Panoramic Stereo Imaging with Complete-Focus Views for Virtual Reality*

Ho-Chao Huang, Yi-Ping Hung, and Sheng-Wen Shih
Institute of Information Science, Academia Sinica, Taipei

Abstract

An image-based virtual reality (VR) system, the panoramic stereo imaging (PSI) system, is proposed in this paper. Traditional image-based VR systems can not to produce stereo views and dos not give the viewer the feel of realistic 3D depth. The proposed PSI system uses a stereo camera set to capture the stereo image pairs simultaneously, and produces the stereo panoramic images by using the camera parameters obtained from camera calibration. Furthermore, our PSI system can construct the clear stereo views from the complete-focus images. With both the stereo and complete-focus features, our PSI system provides more realistic and clearer views images than other image-based VR systems do.

1 Introduction

Virtual reality has become one of the hottest commercial and research topics[1, 2]. The virtual reality (VR) world building is the major time-consuming work for establishing the virtual reality environment. The image-based approach (eg., Quick-Time VR[3, 4, 5]) and the model-based approach (eg., 3D Studio) are commonly used for building the VR world. The model-based approach constructs the 3-D models of the real world objects first and then reconstructs the VR images by rendering the 3-D models. The model-based approach produces good picture quality but the 3-D models building is a tedious work. On the other hand, the image-based approach builds the panoramic VR views by seaming the photographs of the real world directly, and enjoys the advantage of easy and fast VR world building.

However, the image-based VR world building approach has some problems described as follows.

1. Hard to produce stereo images —

   In the image-based approach, no 3-D models are established, and hence it's hard to produce stereo views for human stereo perception.

2. Hard to focus everywhere —

   Most of the image-based approach perform best for scenes containing no near object. When establishing the VR world containing nearer objects, the out-of-focus problem occurs.

*The conference version of this paper is at *Proceeding of RAMS'96, The Second Workshop on Real-Time and Media Systems, July, 1996, Taipei, Taiwan, ROC.*, pp. 405–410.

