Ego-Motion Estimation for Efficient City Modeling by Using Epipolar Plane Range Image Analysis

Shintaro Ono   Hiroshi Kawasaki   Kiyotaka Hirahara
Masataka Kagesawa   Katushi Ikeuchi

Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505 JAPAN
Tel: +81-3-5452-6242   Fax: +81-3-5452-6244
{onoshin,h-kawa,hirahara,kagesawa,ki}@cvl.iis.u-tokyo.ac.jp
http://www.cvl.iis.u-tokyo.ac.jp

Abstract

This paper proposes a novel notion of Epipolar Plane Range Image (EPRI). A line-scanning laser range sensor is mounted on the side of our data acquisition vehicle, repeating drawing horizontal scanning lines. Laminating the line range data along time axis, we can follow temporal continuity of horizontal cross section of the geometry seen from the scanner, hence can estimate the motion of the vehicle. Applying these information to another vertical-line-scanning range data we can align the position of the scanning lines, and can efficiently get correct 3D geometric model of the real urban space, even if the vehicle travels in arbitrary speed without any external devices as GPS.

INTRODUCTION

Construction of a 3-D geometric city model in a virtual world has become a highly interested research topic among a lot of research fields, such as computer vision/graphics, virtual/mixed reality, sensing, architectonics, etc. Urban environment models are used in various applications, including urban planning, disaster prevention, intelligent transport systems, or propagation simulation of radio waves for the cell phone industry.

Such extensive environments as urban scene are not able to be scanned by one sensing device at once, because current, commercially available sensing devices can not cover whole of the enverionment without occlusion and with enough resolution at the same time. General solutions for this problem are as follows:

- “Stop-and-go” scanning
- “Continuous” scanning

“Stop-and-go” scanning is a method to scan objects from fixed point, and move to another point, scan, repeatedly. Since the scanning is done from fixed points, the acquired data becomes relatively dense and accurate. Meanwhile, it takes long time to apply to extensive environment as urban scenes.

“Continuous” scanning is a powerful and efficient method to solve the demerit of “Stop-and-go” scanning. It is a method to scan objects with the sensor continuously moving. The sensor is mounted on some kind of movable body, such as vehicle, helicopter, airplane, etc. This method is suited to scan extensive areas, however, the ego-motion of the scanner, i.e. the velocity/position of the vehicle, helicopter, airplane, etc. must be known by some techniques.

The most popular solution for this problem is to mount other external devices as GPS or INS on the scanning vehicle([3]). It is a highly simple and convenient technique, and the position or the velocity can be acquired in an easy way. However, the accuracy of the positioning result muchly depends on the condition, especially on the reception situation of the radiowave signal. Though the situation improved since the scramble noize addition to the GPS signal has been abolished, the accuracy is still not enough in the ground level of urban area, because of the occulusion by buildings or raised expressway road.

Another approach is proposed by C.Frueh et.al([4]). They mount a line-scanning laser range finder on the data acquisition vehicle, and let it sweep horizontal scanning lines. By matching each result of scanning frame pairwise, the ego-motion of the data acquisition vehicle can be acquired. They constructed 3-D virtual textured model of the urban space by using this technique([5, 6]).

We also agree to the spirit of this approach: external devices as GPS are suited to acquire initial position of the scanner, and detail position is acquired by computational process of acquired geometric data. In concrete terms, the positional information is acquired by horizontal-scanning laser range finder on our data acquisition vehicle in a similar way. But one of the characteristic aspect of our approach in contrast to [4] is the consideration of temporal continuity of each scanning frames. We propose a novel notion of Epipolar Plane Range Image(EPRI) from Epipolar Plane Image(EPI), which is a classical analysis method of moving image. EPRI simultaneously represents spatial feature and temporal continuity. We estimate the ego-motion of our data aquisition vehicle by analyzing the EPRI.

This paper is composed of five sections. After the introduction is brought up in this section, the notion of epipolar plane range image which we propose is described first. Secondly, the EPRI technique is adopted to estimate the ego-motion of the data acquisition vehicle, with outdoor experiment in practice. Finally, the paper is concluded.

EPIPOLAR PLANE RANGE IMAGE

In this section, we introduce a novel notion of Epipolar Plane Range Image(EPRI), which is a coined term in this paper. EPRI is a spatio-temporal range image representing both spatial feature and temporal continuity simultaneously. The notion of EPRI is derived from Epipolar Plane Image (EPI). Here we firstly explain the EPI analysis, and secondly the EPRI.

