

A VIRTUAL WORLD CONSTRUCTION METHOD USING CAPTURED IMAGES - APPLICATIONS FOR VIRTUAL SHOPPING

Nobuyuki Ozaki and Kaoru Takada

Toshiba Corporation

Tokyo, Japan

ABSTRACT

Procedures for interactively employing captured images for virtual world construction are described using prototypes of two applications for virtual shopping. Virtual reality technology, especially the image based rendering technique, enables us to exhibit products in a sophisticated way. With a concert ticket reservation system the customer can check the view of the stage from any seat. These images are synthesized from captured images and are not created from less realistic computer graphics. With a catalog shopping system that uses captured images to show products, the customer can observe them stereoscopically and vary the viewing angle interactively. These systems are implemented into kiosk-type terminals. Purchases can be made with IC-card type electronic cash.

1. Introduction

With the growing popularity of electronic commerce, many virtual malls have appeared. These virtual malls describe a variety of merchandise through the internet. The user can check what is being sold as well as make purchases. The navigation is mainly done by navigating through computer graphics, which are mainly written in VRML or Java [1]. If the shopper is interested in a specific product, he or she can usually get additional information, including an images or even a video of the product. The views of the products aids can be greatly enhanced, however, by immersion and interaction with the images.

We developed suitable ways of showing the products for two purchasing processes: concert ticket reservation and catalog shopping. We designed the concert ticket reservation system so the customer can book any seat and confirm how the stage looks from that location. The view of the stage from the seat is interpolated from the view from a different seat which has been captured. The traditional way to get interactivity with the captured images is to obtain a depth map, equivalent to a 3D model, from images. An image from a different viewpoint can be synthesized from this depth map [2][3]. The preparation, however, requires a camera device and a sophisticated calibration technique, which might not be generally available.

Considering the size of the target object, a stage, and the range of user's movement, this view can be modelled as a walkthrough around one block in a town. For this outdoor size, Image Based Rendering, with no camera calibration procedure, is a promising approach. Hirose's approach [4] for generating wide-range virtual spaces proves to be effective. However, based on view morphing [5], which assumes that the depth is constant, this is not suitable for our case because the depths of the stage can not be considered as constant. Our approach of Image Based Rendering uses a surveying instrument.

With a catalog shopping system, the customer can observe products stereoscopically from any angle. Our approach is one example of the extension of Apple's QTVR to stereo mode. These systems also provide money transaction procedures using IC card-type digital cash. Section 2 focuses on practical procedures for capturing images and interpolating them for a ticket reservation system. Section 3 addresses procedures for capturing images in stereo and displaying them stereoscopically.

2. Concert Ticket Reservation System

2.1 SYSTEM CONFIGURATION

Figure 1 shows the image capture side and image display side configurations. On the image capture side a digital still camera is used to capture the stage view from various seats. These images are stored in a PC with some additional information, such as the seat location. A surveying instrument is used to survey characteristic points on the stage in three dimensions and the positions of the camera. This is explained in detail later. On the image display side the system uses a touch monitor to sense which seat is indicated for reservation. While displaying the stage view, a musical note is also played in stereo with a sound localizer to roughly locate the sound direction.