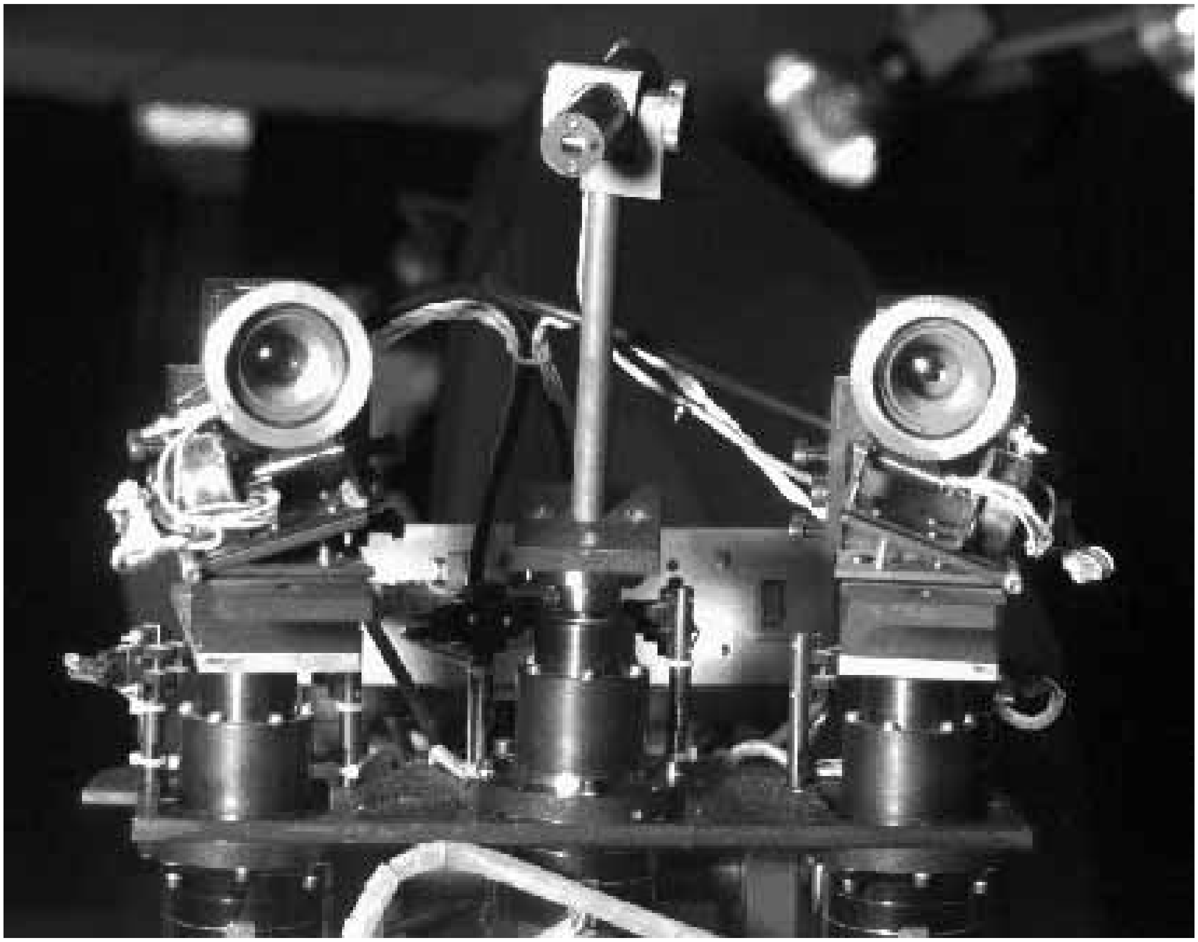


Figure 1: The IIS head.

3. Hard to travel within the virtual reality world arbitrarily.

4. The object occlusion problem.

This paper presents a Panoramic Stereo Imaging (PSI) system which takes the images from a stereo camera set, eg., the IIS head[6, 7, 8, 9, 10, 11], and provides clear stereo views of the VR world. While the IIS head is not a prerequisite to our PSI approach, it does serve as an excellent experimental tool in our initial experiments for system design and performance assessment. Figure 1 is the photograph of the IIS head. Just as that done by most image-based VR systems, our PSI system seams the surround views automatically to build the VR world. The unique feature of our PSI system is that it can provides complete-focus stereo views without building 3D models, by using an auto-focusing technique to select the correctly focused image for each pixel.

2 Preliminary Review

This section briefly reviews the panoramic imaging system. There are two commonly used method for creating such panoramic images. The first one is to use a panoramic camera. The panoramic camera has the capability to capture the 360 degrees panoramic image with one shot. The advantages of using the panoramic camera are that it is easy to capture the panoramic picture and the panoramic imaging system can easily create the panoramic view by using the panoramic picture. However, the panoramic camera is quite expensive and is not popular for virtual reality world builders.

The second way for creating panoramic images is to use a general-purpose off-the-shelf camera. Most panoramic imaging systems, such as the famous QuickTime VR system, create the panoramic view by capturing a series of overlapped shots using a 35mm film camera and a customerized tripod rig. Alternatively, video or digital cameras can be used. The panoramic imaging system sticks the overlapped shots on a cylinder or a sphere and generates the 360 degrees panoramic images accordingly. The advantage of the second approach is that the general-purpose camera is cheap and popular, but the difficulties of this approach are that the panoramic imaging system should have the capabilities to warp the input images and to stick the images on the proper position of the cylinder or the sphere. Furthermore, the virtual world builder need to make sure the the rotation center, when capturing the overlapped shots, is located at the lens center of the camera.

3 Panoramic Stero Imaging (PSI) System

This section presents the detail of the proposed PSI system. In order to generate the stereo views, two cameras are used for simultaneous image acquisition. Figure 2 shows the block diagram of the PSI system.

3.1 Stereo Image Acquisition

The first step of image-based VR world building is to capture the panoramic images. In our initial experiments, images are taken with the two cameras on the IIS head when the IIS head rotates ten degrees step by step, and the rotating center is the same position as the center of the left camera. For each step, seven image pairs with different focus settings are captured, which are indexed from 0 to 6, and the image pair with index 6 has the nearest focusing distance. After the stereo images acquisition process, a sequence of images $IM_{c,f,r}$ are obtained, where the possible values of c (camera selector) are $left$ and $right$, the range of f (the index of the focal distances) is from 0 to 6 and the angles of r (rotation) is from 0 to 180 degrees, stepped by 10 degrees. In our experiment, the rotation angle is limited to the range of 0 to 180 degrees due to the panning capability of the IIS head.

