FUSING STEREO LINEAR CCD IMAGE AND LASER RANGE DATA FOR BUILDING 3D URBAN MODEL

Masafumi NAKAGAWA*, Ryosuke SHIBASAKI**, Yoshiaki KAGAWA*

Graduate School of Frontier Sciences, Institute of Environmental Studies*
and
Center for Spatial Information Science**
University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505
- mnaka@iis.u-tokyo.ac.jp

Commission IV, WG IV/7

KEY WORDS: 3D urban model construction, Data fusion, Line sensor image, Laser range image, Stereo matching Image segmentation, High-resolution image

ABSTRACT:

Photogrammetry is a current method of GIS data acquisition. However, as a matter of fact, a large manpower and expenditure for making detailed 3D spatial information is required especially in urban areas where various buildings exist. There are no photogrammetric systems which can automate a process of spatial information acquisition completely. On the other hand, laser range finder has high potential of automating 3D spatial data acquisition because it can directly measure 3D coordinates of objects, though the resolution is so limited at this moment. With this background, the authors believe that it is very advantageous to integrate laser scanner data and stereo CCD images for more efficient and automated acquisition of the 3D spatial data with higher resolution. Laser scanner can measure coordinate values of an object directly, but it is rather difficult to recognize the object with only laser data, for its low resolution. On the other hand, CCD image represent detailed texture information on the object's surface and has potential of detailed 3D data generation as long as careful manual operation is applied. In this research, by using laser scanner data as an initial 3D data to determine the search range and to detect possibility of occlusions, a method to enable more reliable and detailed stereo matching method for CCD images is examined.

1. INTRODUCTION

1.1 General introduction

Recently, in GIS field, vigorous demand for detailed simulations and analyses, especially, in urban areas are observed. Examples are a propagation analysis of electric wave for wireless communication, a flood analysis, an analysis of wind caused with high-rise building and landscape simulation. 3D spatial information faithful to the real world is needed for such simulations. Photogrammetry is a current method of the data acquisition. However, as a matter of fact, a large manpower and expenditure for making detailed 3D spatial information is required especially in urban areas where various buildings exist. There are no photogrammetric systems which can automate a process of spatial information acquisition completely. On the other hand, laser range finder has high potential of automating 3D spatial data acquisition because it can directly measure 3D coordinate of objects, though the resolution is so limited at this moment.

With this background, the authors believe that it is very advantageous to integrate laser scanner data and stereo CCD images for more efficient and automated acquisition of the 3D spatial data with higher resolution. Laser scanner can measure coordinate values of an object directly, but it is rather difficult to recognize the object with only laser data, due to its low resolution. On the other hand, CCD image represent detailed texture information on the object's surface and has potential of detailed 3D data generation as long as careful manual operation is applied. In this research, by using laser scanner data as an initial 3D data to determine the search range and to detect possibility of occlusions, a method to enable more reliable and detailed stereo matching method for CCD images is examined.

At the first step of a data processing flow, straight edges extraction and region segmentation have been done with the TLS image, and they are individualized. Next, removing non-building areas have been done for reliable extracting buildings. Then, 3D coordinates of features like edges and regions are computed. As a consequence, identifying individual buildings is done. Thus, finally, 3D building models are generated.

1.2 Three Line Scanner TLS

Three Line Scanner (TLS) is an optical sensor for aerial survey. TLS is composed of three linear CCD arranged in parallel, and it 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 plane, a treble stereo image of a ground object can be acquired (See Figure 1). As a result, occlusion area can be extremely reduced. Using two images of the three, it is also possible to get 3D coordinates by stereo matching.

As one of advantages of a linear CCD sensor, more pixels can be arranged in a single scene compared with an area CCD sensor. This means that a linear CCD sensor can achieve a resolution comparable with that o f an air photo. Though a linear CCD sensor can acquire data only by one line at a time (The ground resolution of TLS data in this research is 3 cm approximately). However, time of acquiring each line image is different. Since position and direction of each line when acquiring image is also different, orientation cannot be done by an existing method of photogrammetry. Moreover, the image is greatly influenced by fluctuation of an airplane position and attitude because TLS is air-borne. But setting up a stabilizer between an airplane and TLS, the fluctuation's influences can be reduced.