EPIPOLAR PLANE IMAGE

Epipolar Plane Image(EPI) analysis is one of the well-known methods to analyze moving image, especially in the field of computer vision. The notion of EPI was firstly proposed in [1, 2], and various kinds of applied reserches has been proposed. EPI can be created by moving a line camera ($n \times 1$ pixel) horizontally, and stacking each image frame vertically(Fig.1).

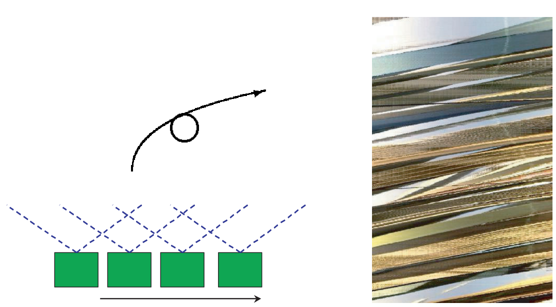


Figure 1: Epipolar Plane Image (EPI)

As shown in Fig.1, when the camera moves horizontally, pixels in each frame of the EPI which represent same point in real world compose a continuous edge. The edge image includes various slopes. These variety of the slopes are derived by the parallax, which is the result of the difference of horizontal position of the camera. Fig.2 represents the change of parallax when the camera moves from C1 to C2. From this figure, the depth D (horizontal distance from the camera) and the parallax u are related with

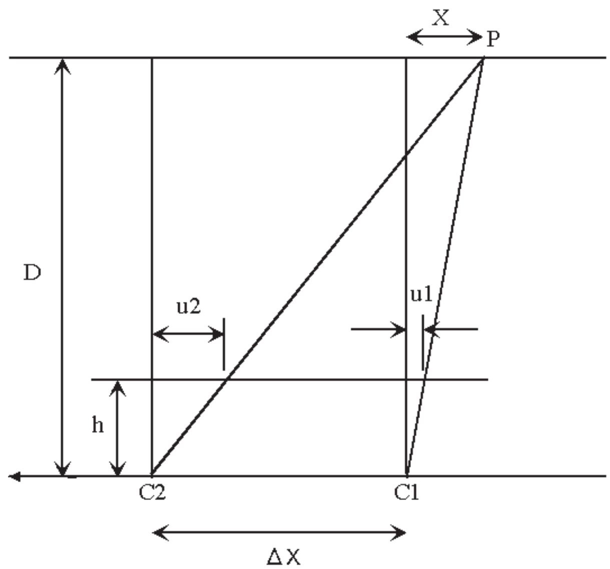


Figure 2: Lateral motion

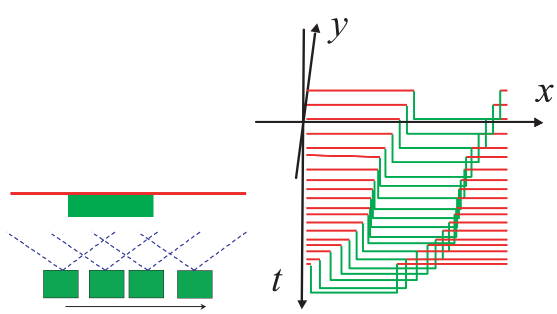


Figure 3: Epipolar Plane Range Image (EPRI)

moving distance ΔX.

$$\Delta U = u_2 - u_1 = \frac{h(\Delta X + X)}{D} - \frac{hX}{D} = \Delta X \frac{h}{D} \tag{1}$$

And the slope of the edge m is related to the depth D of the point which composed the edge, as follows:

$$mV \propto D \tag{2}$$

where V is the moving velocity of the camera. By using this equation, D can be estimated from m, V; e.g. the geometry of the object is reconstructed to some extent.

Edge detection is usually processed by image binarization and hough transformation. Because of this, in EPI analysis the camera is supposed to move in constant speed.

EPIPOLAR PLANE RANGE IMAGE

Here we propose a notion of EPRI. It can also be created in a similar way as EPI, but the sensing device is not a line camera, but a laser range sensor which can also get n points in a horizontal line per one frame. Fig.3 shows the model of EPRI.

Table 1 shows the contrast of EPI and EPRI. EPRI has some characteristic features, compared with EPI.

Table 1: EPI and EPRI

	EPI	EPRI
Sensing device	image sensor	range sensor
Depth information	known	unknown
Depth and the slope of the edge	dependent	independent
Sensor's moving speed	assumed to be constant	not assumed to be constant
Pixels in the (range) image	even	uneven

1. Depth of each point is already known
2. Depth of a point and the slope of the edge are independent.
3. Moving speed of the sensor is not assumed to be constant.