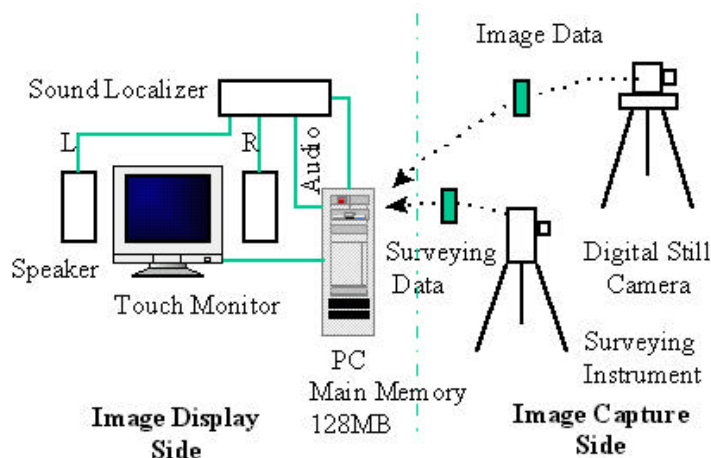


Figure 1: System Configuration

2.2 CONCEPT FOR IMAGE INTERPOLATION

A common method for synthesizing an image from a novel position using two captured images with points' correspondence is calculation based on the fundamental matrix, which can extract from the constraint of epipolar geometry [6][7]. This method requires camera calibration in advance. Since this approach seemed to have some difficulties due to the wide angle camera calibration, noise sensitivity of the fundamental matrix and a camera setup, we developed a very simple method. We measured directly characteristic points on the stage and camera positions with a surveying instrument. Once the system has the stage model with these three-dimensional positions, the system, by simply calculating perspective transformation, can project the model onto the camera image plane assumed to be a view from a seat. After projecting the points, the system maps the corresponding textures from an appropriate captured image. A photo of a female model, whom we assumed to be a singer, is also placed on the stage to help the observer grasp the size of the stage.

2.3 PRACTICAL PHOTOGRAPHING PROCEDURES

2.3.1 Hall Capacity and Stage Model

The hall capacity is approximately 1500 seats and its rough size is as shown in Table 1. The stage is modelled as shown in Figure 2 with four planes: right wall, left wall, back wall, and floor.

Clues for area	Distance
From the stage to the farthest seat on the 1 st floor	approx 25m
Width of the 1 st floor	approx 20m
From the stage to the nearest seat on the 2 nd floor	approx 24m (elevation 5m)
From the stage to the nearest seat on the 3 rd floor	approx 30m (elevation 10m)

Table 1: Hall Area

2.3.2 Procedures

Figure 2 shows 44 points selected to represent the characteristic shape of the stage. Figure 3 shows how these points were measured using a surveying instrument. Several of the forty four points are lighted in the figure.

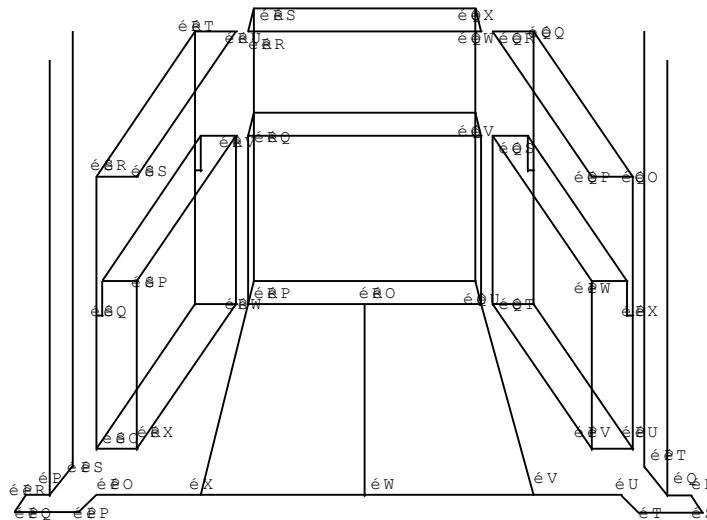


Figure 2: Stage Model showing key points



Figure 3: Using a survey instrument to get key points on the stage

Figure 4 shows the process of photographing the stage and measuring the camera position by a surveying instrument. After photographing the stage, we also photographed a female singer with a green screen in the background to make the extraction of her shape easier.



