Synthesizing the Arbitrary Views of a Large Scale Object Using a Hand-Held Camera

Myint Myint Sein† and Michihiko Minoh†‡

†Keihanna Human Info-Communication Research Center, CRL, 3-5 Hikaridai, Seikacho, Sorakugen, Kyoto619-0289, Japan
‡Kyoto University, Yoshida-Nihonmatsucho, Sakyoku, Kyoto606-8501, Japan

Abstract

This paper proposes a new approach for synthesizing the arbitrary views of a large-scale object. The images are taken with a hand-held camera. The epipolar geometry approach is basically applied for determining the corresponding pairs among the stereo image pairs. The exact depth measurement of each pixel of an object has been determined from the multiple image pairs. The synthesized images observed from new viewing points are generated by the disparities among the different viewpoint images. A few experiments with images of a real scene have been done to confirm the effectiveness of the proposed method.

Keywords: Arbitrary view, synthesized image, depth measurement, disparity image, panoramic image

1. Introduction

The research of the synthesizing the virtual view has been paid substantial attention by the researchers. Image mosaicing technique is very useful for constructing a panoramic image from the sampled images of a scene. This paper presents the recent results of our research work related to construct the arbitrary views of an object from the different viewpoint images of a scene. Especially, a full view image of the large- scale object is reconstructed. The images are grabbed by a hand held camera from different camera positions. The corresponding pairs among the stereo image pairs are determined from the relation of the camera positions. The perspective projected geometry method is applied for detecting the relation between the camera and image coordinate system from the image pairs. The exact depth measurement of each pixel of an object has been determined from the multiple image pairs. The panoramic image has been generated by disparities among the different viewpoint images.

Some reports have already been presented concerning these researches to synthesize the arbitrary view for large-scale virtual environment[1-6]. T. Takahashi et al.[1] proposed a method for rendering views for large-scale scenes. An omni directional camera is used to capture the panoramic image running along a straight line. N. Chiba[2] proposed the feature based image mosaicing technique for arbitrary depth scene. The dense depth measurement is important to reconstruct the 3D stereoscopic view of the object related to the desire viewpoint. CMU has been developed a video-rate stereo machine to obtain these depth map of a scene. In their systems, the distance map has been produced from the multiple images grabbed by the fixed cameras configuration. J. Mulligan and K Daniilidis[5] explored the computation of dense trinocular disparity maps for non-planar camera configurations that arise when cameras are set surrounding the object, which is to be modeled. In many approach, the images are acquired by a fixed camera system. Furthermore, the images of an object and a scene are nearly flat configuration because of the cameras have been set sufficiently far from the object. Unlike the other virtual view synthesizing system, we acquire the images not only far from the object but also near from the object by using a hand-held camera. In this paper, we consider the case of reconstructing the full view of a large-scale object. To obtain a full view image of the large-scale object, it needs the sufficient distance between the object and a camera. Due to some preventions such as some obstacles and narrow or small experimental room, it is not impossible to acquire the far image or full view image directly. For this problem, we consider that a way of reconstructing technique by synthesizing the closed view images. Figure 1 illustrates the generating the arbitrary views from the closed view images taken by a hand-held camera.

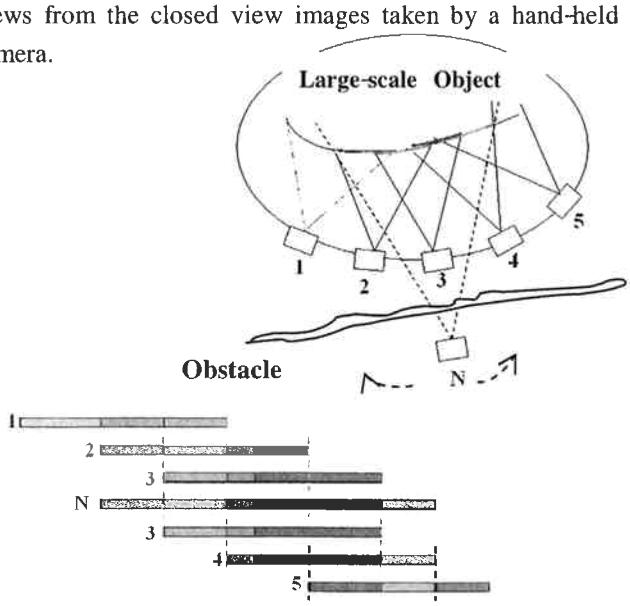


Fig 1. Illustration of the arbitrary views generation system

This approach can be applied directly in the field of geometric modeling in the virtual environment. It is also possible to generate the 3D virtual views and a new scene of the object by using the depth measurement of the object.