1. is due to the essential feature of the sensing device, range sensor.

2. is because a range sensor does not have a projection plane as a camera. The difference of the slope of edges in ordinary EPI is arisen by the existence of projection plane. In EPRI, the distance which a feature point (ex. corner) moves between adjacent frame is equal to V/f, where V is sensor's moving velocity, and f is frame frequency of the sensor. Instead of Eq.2, the following equation is lead in EPRI.

$$mV = const. \tag{3}$$

3. is because edges in EPRI are easily detected compared with EPI. In EPI, Hough transform is used for edge detection, hence the edge must be straight. This means that the moving speed of the camera is assumed to be constant. On the other hand, each point in EPRI composes some cluster planes (Fig.4). These planes are geometrically separated. Therefore, their edges are easily estimated, even if they are not straight. This means that the moving speed of the range sensor is not assumed to be constant.

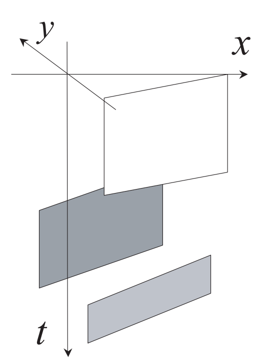


Figure 4: Each range point in EPRI composes cluster planes

Now we propose a novel way to utilize EPRI. In EPI, the depth of each point is unknown variable, and it can be estimated by Eq.2, where the moving speed of the camera is assumed to be constant. On the other hand, in EPRI, it can be the moving speed of the sensor that is assumed to be unknown and inconstant variable, because the depth of each point is given. It will be inversely estimated by using the same method as EPI analysis (Eq.3). Then, the horizontal position of the sensor will be also estimated(Fig.5 upper).

Moreover, the horizontal position *against* the object (the moving speed approaching the object) can be also estimated from EPRI(Fig.5 lower). This is because this problem is returned to estimate certain parameter on the cluster planes(ex. curvature).

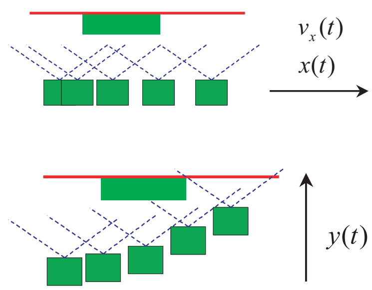


Figure 5: Sensor's position can be obtained from EPRI

As mentioned above, EPRI analysis can be utilized as a kind of "positioning system". The following section will introduce a concrete technique for position estimation.

EGO-MOTION ESTIMATION FOR URBAN SPACE MODELING

In the previous section, we introduced an EPRI and described that the motion of the sensor can be estimated by analyzing it. In this section, a concrete technique for the estimation is explained. Also, our data acquisition system is introduced for 3-D geometry reconstruction.

ESTIMATION ALGORITHM

Here we explain the concrete algorithm for sensor motion estimation. This time we assume that the sensor moves in a variable speed, but straightly. Fig.6 shows the process of estimatating sensor velocity and position. The process is described as follows:

1. Data acquisition and creation of EPRI from scanning data.
2. Segment each cluster in EPRI.
3. Fit an analytical curve to each segment.
4. Smoothly connect the derivative of the curve.

The details are described in the following.

Step 1

A line-scanning laser range sensor is mounted on our data acquisition vehicle(Fig.9). The detail system is described in the next subsection. EPRI can be created by vertically stacking each frame of horizontal-directional range data.

Step 2

As described in previous section, Each point in EPRI actually make cluster as shown in Fig.7. In this step, these clusters are carved into each segment. The result of the segmentation is shown in Fig.8. Each segment represents transition of our data acquisition vehicle.

Step 3

In the each segment, the scanned object are represented as moving gradually. This is equivalent to the motion of the sensor itself. Actually, however, the segment are not composed of even plane because