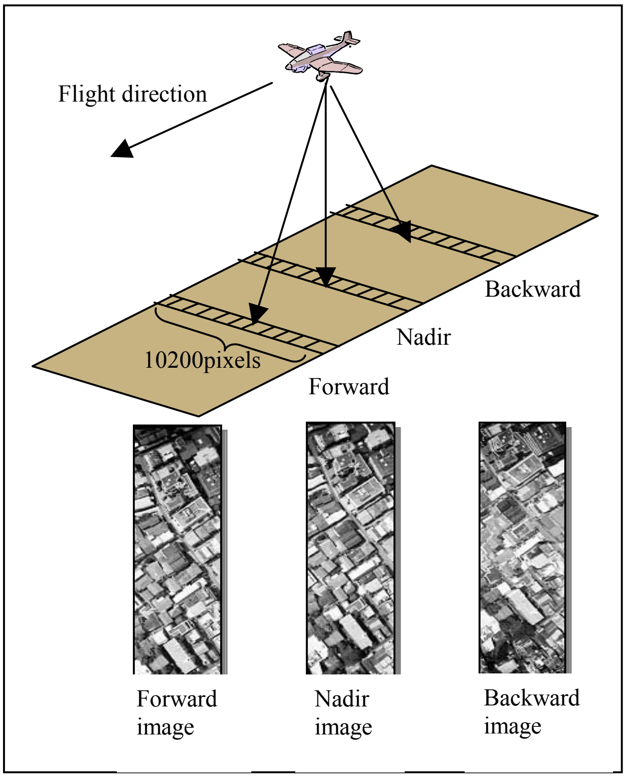

Figure 1. Method of TLS data acquisition

1.3 Laser Scanner Data

Laser scanner is a sensor, which irradiates laser to an object and measure a distance to the object by measuring returning-time as the laser reflects. If the position of sensor and the angle of irradiating laser beam are known, 3D coordinates of a place, where laser hits, can be acquired easily. And, an automation of measurement with laser scanner is simply processed too, because 3D coordinates of a measurement point can be acquired directly.

However, it is difficult to know what the object is, because laser data provides only information about object's coordinates and reflection strength. And, in general, by the restriction of laser output, a ground resolution of airborne laser is 2 m. It is not high resolution to make use of the laser data as 3D city data. To extract a building in a city area from laser data, auxiliary data is required.

Laser scanner data using in this study is acquired synchronizing with TLS. The resolution is 2 m approximately. The laser scanner data is depicted (Figure 2).

2. METHODOLOGY OF INTEGRATING LASER RANGE IMAGES AND TLS IMAGE

2.1 The outline of our approach

The outline of our approach for data fusion has following steps.

- Remove non-building areas,
- Compute 3D coordinates of features like edges and regions,
- Identify individual buildings.

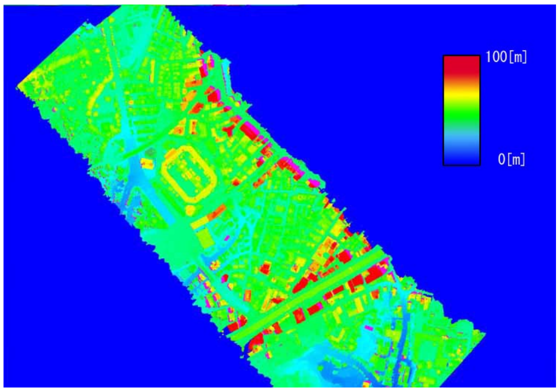

Figure 2. Data of laser scanner

Finally, 3D model can be generated. In this procedure, laser data is used for constructing rough 3D data as an initial data, and TLS images are used for refining the rough 3D data with very high-resolution images. In a sense, TLS images are used as a main data and laser data is used as an auxiliary data (See, Figure 3).