2. Projective transformation

In this Section, we will consider the principal of the projective transformation between two different viewpoints V_1 and V_2. The transformation matrix **H** can be defined by the image coordinates m_1 and m_2 as:

$$k\, m_2 = \mathbf{H}\, m_1 \quad , \qquad (1)$$

where k is an arbitrary constant and $m_1=(x_1,y_1)$ and $m_2=(x_2,y_2)$ are the projected image points of the point M on the object, respectively. Eq. (1) can be rewritten as

$$m_2 = k\ \mathbf{H}\ m_1 \quad , \qquad (2)$$

where k is an arbitrary nonzero scale factor. The linear solution of the transformation matrix **H** can be determined by using the four or more corresponding pairs of feature points.

3 An Algorithm for reconstructing the virtual view

An algorithm is developed for generating the virtual view of a scene from the stereo image pairs grabbed by the multiple camera system. The model of an object in new view has been generated based on the registration approach. The general process for generating system is as follows:

[Step 1] Grabbed images of an object and a scene by a hand-held camera from the different camera positions.

[Step 2] Preprocessing, Detecting the correspondence points pairs of object in multiple images.

[Step 3] Computing the disparities measurements among them.

[Step 4] Averaging the overlapping region in both images.

[Step 5] Reconstructing the arbitrary view point image and virtual views of a scene.

4. Experimental result and Conclusions

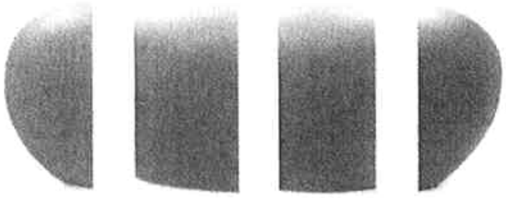
A few experiments have been done for generating the virtual view reconstruction of an object and a scene. To analyze the data, the high performance computation and visualization software visual c++ version 6 is used. Figure 2 (a), (b) and (c) show the images of the restaurant building acquired from the three different camera positions by a hand-held camera. The arbitrary viewpoint of a scene shows in figure 2(d). Figure 3 illustrates the four original images of a vase and its reconstructed panoramic image.

The recent result of our research work related on the virtual view generation of an object and a scene has been presented. Especially, a full view image of the large-scale object is reconstructed form the closed view images taken by a hand-held camera. The effectiveness of our method is confirmed through the experimental results. In the next step, we should take to consideration the reconstruction of a full view of the large object included a few textures.

(a) 1st frame (b) 2nd frame (c) 3rd frame

(d) Reconstructed image of the whole scene

Fig. 2 Four real images and panoramic view



(a) (b) (c) (d)

(e) Reconstructed panoramic image

Fig. 3 Images of a vase and its reconstructed panoramic image

Reference

[1] T. Takahashi et al., "Arbitrary View Position and Direction Rendering for Large-Scale Scenes", IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1063-6919, vol. 2, 2000.

[2] N. Chiba," Image mosaicing for arbitrary depth", Image Labo, 11(8)220-230, August 2000.

[3] N. Chiba et al., "Accurate and First Panoramic Image Mosaicing Software Based on Image Features", Sanyo Technical Review, Vol. 35, No. 1 pp. 75-82, June 2003.

[4] M. Okutomi and T. kanade, "A multi-baseline stereo" IEEE Trans. On PAMI, pp. 353-363, Vol. 15, No. 4, April 1993.

[5] Y. Nishiyama et al., " Estimation of 3D Model and Object Color Using Multiple Color Images", IEICE-DII, Vol. J82, No. 6, pp. 990-1000, 1996.

[6] T. Yamamoto et al., "Correspondence Search of Trinocular Stereo Using Dynamic Programming", IPSJ SIGNotes Computer Vision, Vol. 046, 1986.