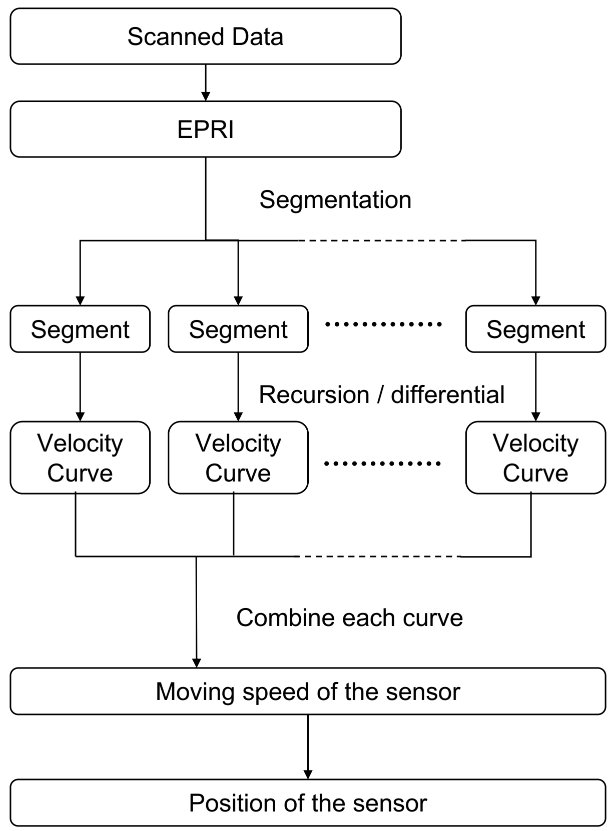


Figure 6: Flow of the estimation process of sensor velocity and position

of scanning noise. In addition, scanned range points are discretely distributed and the edge of each segments are not smooth, because the resolution of scanning angle or scanning frequency is finite.

On the other hand, the range sensor is mounted on a vehicle in this research. Generally, the movemant change of a vehicle is considered to be smooth due to its mechanical acceleration principle as long as intentional sudden acceleration or stop is not carried out, To get the general tendency of transition, we carry out the regression analysis and fit an analytical function to each segment, for vehicles can be assumed to make continuous and smooth movement in general.

Step 4

Though the regression curves describe the transition of our data acquisition vehicle, they are valid in each segment only. The segments are separated, therefore it is impossible to obtain the transition from start to end directly. Here we estimate the velocity of our vehicle. The derivative of each curve represents the velocity of our vehicle. By smoothly connecting each curve acquierd from each segments, the time variation of the velocity from the start point to the end point is obtained.

Finally, by integrating the time variation of the velocity, we obtain the transition history of the vehicle. Using the positioning information, we can align the vertical scanning line according to the correct position, even if the velocity varies.

OUR SYSTEM AND RECONSTRUCTION OF URBAN MODEL

Fig.9 shows our data acquisition vehicle. Four line laser range sensor is mounted on the vehicle: one repeats horizontal-diredctional line scanning to acquire EPRI, and the other vertical to get the 3-D geometry of the object in the urban scene.

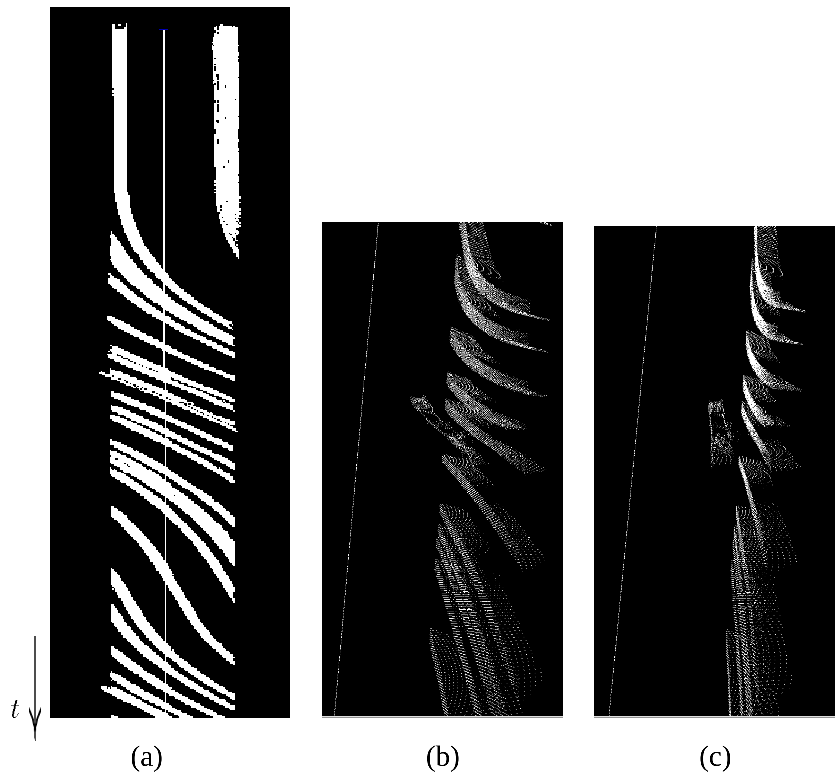


Figure 7: An actual example of Epipolar Plane Range Image. (a) Both spatial and temporal continuity can are simultaneously represented. (b)(c) Each point make cluster.

