*MVA '98* IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

# 13—24 3-D Objects Mapping Using Model and Image Based Approach

Tetsuo Shoji, Eiichi Iwanari and Yoshihiro Okada *
Ryukoku University

## Abstract

*In this paper, we propose 3-D objects mapping for 3-D model using model and image based approach. Our approach is the following: First, the model based approach, we generate a 3-D floor model from a floor plan by using given feature points, and acquire several disparity maps by using the model. Secondly, the image based approach, we acquire several disparity maps by using a stereo pair of panoramic images. Thirdly, we generate several 3-D models by combining virtual disparity map and real disparity map. Fourthly, we register the several 3-D models, and integrate the registered 3-D models into a single 3-D model. Finally, we extract 3-D objects from generated 3-D model at any positions, and we place them in the 3-D model at another position. Our method is useful for reducing the costs of processing for generating 3-D model and improving its quality.*

## 1 Introduction

The process of reconstructing 3-D models from 2-D images has long been a principal problem of the computer vision, and the process of rendering such recovered structures are as subjects which have recently increased interest in computer graphics. Recently, four particular areas of research have provided results that are applicable to the problem of modeling and rendering real scenes: Detemining Structure from Multiple Views[1], Stereo Correspondence[2][3][4], Modeling from Range Images[5], and Image-Based Rendering[6]. In particular, a great deal of effort has been made on generation of models from images. What seems to be lacking, however, is generation of convenient, accurate, and photorealistic 3-D models.

Therefore, we have applied an approach that combines model-based and image-based method.

Our approach consists of the following fave steps: First, the model based approach, we generate a 3-D floor model from a floor plan by using given feature points, and acquire several omnidirectional disparity maps by using the model at several positions on it, and we call these disparity maps *virtual disparity maps*.

Secondly, the image based approach, we acquire several omnidirectional disparity maps by using a stereo pair of panoramic images at the same positions as the model in the indoor scene, and we call these disparity maps *real disparity maps*.

Thirdly, we generate a 3-D model from the disparity map acquired by combining the virtual disparity map and the real disparity map acquired at each position.

Fourthly, we register the several 3-D models, and integrate the registered 3-D models into a single 3-D model.

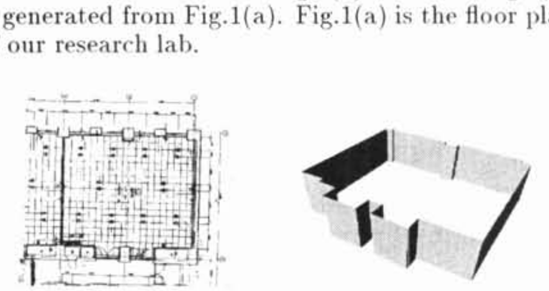
Finally, we extract 3-D objects from this 3-D model and place them at any position in the 3-D model using combinated disparity maps and virtual disparity maps. We call these process *3-D objects mapping*.

## 2 Model and Image Based Approach

### 2.1 Model Based Approach

#### 2.1.1 3-D Floor Model

A 3-D floor model is generated from a floor plan by using given feature points. The feature points are given corners of a floor in Fig.1(a) when the Fig.1(b) is generated from Fig.1(a). Fig.1(a) is the floor plan of our research lab.



(a) (b)

Figure 1: (a) Floor plan, (b) 3-D floor model

#### 2.1.2 Virtual Disparity Maps

In this section, we explain how to acquire virtual disparity maps from the 3-D floor model.

In case of a parallel stereo as shown in Fig.2, a coordinates of 3-D point can be calculated as:

*Address: 1-5 Seta, Otsu, Shiga 520-2194, Japan. E-mail: {shoji, eiwanari, okada}@mail433.elec.ryukoku.ac.jp

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \frac{ub}{d} \\ \frac{vb}{d} \\ \frac{fb}{d} \end{bmatrix}, d = u - u' \tag{1}$$

where $(X, Y, Z)$ is a camera coordinates system, $(u, v)$ is a coordinates of left image, $(u', v')$ is a coordinates of right image, $b$ is the baseline between the parallel stereo, $f$ is the focal length, and $d$ is the disparity.