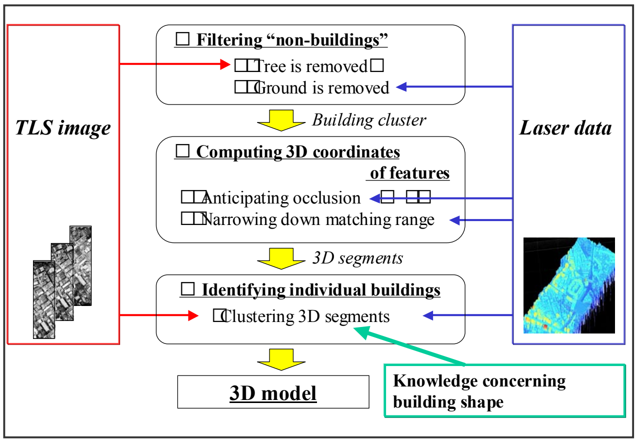

Figure 3. The outline of our approach

2.2 Filtering "non-buildings"

At the first step of the data processing flow, non-building areas such as trees or roads are removed to extract buildings. To remove tree, at first, edges are extracted from TLS image. Furthermore, straight lines are extracted from the edge image. By removing very short straight lines and curved edges, trees are remove (See, Figure 4).

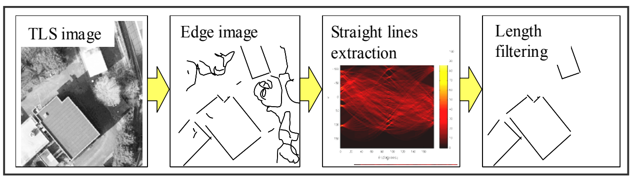

Figure 4. Removing trees

Additionally, relatively flat ground such as roads is removed approximately using the laser range image. At first, using height information of laser data, building area and ground area are divided. In the next stage, a mask image is generated with this ground area. After masking TLS image with this mask image, region segmentation is applied (See, Figure 5).

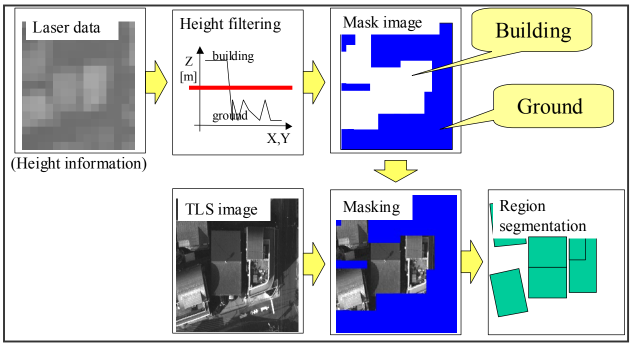

Figure 5. Removing roads

As a result, building region segments can be extracted. By doing these two-stage operations, almost non-building areas can be removed from TLS image, that means, features which may belong to buildings are extracted.

2.3 Computing 3D coordinates of features

Previously, regions and straight edges are extracted from TLS image with segmentation. In this stage, features on the forward / backward images are corresponded to features on the nadir image for computing 3D coordinates of them.

2.3.1 The method of laser range data using: Next step is computing 3D coordinates of those features. Anticipating occlusion and narrowing down search range in stereo matching using laser range images is conducted to realize more reliable stereo matching in this stage. The detail of this step is given as follows.

1) Anticipation of occlusion: In TLS image, if a feature point cannot be seen from one direction (e.g. in a forward image), the same feature point may be likely seen from the opposite direction (in a backward image). If occlusions are anticipated, appropriate image pairs (in this case, nadir image and backward image) can be selected. For anticipating the occlusion, the laser data are used as the height information (See, Figure 6).