Figure 4: Photographing the stage

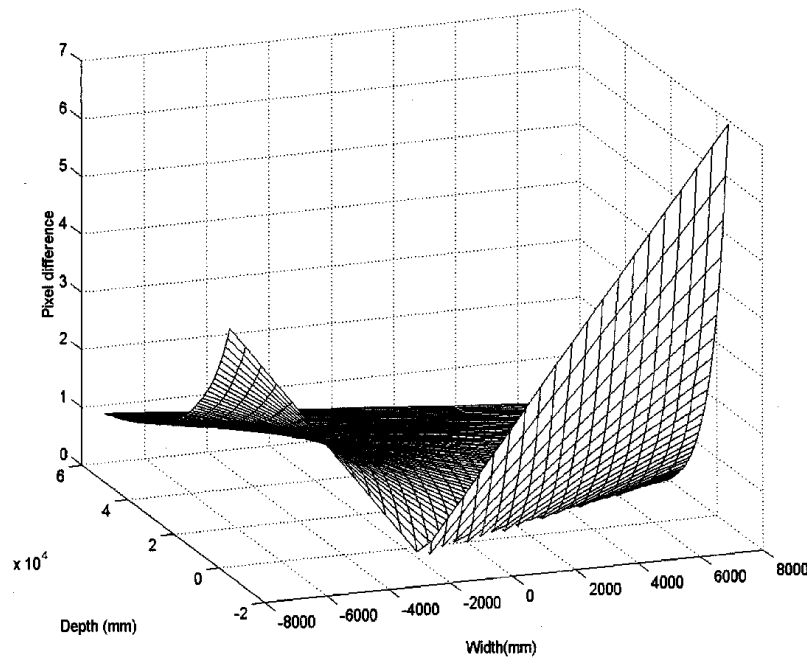


Figure 6: Distance Estimation

2.4 IMAGE SYNTHESIS AND AUDIO DISPLAY

Once the seat to be reserved is selected, shown by a lighted point in Figure 7, the following steps are performed to synthesize the view in the upper right corner of the figure:

1. The system interpolates the seat's 3D position from the two known photographing locations.
2. With the seat's 3D position assumed to be a image center, the stage points are projected onto the image plane of the camera by perspective transformation.
3. The system selects an appropriate captured image and maps its texture to the corresponding areas. For texture mapping the stage image must be divided into several triangular or quadrilateral patched areas. Using some of the 44 characteristic points, these patched areas are defined through a user graphical interface (GUI) as shown by red lines in Figure 8.
4. The plane of a female singer is also placed on the stage, based on the perspective transformation.
5. The intensity of the left and right stereo sound channels is adjusted with a sound localization processor establish to the direction of the virtual female singer.



Figure 7: Synthesized Stage Image



Figure 8: Definition of the Patched Area

3. Catalog Shopping System

The catalog shopping system also has image capturing and image displaying sides as shown in Figure 9. On the image capture side, two CCD cameras are placed on a bar at roughly the separation of the human eyeballs. The bar is attached to a manipulator that can rotate 90 degrees vertically and 360 degrees horizontally. Thus the cameras can change their orientation in a hemisphere at the center of which an object is placed. On the image display side an observer wears stereo glasses to observe an object in stereoscopic mode. The observer can interactively change his observation angle using a 3D mouse.

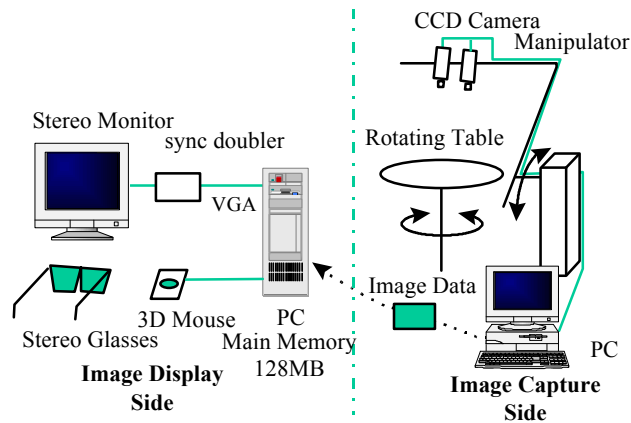


Figure 9: System Configuration

3.1 PRACTICAL PROCEDURES FOR TAKING THE PHOTOS

3.1.1 Image Capturing with the Manipulator

Images of an object are captured at intervals of 10 degrees both vertically and horizontally in the hemisphere, making a total of 162 image patches. Figure 10 illustrates photographing with the manipulator. Wide angle lenses are used to photo large objects with a distance of the arm's length. However they introduce distortion.