3.2 Generating of Complete-Focus Images

For each camera and each orientation, seven images with different focal distances are captured during the image acquisition process. Our PSI system generates the complete-focus images by selecting the correctly focused image for each pixel. This process contains two stages, the

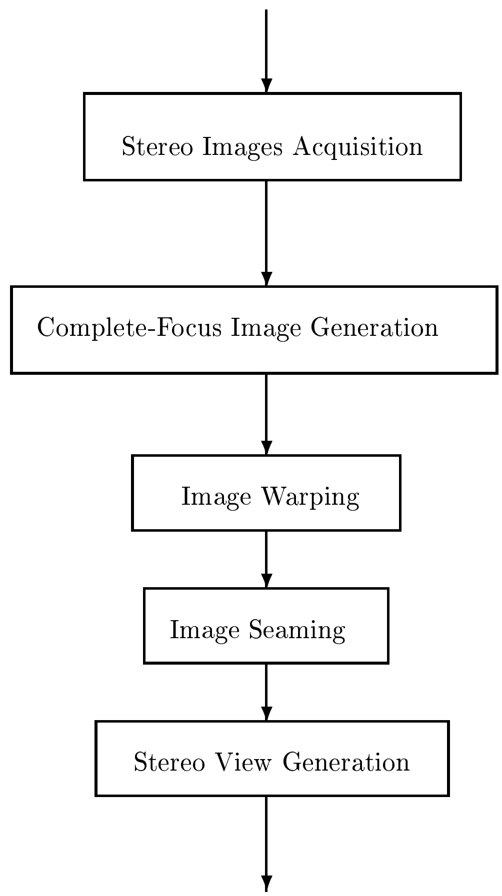


Figure 2: The block diagram of the panoramic stereo imaging system

gradient calculation and the medium filtering. The gradient calculation stage uses equation (1) to calculate the gradients of each pixel.

$$G_{(c,f,r)}(x,y) = \sum_{n=-3}^{n=2} |IM_{(c,f,r)}(x+n+1,y) - IM_{(c,f,r)}(x+n,y)| + |IM_{(c,f,r)}(x,y+n+1) - IM_{(c,f,r)}(x,y+n)|, \quad (1)$$

where the (x, y) is the x and y coordinates of the pixel in the image.

For each pixel, the index of the focal distance is selected if its corresponding $G_{(c,f,r)}(x, y)$ is maximum among all focal distances. The selection function is defined in (2).

$$F_{(c,r)}(x,y) = \max_{f} G_{(c,f,r)}(x,y) \quad (2)$$

Due to the blurring effect caused by out-of-focus, the result of (2) is noisy especially on the boundaries of the objects with different distances and with different intensity. Therefore, a medium filter defined in (3) is used for removing the noises on object boundaries.

$$F'_{(c,r)}(x,y) = medium_{5 \leq i \leq 5, 5 \leq j \leq 5} F_{(c,r)}(x+i, y+j) \quad (3)$$

Finally, a complete-focus image is generated by using equation (4).

$$IF_{(c,f,r)}(x,y) = IM_{(c,F'_{(c,r)}(x,y),r)}(x,y) \quad (4)$$

3.3 Image Warping

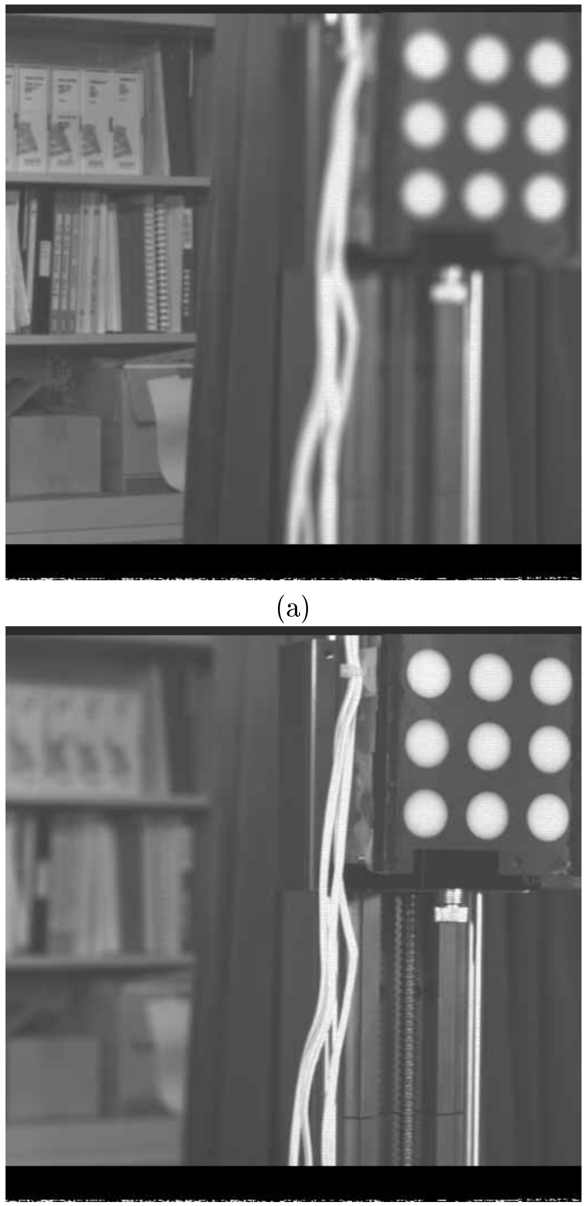
Since the panoramic image is mapped to a cylinder, the complete-focus image IF should be warped to be fitted on the cylinder. Furthermore, due to the optical lens of the cameras are not perfect, the distortion due to the optical lens also need to be compensated during the warping process. Since the IIS head used in the PSI system has been calibrated, and the intrinsic and extrinsic parameters of the camera are available, the image warping of the PSI system is easy and accurate.

