MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

13—35

From Digital Map to 3D Map: Creating 3D Map by Motion Stereo Utilizing 2D Map

Masafumi UEHARA and Heitou ZEN *
Tokyo University of Mercantile Marine

Abstruct

This paper presents the framework for creating of 3D map. The 3D shape and the textures of a building are acquired by analyzing moving picture taken from a running vehicle. However, in the real city environment, it is difficult to obtain 3D data of buildings accurately, because the slight motion of images due to the vehicle motion cannot be removed completely. This paper describes a approach to reconstructing 3D city model utilizing existing 2D digital map in addition 3D measured data. The obtained shapes and textures are compared with 2D maps to refine their location and shape. The implementation of the alogorithm has been applied to several real urban scenes and successfully created 3D map.

1 Introduction

3D model of urban scene plays an important role in many fields, such as city engineering, traffic engieering, amusement. Under present conditions, these models are made manually with CG modeling tools. Hence, huge time and cost are wasted, then the utilization of these models are restricted. Therefore, in order to construct model efficiently and automatically, many approaches have been suggested. These approaches are roughly divided into two methods. First methods is called IBR(Image Based Rendering) using stored 360° real images from observation point without 3D data[1][2]. It is possible to realize photorealistic urban scene with this method. But on the other hand, it is necessary to use a large quantity of data to utilize many images. The second method is to make model using measured 3D data of buildings with the stereo method[3] or a laser measurement[4]. This method enables to move a viewpoint freely and construct model with a small qantity of data. We also constructed urban scene model based on 3D data.

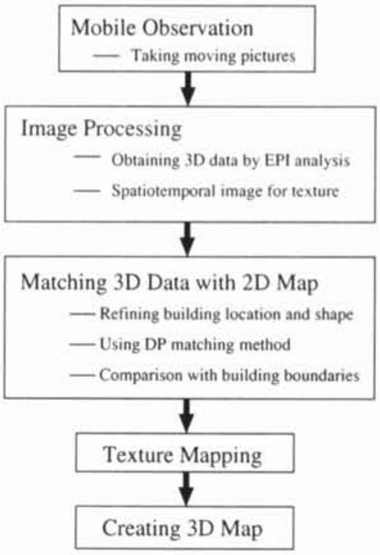


Figure 1: Summary of our approach

The scheme of the method we proposed is shown at Figure 1. First, 3D data of buildings are obtained from continuous images of urban scenes by using Epipolar Plane Image (EPI) analysis[5][6]. The textures are obtained from slit spatiotemporal images. The slit spatiotemporal images are gathered pixels on a horizontal slit from continuous images[7]. Next, to refine building location and shape, the obtained 3D data are compared with existing 2D digital maps by DP matching method (see Figure 2). Then, slit spatiotemporal images are divided several images into each buildings which are utilized as building textures. Finally, the shape of buildings based on a digital map are constructed, and the textures are mapped onto these shape of buildings. The 3D map of the city can be created by taking moving pictures of all streets in the city.

In this paper, we mention the method to obtain 3D data and textures of buildings in chapter 2. Chapter 3 describes about the matching method between measured 3D data and a 2D digital map. The experimental results are shown in chapter 4, and chapter 5 describes the conclusion of this research.

---

*Address: 2-1-6 Etchujima, Koto-ku, Tokyo 135-8533, Japan. E-mail: `{muehara,zen}@ipc.tosho-u.ac.jp`

Figure 2: Example of 2D digital map

2 Obtaining 3D data and textures

The depth data from camera path to buildings are obtained with Extended EPI analysis improved by Notomi et al[7]. It is necessary to prepare many cameras for wide areas observation in EPI analysis. In EEPI, all flows of stationary objects appears from FOE (Focus of Expansion) are picked up by the camera whose the optical axis is parallel to the camera path. Thus a wide region can be observed at once by a small number of camera (see figure 3).

