Japan)

3 EXAMPLE OF GENERATED 3D MODEL

Generated 3D building models are shown in figure 7. Matching result is not only performed precisely, but also the building's shapes are expressed correctly.

4 CONCLUSIONS

In this study, semi automatic 3D mapping technique is developed. In this system, feature points are only inputted on the nadir image manually. The template extracted by this way was moved to match other images automatically using epipolar constraint. As a result, the semi automatic 3D mapping with good accuracy could be possible. Thus, the precise 3D model can be constructed without conducting a stereoscopic observation.

As the future works, it is necessary to automate a calculation system. Particularly, the matching templates extracted by manual edition in this time will be replaced with the results of automatic segmentation. Furthermore, the automation level will increase by developing the algorithm that edits the result of image segmentation.

ACKNOWLEDGEMENTS

The authors would like to thank STARLABO Corporation, Tokyo, Japan for their support and provision of the TLS image data sets.

References

[1] M.NAKAGAWA, R.SHIBASAKI, Y.KAGAWA, Fusing stereo linear CCD image and laser range data for building 3D urban model, ISPRS Commission IV,WG IV/7, 2002.