3.4 Image seaming

Next, all warped images should be seamed to generate a panoramic image. The relative orientations of each images are available in the camera parameters, and our PSI system uses the orientation information to calculate the offsets of the adjacent images. Some pixels in the panoramic image are only covered by one warped image, and some pixels in the panoramic image are covered by two adjacent warped images. We defined the pixels covered by two warped images as the transition pixels. The value of transition pixels is the linear combination of the values of the corresponding pixels in both covering images.

3.5 Generation of Stereo Views

Both the right-eye and left-eye panoramic images are constructed by using the previous procedures. Users can see stereo views produced by the PSI system through a pair of stereo



(a)

(b)

Figure 3: Two images with different focus settings.

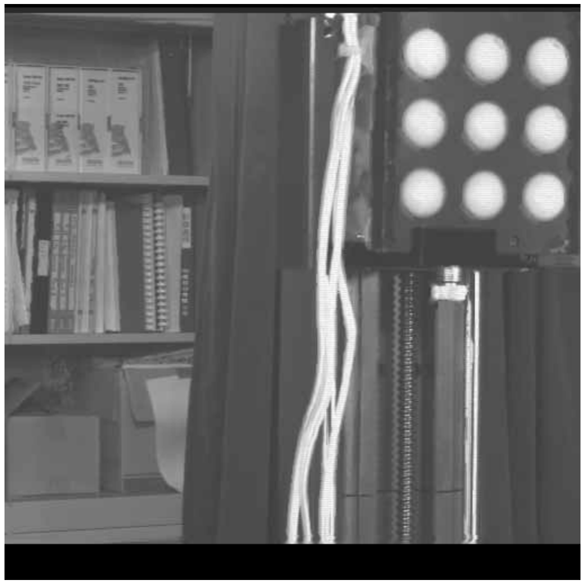


Figure 4: A complete-focus image.

glasses, and control the viewing position and orientation through a 6-D mouse. The turn-left, turn-right, turn-up, turn-down, zoom-in and zoom-out functions are all available in our PSI system with a natural and user-friendly way.

The distance between the left and right cameras determines the disparity between two eyes of the stereo movie. However, the distance between two cameras may not be the same as the distance of human left and right eyes, and due to the zoom-in function of the PSI system, users may feel uncomfortable when gazing at the near objects if the disparity of the PSI system is fixed. Fortunately, when the PSI system making the auto-focused images, the roughly distance of each pixels can be obtained. Thus, the proposed PSI system has the capability to automatically determine the comfortable disparity of the stereo views when the user is looking at any object.

4 Experimental Results

Several pictures are shown in this section to illustrate the process and the results of our PSI system. Figure 3 shows two of the original images captured by the camera of the IIS head with different focal distances. Since the distances of the objects in the images are different, none of the images contain clear image of every object. For example, in Fig. 3(a), the bookshelf is clean but the nine-dot calibration plate is blurred. On the other hand, the bookshelf in Fig 3(b) is blurred and the calibration plate is clearly imaged. Figure 4 shows an complete-focus image generated by our system. All objects in the complete-focus image are clear and sharp. The result of the seaming process is shown in figure 5. Figure 6 shows a user touring the VR world with our PSI system by using a stereo glasses and a 6-D mouse.

Another experiment on creating outdoor panoramic stereo images is described below. Figure 7 shows three shots of the images for the left eye. The warped images are shown in figure 8

(a) A complete-focus image viewed by the left eye.

(b) A complete-focus image viewed by the right eye.

Figure 5: A complete-focus stereo view seen on the screen by cutting off from the panoramic stereo image pair.

Figure 6: The photograph of a user using the PSI system.

and the seamed image is shown in figure 9. Finally, the whole panoramic stereo images are shown in figure 10.