2) Narrowing down search range in stereo matching: In stereo matching, the range of the search for feature point is relatively very wide, because there is no information on approximate height of features such as buildings. But, by using laser data, there are good estimates of height, which allow us to narrow down the range of the search (See, Figure 7).

With these methods, and it is possible to decrease mismatching and processing-time drastically.

2.3.2 Region matching: In this stage, region matching have been done for extracting outlines of building objects. Selecting a region on the nadir image, the region is cut out as a template with region information. Projecting the template on the forward / backward image, a corresponding region is searched on the epipolar line (See, Figure 8).

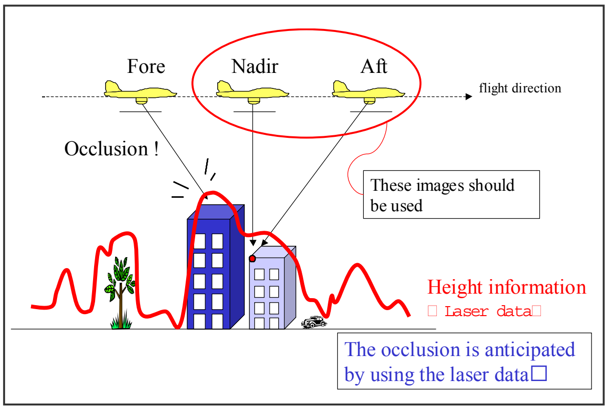

Figure 6. Anticipating of occlusion

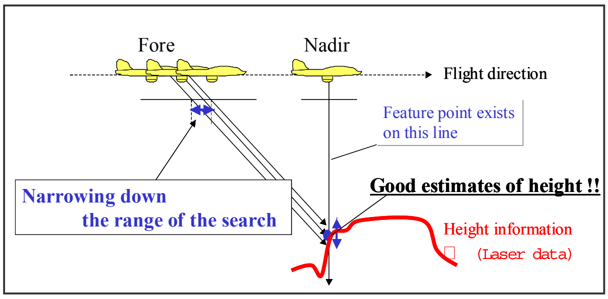

Figure 7. Narrowing down search range in stereo matching

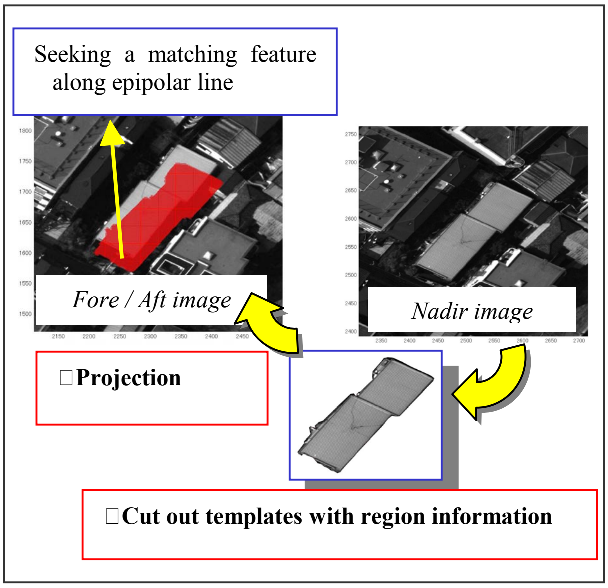

Figure 8. Outline of the region matching (1)

When region matching is done, it is necessary to define corresponding features. In this study, it is defined as follows.

Definition 1: Shapes and sizes of features are resembled.
Definition 2: Colors of features is resembled.