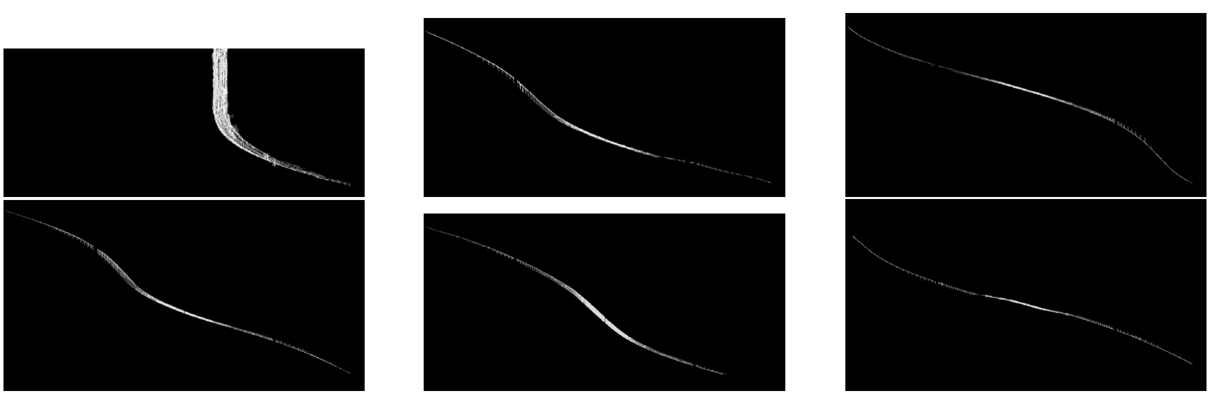


Figure 8: Segmented EPRI

Figure 9: Our data acquisition vehicle

Figure 10: Photo of our campus building: vertical columns are actually arranged at every 6m.

Urban model is reconstructed by arranging vertical-directional scanning lines to appropreate positions. When the vehicle travels in arbitrary unknown speed, the result of the arrangement becomes expanded or shrinked before motion estimation, assuming that the vehicle had traveled in a constant speed. After the motion estimation by the method described in previous section, the arrangement result will become appropreate.

EXPERIMENT AND DISCUSSION

OUTDOOR EXPERIMENT

We have made an outdoor experiment to confirm that our alogrithm works well. The place was our campus, Institute of Industrial Science, The Univ. of Tokyo (Fig.10). This time we assumed a linear motion to the vehicle, parallel to the wall of our building.

The aspect of the regression analysis is shown in Fig.11. We fit 6-dimensinal polynomial function to the cluster segment. Applying the process, we could estimate the time variation of the velocity of the vehicle as Fig.12. It approximately agrees to the correct answer, which was manually calculated from actual measuring of our building and the frame rate of the sensor. The maximum error was 8–12%.

According to the result of estimation, we can align the position of vertical scanning line. Fig.14 shows the result of modeling our building, with before and after of the alignment process. Columns of the buildings are arranged in equal pitch, as seen in Fig.10. Fig.13 shows the detail of the result. The gap between each scanning line is adjusted according to the vehicle's velocity.