5 Conclusions

A panoramic stereo imaging (PSI) system is presented in this paper. The PSI system generates stereo panoramic stereo view from a series of images. Clear images (or complete-focus views) are constructed by using the gradient maximization and the medium filter. Due to the stereo and the complete-focus features of our PSI system, the image-based virtual reality system becomes more attractive than before.

In the current implementation of our PSI system, the rotation center for capturing the panoramic images is at the same position as the lens center of the left camera. Because the rotation center is not at the same position as the lens center of the right camera, the distortion of the images for right eye due to the disparity occurs. This distortion can be alleviated by rotating with respect to the middle point between the two cameras, and by appropriate image warping.

Moving objects can cause problems for image-based virtual reality system. We are currently working on solving this problem by using video, or more precisely, multiple images with same viewpoint. Hopping discontinuity is another open problem for the image-based virtual reality systems. In the future, we intend to solve this problem by generating 3D model using the stereo image pair.

References

[1] J. M. Moshell. Three views of virtual reality: Virtual environments in the U. S. military. *IEEE Computer*, 26(2):81–82, February 1993.

[2] G. B. Newby. Virtual reality. *Annual Review of Information Science and Technology*, 28, 1993. Medford, N,J,: Learned Information.

[3] Apple Computer. QuickTime VR. software package, 1995.

[4] S. E. Chen. QuickTime VR — an image-based approach to virtual environment navigation. *SIG GRAPH '95*, pages 29–38, 1995.

[5] W.K. Tsao and M. Ouhyoung. An alternative approach of rendering high quality images for virtual environments using scanned images. *HDTV '95*, pages 7B–1–7B–8, 1995.

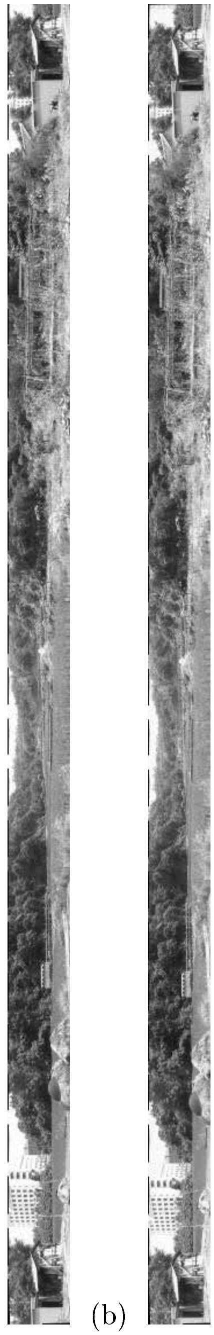
[6] S. W. Shih, J. S. Jin, K. H. Wei, and Y. P. Hung. Kinematic calibration of a binocular head using stereo vision with the complete and parametrically continuous model. In *SPIE Intelligent Robots and Computer Vision XI*, volume 1825, pages 643–657, 1992.

[7] S. W. Shih, Y. P. Hung, and W. S. Lin. Kinematic parameter identification of a binocular head using stereo measurements of single calibration point. In *IEEE International Conference on Robotics and Automation*, pages 1796–1801, Nagoya, Japan, 1995.

Figure 7: Three original images of the panoramic images for left eye.

Figure 8: Warped images of the panoramic images for left eye.

Figure 9: Seamed images of the three panoramic images for left eye.



(a) (b)

Figure 10: Seamed images of the panoramic stereo images for left (a) and right (b) eyes.

[8] S. W. Shih, Y. P. Hung, and W. S. Lin. Comments on 'a linear solution to kinematic parameter identification of robot manipulator' and some modifications. *RA*, 11(5):777–780, 1995.

[9] S. W. Shih, Y. P. Hung, and W. S. Lin. Accuracy analysis on the estimation of camera parameters for active vision systems. In *IEEE International Conference on Pattern Recognition*, Vienna, Austria, August 1996.

[10] S. W. Shih, Y. P. Hung, and W. S. Lin. Error analysis on closed-form solutions for kinematic calibration. Technical Report TR-IIS-96-xxx001 (in preparation, ftp://140.109.23.242/pub/TechRpt/TR-IIS-96-xxx001.ps.gz), Institute of Information Science, Acdemia Sinica, Nankang, Taipei, Taiwan, 1996.

[11] S. W. Shih, Y. P. Hung, and W. S. Lin. Calibration of an active binocular head. Technical Report TR-IIS-96-007 (ftp://140.109.23.242/pub/TechRpt/TR-IIS-96-007.ps.gz), Institute of Information Science, Acdemia Sinica, Nankang, Taipei, Taiwan, 1996.