In this stage, regions are corresponded each other with two steps.
At first, region matching, based on feature's shapes and sizes, is applied as the first step. And, obtaining a matching coefficient, corresponded features are selected roughly. The matching

coefficient is a reference value that shows a corresponding level. In this study, a matching coefficient is calculated as follows (1).

$$R = (N / Nt)\ (N / Nf) \tag{1}$$

where N = pixel number which is overlapped
Nt = pixel number of a template
Nf = pixel number of a corresponding feature

It is calculated a value (N / Nt) which is a ratio of pixel number on a template and numbers of pixels which overlap with other features on the forward / backward image. And, it is calculated another value (N / Nf) which is a ratio of pixel number on corresponding features and numbers of pixels which overlap with a region on the nadir image. Multiplying (N / Nt) with (N / Nf), the matching coefficient R is calculated. As a result, it is used for evaluation of the corresponding level.
For example, the following figure shows a result that is obtained with the region matching (See, figure 9). The feature's label number is shown on X-axis, the seeking number is shown on Y-axis, and correlation values are shown on Z-axis. To discover the peak in this result, regions are corresponded.

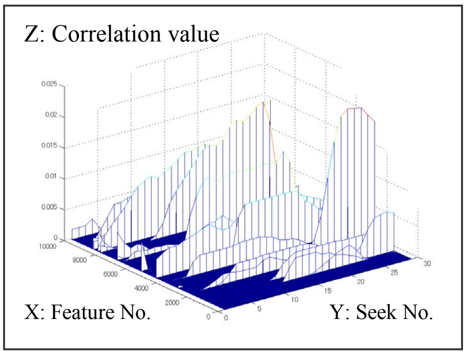

Figure 9. Outline of the region matching (2)

As the next step, region matching, based on feature's colors, is applied. In this study, the area correlation method is applied. At this time, using the result that is obtained in the first step as an initial value for the searching a matching point, the searching range can be narrowed. As a consequence, corresponded features are selected exactly.
Finally, 3D region is generated with stereo matching (See, Figure 10). The height information obtained here becomes an initial value in the next stage (edge matching).
Moreover, using the laser data, rough corresponding features can be done instead of the first step, because the laser data can be used as an initial height value for the region matching. Therefore, substantially, region matching becomes one step only of the use of the area correlation method as the region matching.

2.3.3 Straight edge matching: In this stage, straight edge matching is done for extracting details of building objects. To select a straight edge around a region in the nadir image, it is projected in the forward / backward image. Also, a corresponding straight edge is searched on the epipolar line (See, Figure 11). At this time, the method in the first step of the region matching is applied.

Moreover, edges are corresponded with similar informations of edges as a clue for the matching in this stage. The clues are as follows, a direction of the edge, a length of the edge, the average density on both sides of edge, a relation concerning position with region, and so on (See, Figure 12).

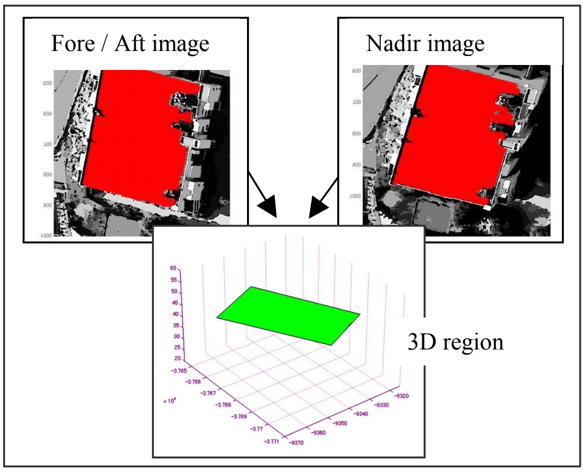

Figure 10. Generating 3D region

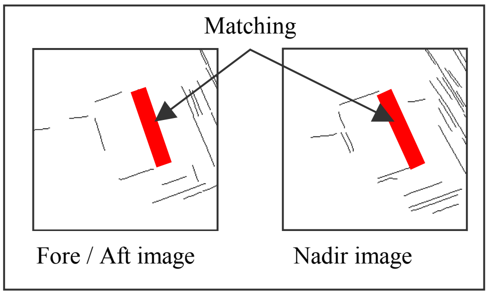

Figure 11. Straight edge matching (1)