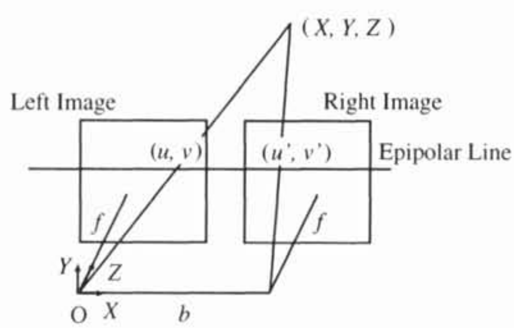



Figure 2: The parallel stereo model

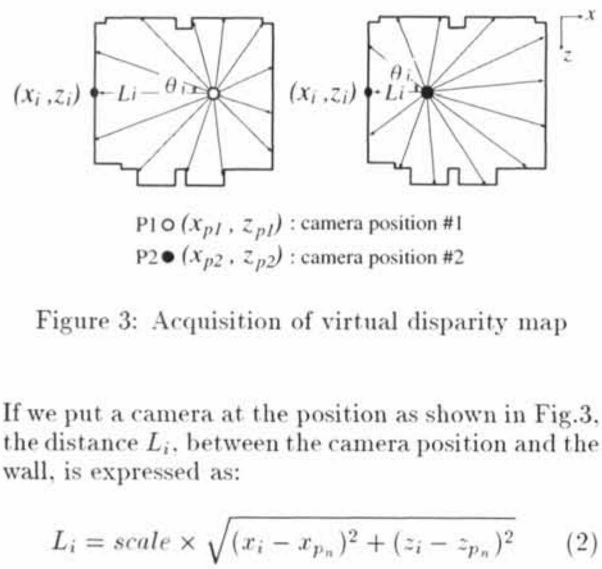



Figure 3: Acquisition of virtual disparity map

If we put a camera at the position as shown in Fig.3, the distance $L_i$, between the camera position and the wall, is expressed as:

$$L_i = scale \times \sqrt{(x_i - x_{p_n})^2 + (z_i - z_{p_n})^2} \tag{2}$$

where $scale$ is the actual length per a pixel, $(x_i, z_i)$ is a coordinates of the wall in the 3-D floor model, $(x_{p_n}, z_{p_n})$ is a coordinates of the camera position in the 3-D floor model.

A coordinates of the wall is given as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} L_i sin\theta_i \\ y_j \\ L_i cos\theta_i \end{bmatrix} \tag{3}$$

where $\theta_i$ (for $i = 0,...,2\pi$) is the direction of view point, and $y_j$ (for $j = 0,...,$height) is the height of the wall.

Therefore, virtual disparity maps are calculated as:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{d}{b} L_i sin\theta_i \\ \frac{d}{b} y_j \end{bmatrix}, d = \frac{fb}{L_i cos\theta_i} \tag{4}$$

Fig.4 shows the omnidirectional virtual disparity map of Fig.1(b).

Figure 4: The omnidirectional virtual disparity map at the camera position #1

## 2.2 Image Based Approach

### 2.2.1 Generation of Panoramic Images

A omnidirectional panoramic image is generated from mutiple rotated camera images, as shown in Fig.5. these images are captured while panning camera around 360 degree. We use the stereo camera is mounted on the tripod. Fig.6 shows the stereo pair of partial panoramic images at the camera position #1 of Fig.3 in our research lab.

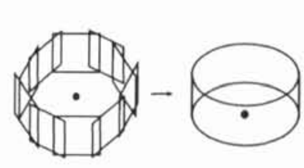


Figure 5: Generation of omnidirectional panoramic images


Figure 6: The stereo pair of panoramic images at the camera position #1(partial image):(a) Left panoramic image, (b) Right panoramic image

### 2.2.2 Real Disparity Maps

In order to acquire real disparity maps from the stereo pair of omnidirectional panoramic images, we use Pixel-to-Pixel Stereo[3]. This algorithm is used to detect depth discontinuities from a stereo pair of images. The algorithm matches individual pixels in corresponding scanline pairs while allowing occluded pixel to remain unmatched, then propagates the information between scanlines by means of a fast postprocessor. This handles large untextured region, uses a measure of pixel dissimilarity that is insensitive to image sampling, and prunes bad search nodes to increase the speed of dynamic programming.