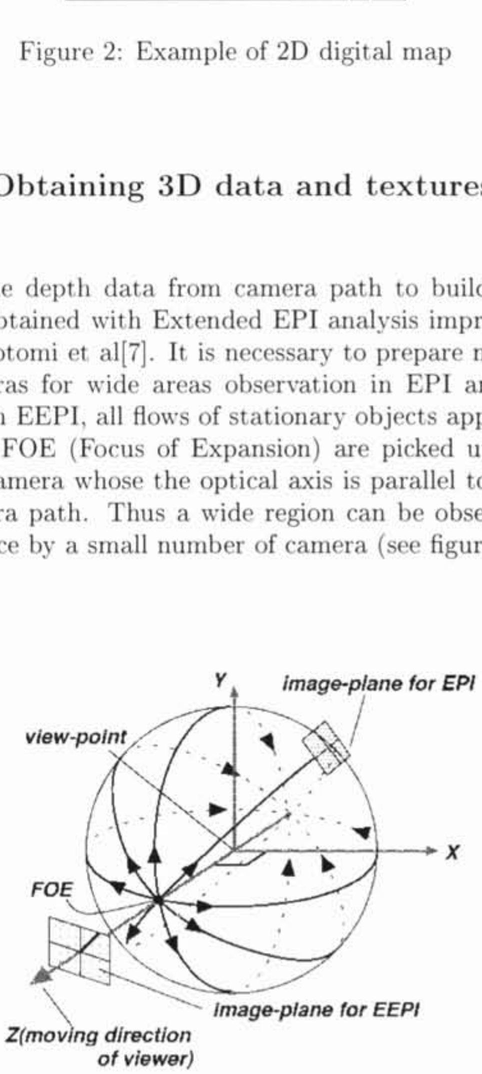


Figure 3: Flow of stationary object under straight moving

The textures of buildings for expression of urban scene are acquired from the image sequence. The slit spatiotemporal plane image utilized as texture data is made by integrating vertical lines of both sides of each frame of image sequence toward the temporal axis (see Figure 4).

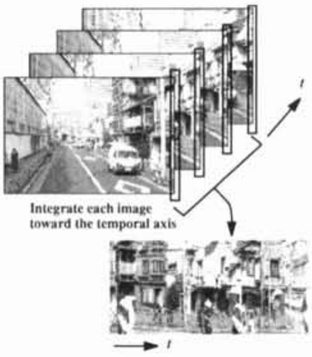


Figure 4: Slit image

3 Matching Measured 3D Data with Digital Map

In the real city environment, it is dificult to measure 3D data of buildings because of the existance of objects, such as a tick growth of trees and guardrails. Moreover, slight motion of each image due to the pitch, yaw and roll from the vehicle motion cannot be removed completely.

In the following section, we describe the matching method by utilizing boundary information between buildings in order to match 3D data with digital map.

3.1 Detecting boundary of buildings

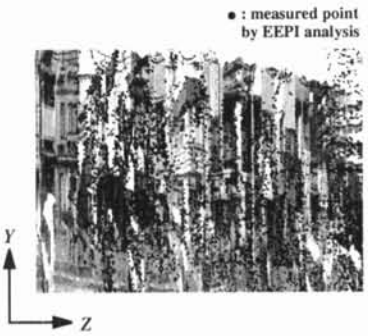
The 3D measured points appear on the parts of the vertical edge on the depth map as the white point shows in Figure 5. These parts are equivalent to steep texture alteration, such as boundaries between buildings, windows and doors.

Therefore, when the histogram of measured points in the direction of the camera path (Z-axis) is made, the peak of this histogram can be likely judged as the prospective boundary of buildings. At this time, we utilize the histogram of the measured points in the direction of the X and Z-axis with the measured points of buildings face the street (see Figure 6).

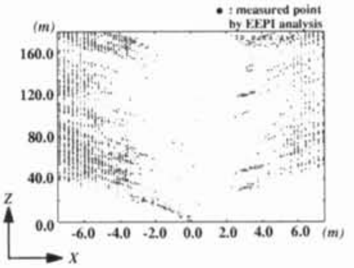
3.2 DP matching

In this paper, we apply DP matching method to match 3D measured data by EPI analysis with 2D digital map. The pattern of buildings boundary from 2D map and the prospective pattern of building boundary from the histogram of 3D measured data are utilized as feature vectors.

Matching urban scene with a 2D map utilized DP matching method has reported in [8]. In this research, the boundary pattern made from obtained



(a) Measured points on $Y - Z$ plane (Depth Map)



(b) Measured points on $X - Z$ plane

Figure 5: Measured points by EEPI analysis

panorama image of urban scene has matched with the building boundary pattern made from 2D digital map. This research utilizes the depth-data to buildings by EPI analysis and 2D map.

4 Experimental Result

On this experimant, we used the car equipped with the gyro sensor and distance sensor in order to record vibration and to obtain the moving distance. GPS is used to record the location of the car. Obtained image sequence is normalized using distance sensor data. Because of data collection on real environment, it is necessary to revise vibration for image sequence. As a result of measurement of gyro sensor, we had known that the vibration of a car running on a road is more influenced by a pitch than by a yaw and roll. Therefore, we shifted one of the two consecutive images up and down, then maximized the calculation of the correlation between two images.