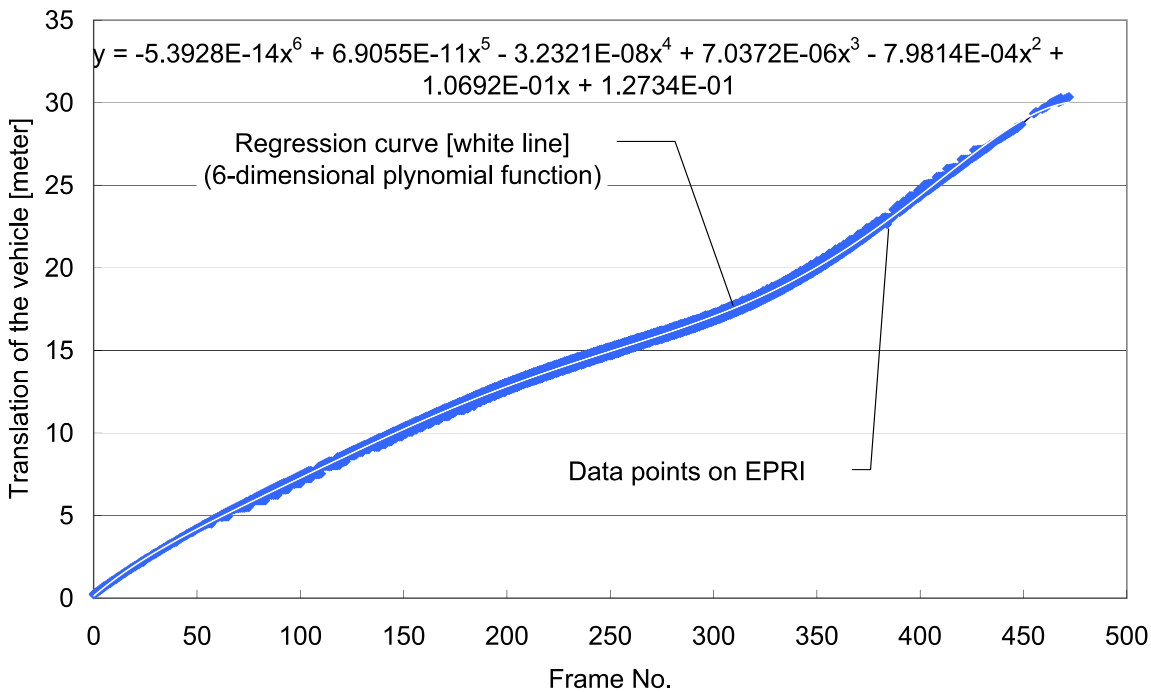


Figure 11: Regression

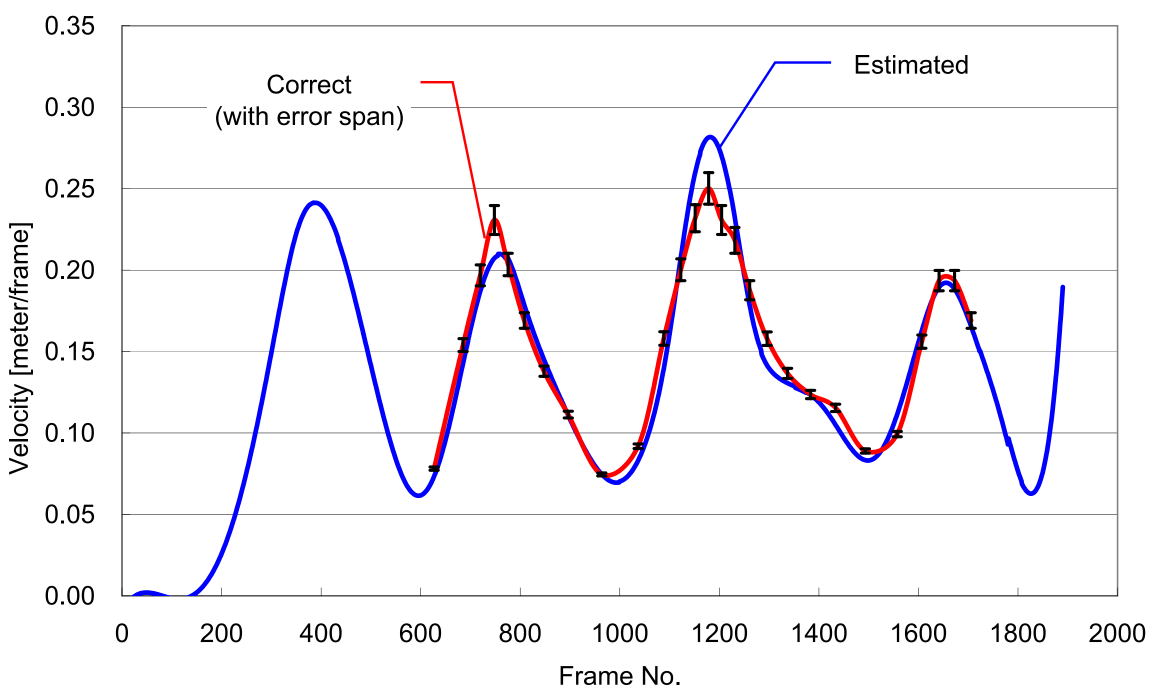


Figure 12: Estimated velocity

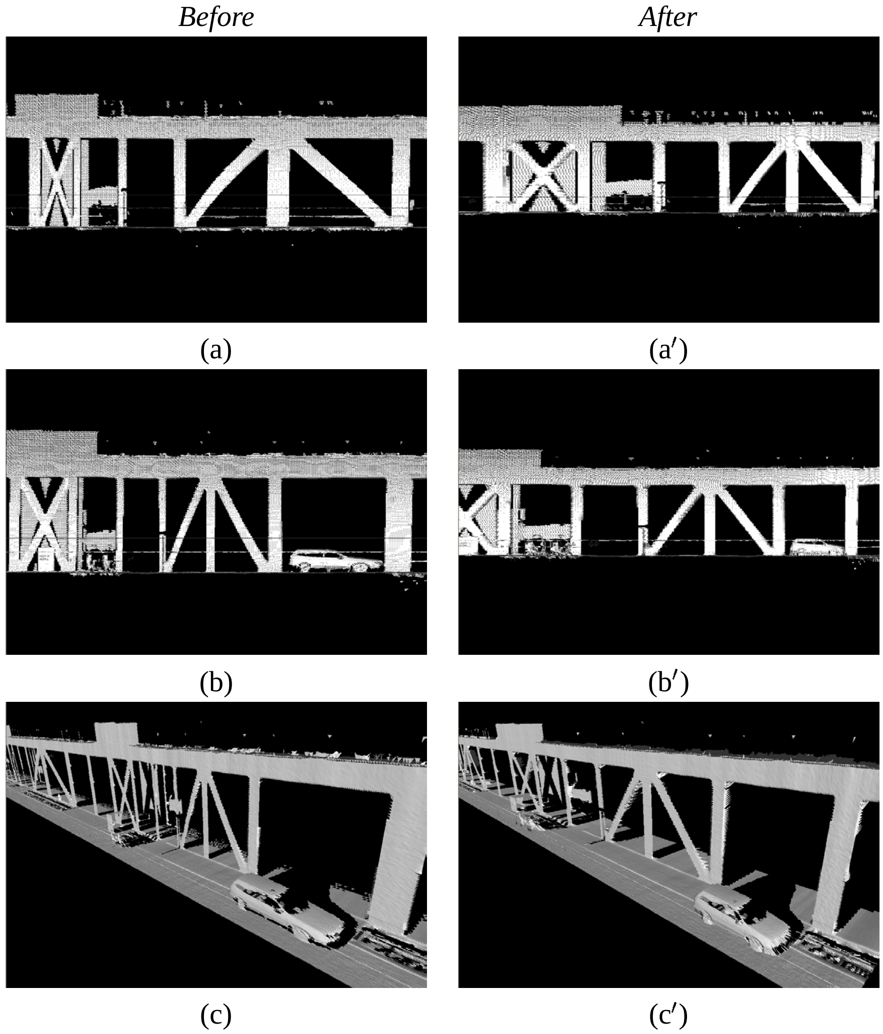


Figure 13: The detail of modeling result, with before(a,b,c) and after(a′,b′,c′) the correction of the vehicle velocity.

DISCUSSION

The experimental result approximately agrees to the correct answer with 8–12% of maximum error. Regarded causes to the error can be raised as follows:

- Regression by a single polynomial:
  In this experiment, the regression process was carried out by one polynomial per segment. The regression equation is not needed to be a polynomial. It must be examined to use more suitable equation for vehicle dynamics.

- Completed process in one segment:
  In this experiment, motion of the vehicle was estimated from each segment by completed process in each segment, and finally they are smoothly combined to estimate whole result. However, the neighbor segment to a certain segment must represent approximately same velocity. Therefore, the accuracy will be improved by considering inter-segment continuity in the regression step.

- No use of reflectance:
  Reflectance values can not be acquired from the laser range scanner used in this experiment. By using reflectance edge in addition to geometric edge in EPRI, the accuracy will be improved.

CONCLUSION

In this paper, we firstly proposed a novel notion of Epipolar Plane Range Image The feature of epipolar plane range image is that it shows both spatial and temporal contiuity simulatneously, and helps obtaining general tendency of the movement, compared to match each scanning frame piecewisily.

And by estimating the velocity of the data acquisition vehicle with this theorem, we have aligned the vertical scanning line which was obtained from another sensor on the vehicle. The vertical scanning lines are correctly aligned according to the vehicle velocity.

All estimation process have been done without using any external positioning devices as GPS, which does not have enough accuracy for geometrical alignment in general urban space with the low radiowave sensitivity by buildings or skyways.

The future works are raised as follows:

- Accuracy improvement:
  The accuracy of the estimation result will be improved by the approach mentioned in the discussion previous section.
- Fully arbitrary motion:
  This time we assumed arbitrary speed and linear motion to the vehicle. By easing this constraint, the EPRI becomes to make curved surface.
- Further experiment in various environments:
  In this experiment, the scanned objects are mainly the column of our building. The experiment in more complicated environments will be inevitable to present the versatility of our method.
- Texture mapping:
  Mounting not only laser range sensors but cameras, the texture of urban scene is acquired.
- Calibration of multiple sensors:
  Position and orientation of each sensor will be estimated by analyzing EPRI, simultaneously with the motion of the sensor using parametric alignment technique as [7].