Fig.7 shows the real disparity map acquired from the stereo pair of partial panoramic images at the camera position #1 of Fig.3.

Figure 7: The disparity map acquired by pixel-to-pixel stereo algorithm at the camera position #1

### 2.3 Combination of Disparity maps

We combine real disparity maps and virtual disparity maps for correcting real disparity maps. To put it concretely, disparity maps for width × height pixel resolution are calculated, for $u$=0,...,width-1 and $v$=0,...,height-1, as:

$$c\_d(u,v) = \begin{cases} v\_d(u,v) & r\_d(u,v) < v\_d(u,v) \\ r\_d(u,v) & \text{otherwise} \end{cases}$$

where, at a coordinates of $(u,v)$, $r\_d(u,v)$ expresses the value of the disparity in the real disparity map, $v\_d(u,v)$ expresses the value of the disparity in the virtual disparity map, $c\_d(u,v)$ expresses the value of the disparity in the combined disparity map.

Fig.8 shows the combined disparity map at the camera position #1 of Fig.3.

Figure 8: The combined disparity map at the camera position #1

In the same way, we can acquired the combined disparity map at the camera position #2 of Fig.3.

## 3 Registration and Integration

### 3.1 Registration

Two 3-D models are acquired from the combined disparity maps by applying (1) and (4) at the camera position #1 and #2. Registration is the process that two 3-D models are brought into alignment. The purpose of registering the individual 3-D models is to place them in the floor plan. In our work, the process of registration is over by putting two 3-D models together, because the camera position is known on the floor plan. In our research lab, Fig.9(a) shows the 3-D model at the camera position #1 and Fig.9(b) shows the 3-D model at the camera position #2. Fig.9(c) shows the result of registering 3-D models acquired at the camera position #1 and #2.

### 3.2 Integration

Integration is the process that registered 3-D models are combined into a single 3-D model. To put it concretely, our integration is the process that two 3-D models are combined into a single 3-D model by average between two 3-D models. Fig.9(d) shows the integration of registered 3-D models.

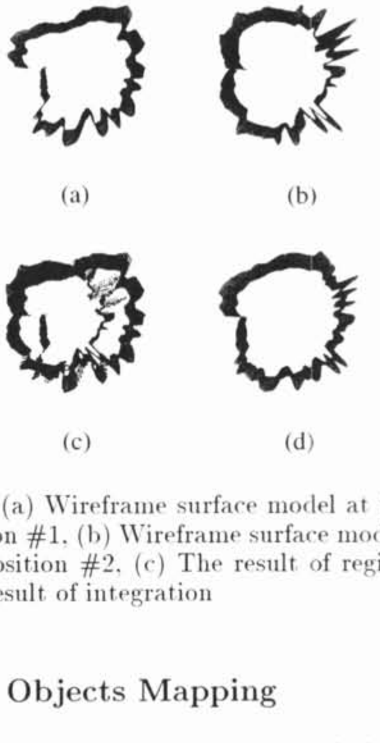



Figure 9: (a) Wireframe surface model at the camera position #1, (b) Wireframe surface model at the camera position #2, (c) The result of registration, (d) The result of integration

## 4 3-D Objects Mapping

Using combined disparity maps and virtual disparity maps, we extract 3-D objects from the 3-D model and place them at any position in the 3-D model. To put it concretely, as shown in Fig.10, the process of 3-D objects mapping is the following steps:

**Step1** : As shown in Fig.10(a), we extract the source image containing objects which are extracted from the omnidrectional panoramaic images.

**Step2** : As shown in Fig.10(b), we extract the disparity map containing objects which are extracted from the omnidrectional combined disparity map.

**Step3** : As shown in Fig.10(c), we extract the disparity map containing objects which are extracted from the omnidirectional virtual disparity map.