First, we analyzed the EPI to make the slit-spatiotemporal image, and built the depth map. Figure 7(a) shows the histogram of 3D measured data by EEPI analysis on the $X - Z$ plane. We can judge that the peak is about $5 \sim 7$ meters from this figure. This part is supposed that the building surface faces the street. Then, in Figure 7(b), the peaks on the histogram are the prospective of building boundary. Figure 8 shows the result of slit-spatiotemporal image which had been divided each building. The lines in Figure 8 signified the building boundaries which have been matched the prospective of building boundary from histogram with the building boundary from 2D map by DP matching method. Next, we constructed the solid shapes of buildings with the digital map by utilizing data of GPS on the car (see Figure 9(a)). Figure 9(b) is the result of mapping texture on this shapes. Moreover, it was easy for modeling the plural paths to utilize the 2D digital map data. Meanwhile, Figure 9(c) shows the result of modeling the plural paths.

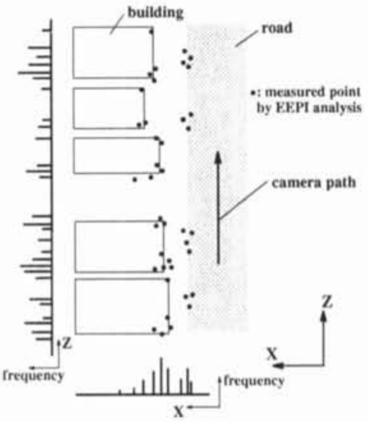


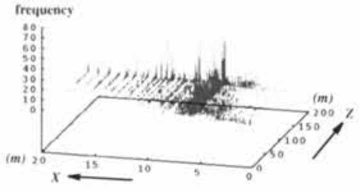
Figure 6: Histogram of 3D measured points

5 Conclusion

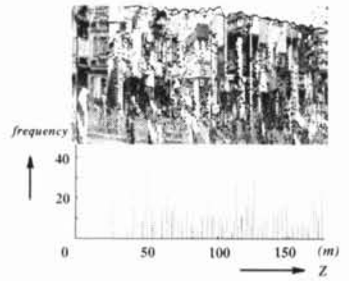
This paper described a technique to create the 3D model of the city from the motion images urban scenes by mobile observation and 2D digital map. We matched 2D map with these data to refine the obtained shapes and textures by analyzing moving picture. A photorealistic urban model was been developed as a result of using this technique. In this model, it is possible to walk through and locate a viewpoint freely. This model is expected the utilization of urban simulations and car navigation systems, and so on. For future works, we aim to create more realistic 3D city by adding elevation data to this model.

References

[1] M.Hirose, S.Watanabe and T.Endo: "Generation of Wide-Range Virtualspaces Using Pho-



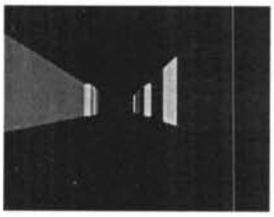
(a) Histogram on the $X-Z$ plane



(b) Depth-map and histogram

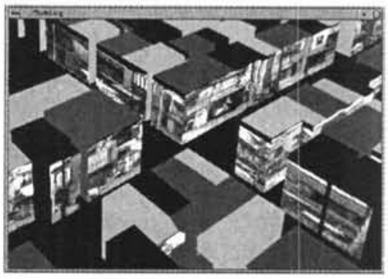
Figure 7: Histogram of 3D measured data

Figure 8: Result of building division



(a) Solid shapes of building

(b) Result of modeling



(c) Result of modeling in plural paths

Figure 9: Creating 3D city

tographics Images," Proc.VRIS'98, pp.234-241, 1998.

[2] M.Hirose, T.Tanikawa, T.Endo: "Building a Virtual World from the Real World," in (Y.Ohta and H.Tamura, eds.), Mixed Reality, Chapter 10, pp.183-197, 1999.

[3] M.Pollefeys, M.Vergauwen and L.V.Gool: "Automatic 3D Modeling from Image Sequences," Proc.ISPRS, Vol.XXXIII, PartB5, July, 2000.

[4] H.Zhao and R.Shibasaki: "Simulation on Automated Reconstruction of Urban 3D Object using Laser Range Finder (LRF) data," Journal of the Japan Society of Photogrammetry and Remoto Sensing, Vol.36, No.4, 1997.

[5] M.Yamamoto: "Determining Three-Dimensional Structure from Image Sequence given by Horizontal and Vertical Moving Camera," Trans.IEICE, Vol.J69-D, No.11, pp.1631-1638, Nov., 1986.

[6] R.Bolles, H.Baker and D.Marimont: "Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion," Int.J.Computer Vision, Vol.1, pp.7-55, 1987.

[7] S.Ozawa, M.Notomi and H.Zen: "A Wide Scope Modeling to Reconstruct Urban Scene," Proc.ISPRS Commision V, International Symposium on Real-Time Imaging and Dynamic Analysis, pp.370-376, 1998.

[8] H.Kawasaki, T.Yatabe, K.Ikeuchi and M.Sakauchi: "Interactive System of Real-World Video Based on Self-Organized Data Structure," Trans.IEICE, Vol.J82-D-II, No.10, pp.1561-1571, Oct., 1999(in Japanese).