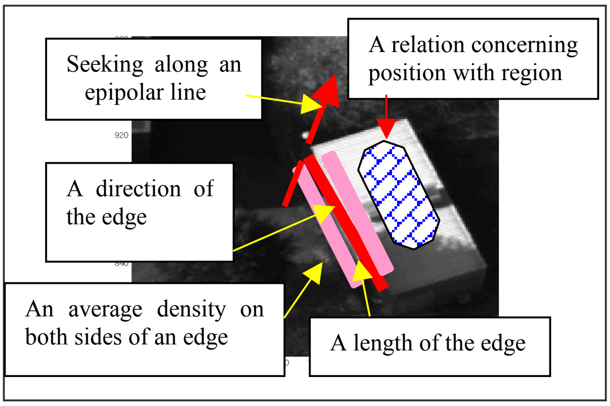

Figure 12. Straight edge matching (2)

Finally, calculating 3D coordinates with forward intersection, 3D lines are generated with stereo matching (See, Figure 13).

2.4 Identifying individual building

Generating 3D region and 3D lines with the region matching and straight edge matching, they are overlaid (See, Figure14).

Figure 14 shows a cross section of the segments. Red arrow is indicating a height level that many 3D lines gather. At the next step, the 3D region is updated with 3D lines, because generated

3D lines with edge matching have more correct values than 3D regions which are generated with region matching (See, Figure15).

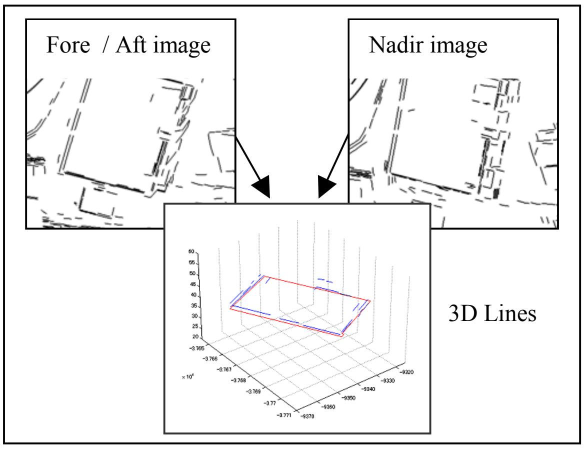

Figure 13. Generating 3D lines

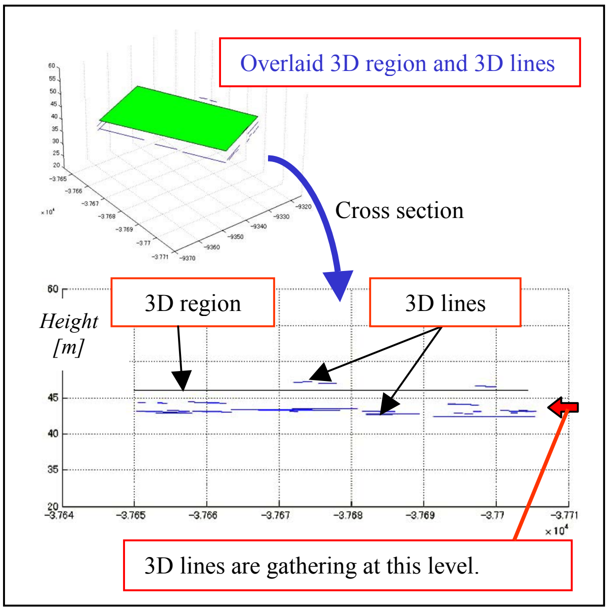

Figure 14. Generating 3D building parts (1)