**Step4** : As shown in Fig.10(d), we acquire the difference disparity map between Fig.10(b) and Fig.10(c).

**Step5** : As show in Fig.10(e), we acquire the 3-D objects model from the difference disparity map by applying (1) and (4). 3-D objects model is made texture mapping using Fig.10(a).

**Step6** : We place 3-D objects model at any position in the 3-D model.

## 5 Experimental Results

We generated 360 degree panoramic image by combining 48 images which were taken at each 7.5 degree. We acquired 2 stereo pairs of omnidirectional panoramic images at 2 positions. Each panoramaic image is 8137 × 240 pixel resolution. Fig.11 shows the result of generating textured 3-D model. Fig.12 shows the result of 3-D objects mapping in the 3-D model.

## 6 Summary

We have described about 3-D objects mapping using model and image approach. Using the floor plan, generated 3-D model from indoor scenes was more convenient, more accurate, and more photorealistic than the methods currently available.

## References


[1] Camillo J. Taylor and David J. Kriegman: "Structure and motion from line segments in multiple images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(11), November 1995

[2] Sing Bing Kang, Richard Szeliski: "3-D Scene Data Recovery Using Omnidirectional Multibaseline Stereo", *International Journal of Computer Vision 25(2)*,167-183(1997)

[3] Stan Birchfield, Carlo Tomasi: "Depth Discontinuities by Pixel-to-Pixel Stereo", *Processings of the 1998 IEEE International Conference on Computer Vision* Bombay, India

[4] Masatoshi Okutomi, Takeo Kanade: "A Multiple-Baseline Stereo", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(4), pp.353-363 1993

[5] Y.Chen and G.Medioni: "Object modeling from multiple range images", *Image and Vision Computing*, 10(3), 145-155, April 1992

[6] Leonard McMillan and Gary Bishop: "Plenoptic modeling: An image-based rendering system", *SIGGRAPH '95*, 39-46, 1995

[7] Paul J. Bsel, Neil D. Mckey: "A Method for Registration of 3-D Shapes", *PAMI*,239-255(1992)

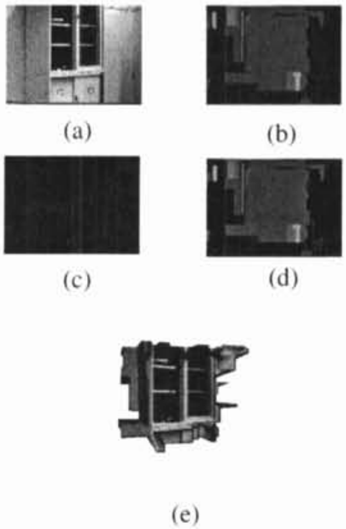



Figure 10: (a) Source image, (b) Combined disparity map, (c) Virtural disparity map, (d) Difference disparity map, (e) Extracted 3-D objects model

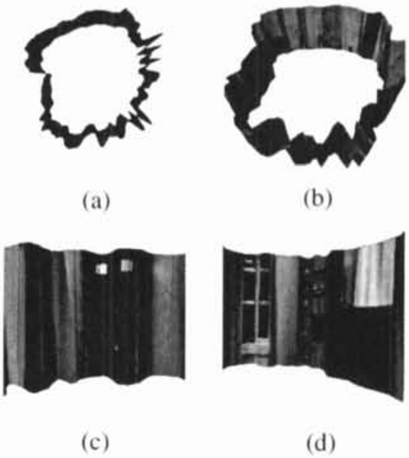


Figure 11: The result of generating textured 3-D model: (a) Top view, (b) Oblique view, (c) Close up view1 , (d) Close up view2

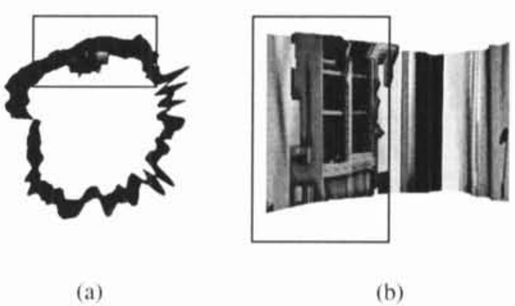


Figure 12: The result of 3-D objects mapping: (a) The 3-D objects mapping (top view), (b) Close up the model