References

[1] R.C.Bolles, H.H.Baker, D.H.Marimont : "Epipolar-plane image analysis: an approach to determining structure from motion", International Journal on Computer Vision, 1, 7-55, 1987

[2] H.H.Baker, et.al : "Generalizing epipolar plane image analysis on the spatio-temporal surface", International Journal on Computer Vision, 3, 33-49, 1989

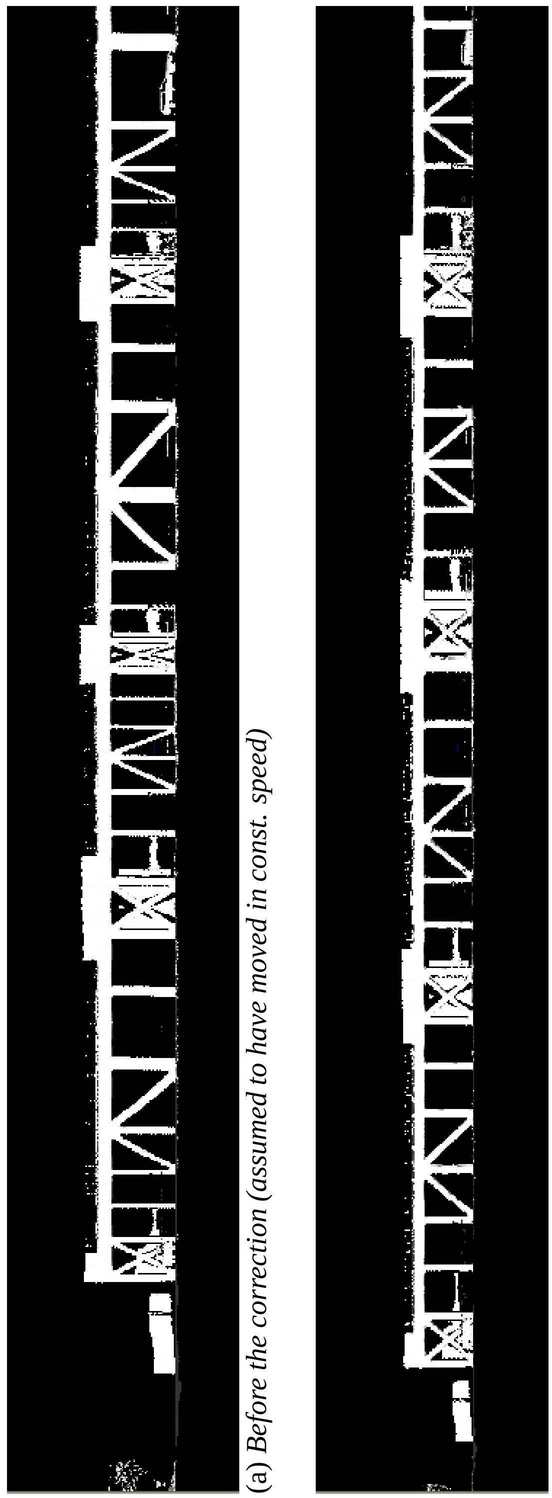
[3] Huijing Zhao, Ryosuke Shibasaki: "Reconstructing Urban 3D Model using Vehicle-borne Laser Range Scanners", International Conference on 3D Digital Imaging and Modeling (3DIM), 2001

[4] Christian Frueh, Avideh Zakhor: "Fast 3D Model Generation in Urban Environments", IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems (MFI), 2001

[5] Christian Frueh, Avideh Zakhor: "3D Model Generation for Cities Using Aerial Photographs and Ground Level Laser Scans", Proc. IEEE Computer Vision and Pattern Recognition (CVPR), 2001

[6] Christian Frueh, Avideh Zakhor: "Constructing 3D City Models by Merging Ground-Based and Airborne Views" Proc. IEEE Computer Vision and Pattern Recognition (CVPR), 2003

[7] E. Boyer and J. S. Franco: "A Hybrid Approach for Computing Visual Hulls of Complex Object", Computer Vision and Pattern Recognition, 2003



(a) *Before the correction (assumed to have moved in const. speed)*

(b) : *After the correction with estimated velocity*

Figure 14: Modeling result: (a)before and (b)after the correction of the vehicle velocity.