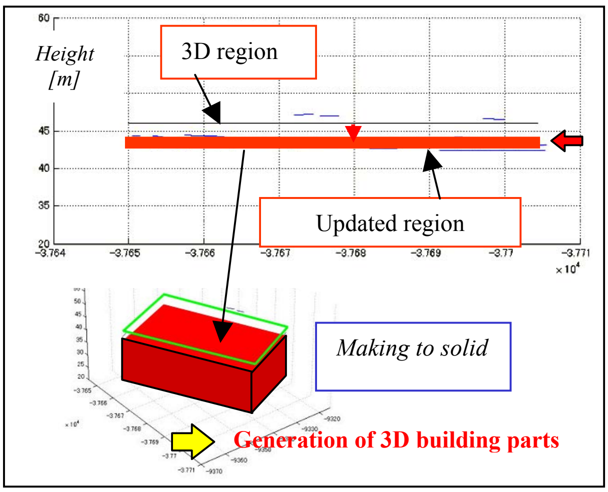

Figure 15. Generating 3D building parts (2)

Then, a solid model is generated with lofting the 3D region. This model is called 'building parts'.
However, a real building cannot be expressed enough only by this model. Then, a building object is generated with merging adjoining 3D building parts (See, Figure16). In this way, 3D segments are clustered.

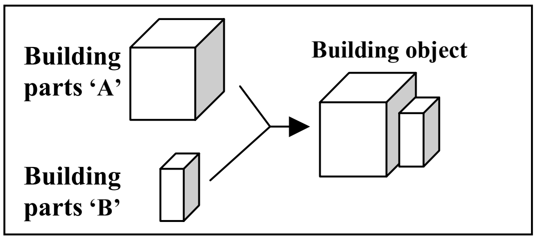

Figure 16. Merging 3D building parts

3. EXPERIMENT

3.1 Study area

The study area is Roppongi, Tokyo, Japan. In this area, buildings with various shapes exist. If buildings are classified by sizes, there are various building such as houses and multistory buildings. If buildings are classified by roof's shapes, there are various building such as gable roofs and flat roofs. Besides the building, there are roads, trees, and parking lots, etc. The terrain of this area is full of ups and downs.

3.2 Outline of the method

The region matching with a laser data and the region matching without a laser data are done respectively. The match accuracy is calculated for 40 buildings (86 building parts). As a result, how using the laser data is effective is considered.

4. RESULT

4.1 Matching result

The matching result is evaluated by measuring the gap of the direction of the line (vertical direction in the image) of regions (See, Figure17). And, the result (region matching) is shown Figure18. Blue points show matching results without laser data, and red points show matching results with laser data. Figure 18 shows that fusing TLS images and laser data makes result better.

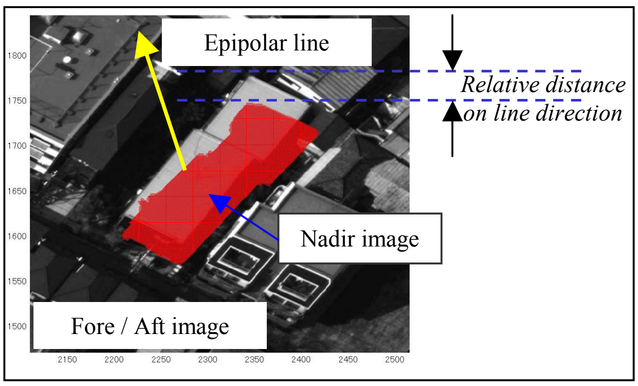

Figure 17. How to evaluate matching results

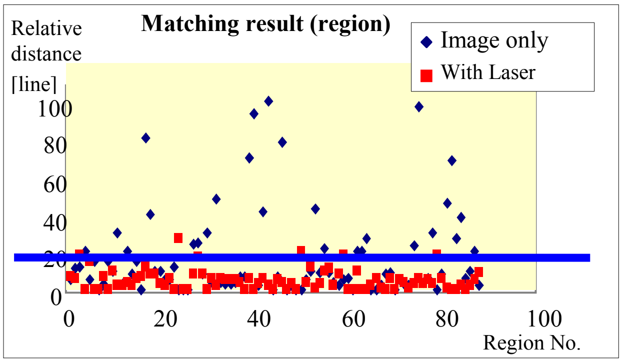

Figure 18. Matching results (Region matching)

By the way, defining 20 line as a threshold of success / failure, the results can be given in the following table (See, Table 1).

	Image only	With Laser data
Success rate	68% (59 / 86 objects)	98% (84 / 86 objects)
CPU time (86 objects) (pentium4 1.8GHz)	711[s]	363[s]

Table 1. Success rate of matching

Table 1 shows that using the laser data makes the success rate in the matching almost 100%. And it makes calculating-time about half. This is just an effect of fusing TLS image and laser data.

4.2 Modeling result

Modeling results are shown as follows. The figure 19 shows a solid model, and the objects are classified by buildings.
Moreover, using TLS nadir image, a textured 3D model is shown figure 20.

5. SUMMARY

In this research, combining TLS images with laser scanner data, a more efficient method of making 3D spatial information with higher resolutions is examined.

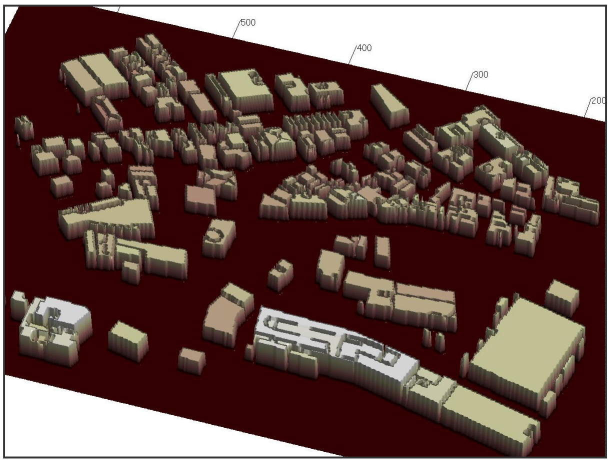

Figure 19. Modeling result (Solid model)

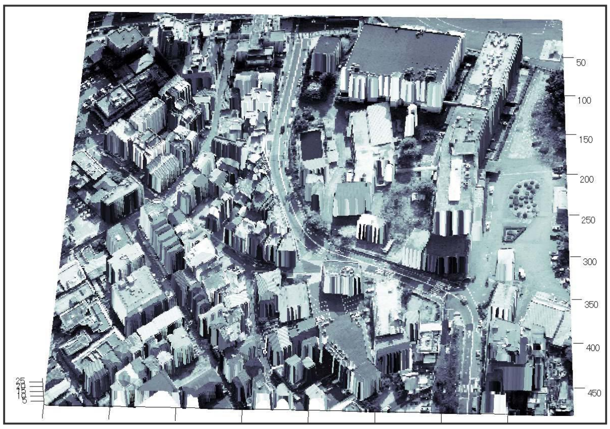

Figure 20. Modeling result (Textured model)

As a conclusion of this research, it can be concluded that basic shapes of buildings can be possible to generate full-automatically by fusing TLS images and the laser data. However, region segmentation has to be done under a good condition.
Moreover, the automation level improves dramatically compared with the method using only images. And a more detailed modeling can be done compared with the method using only laser data.
As effects of fusing the stereo image and the laser data, the authors can enumerate the improvement of the matching accuracy (improving it from 68% to 98%) and shortening of calculating-time (about 50%) for the matching.
The data obtained here becomes a basic 3D spatial data for city modelling, and the authors believe that it becomes a big support to digital mapping.

References

[1] Yoshiaki, KAGAWA, 2001. Automatic acquisition of 3D city data with air-borne TLS (Three Line Scanner) and Laser Scanner. In: *Graduate School of Frontier Sciences, Institute of Environmental Studies, University of Tokyo,* Japan.

[2] Mikio, TAKAGI, & Haruhisa, SHIMODA. HANDBOOK OF IMAGE ANALYSIS.