Figure 10: Photographing with the manipulator

3.1.2 Distortion Cancellation

When a wide angle lens is used, image distortion must be cancelled. Tsai[9] suggested a camera calibration technique including a distortion parameter, which can be modelled as follows:

$$\begin{aligned} X_d(1 + \mathbf{k}_1(X_d^2 + Y_d^2)) &= X_u \\ Y_d(1 + \mathbf{k}_1(X_d^2 + Y_d^2)) &= Y_u \end{aligned}$$

where \mathbf{k}_1 is the radius lens distortion parameter, X_d, Y_d are the distorted or true image coordinates, and X_u, Y_u are the ideal undistorted image coordinates. Using the calibration of two CCD cameras, we calculated the distortion parameters to be $1.79 \times 10^{-2} \text{mm}^{-2}$ and $1.62 \times 10^{-2} \text{mm}^{-2}$, respectively. Distorted images can be stretched to undistorted images by solving these two nonlinear equations.

3.1.3 Image Storage

The left and right images are merged into a single image with 640 by 480 resolution, which is stored as a single .avi file.

3.2 STEREO IMAGE DISPLAY

Figure 11 shows the author wearing stereo glasses to observe a product stereoscopically. He can change the viewing orientation with the 3D mouse held in his right hand.



Figure 11: Image Display

4. Conclusion

This paper presented practical methods for constructing virtual worlds with captured images for virtual shopping. Since existing malls have few attractive displays for showing products, we selected two purchasing processes, ticket reservation and catalog shopping, to study how the computer system can show things attractively and attractively. Since we believe that captured images give us more a realistic impression than computer graphics, we applied them to these two purchasing processes with interactivity for shoppers.

The concert ticket reservation system can show a view of the stage from every seat. The catalog shopping system can show products stereoscopically from any angle. These systems, combined with an e-cash transaction demonstration, were exhibited at an exhibition where the systems attracted many visitors. We found that the demand for the feelings of immersion created by the interactivity with captured images is strong. Computer graphics can create realistic images but with less interactivity because of the time-consuming calculations. Many cost effective and practical ways should be developed for applications. This construction method provides one solution. For internet-based shopping, sophisticated enhancements, such as image compression to lower the bandwidth for the internet, need to be developed.

REFERENCES

- [1] A. Dahl and L.Lesnick: "Internet Commerce", New Riders Publishing, 1996.
- [2] M.Okutomi and T.Kanade: "A multiple-baseline stereo", IEEE Trans. on Pattern analysis and Machine Intelligence, vol.15, no. 4, pp 353—363,1993.
- [3] P.W.Rander et.al.: "Virtualized Reality: An immersive visual medium", IEEE Multimedia, vol 4, no.1, pp 34-47, Jun.1997.
- [4] M. Hirose and T.Endo: "Building a Virtual World from the Real World", Mixed Reality, pp 183-197, Springer-Verlag, 1999.
- [5] S.M.Seitz and C.R.Dyer: "View morphing", Proc. SIGGRAPH'96, pp 21-30, 1996.
- [6] Oliver Faugeras: "Three-Dimensional Computer Vision", 169/174, The MIT Press (1996)
- [7] Quang-Tuan Lung, et.al.: "On Determining the Fundamental Matrix: Analysis of Different Methods and Experimental Results", INRIA Sophia-Antipolis BP93-06902 (1993)
- [8] N.Ozaki and S.Tachi: "Development of system that creates natural scene as a virtual world", Journal of Virtual Reality Society of Japan, Vol .2, pp 1-8, 1997.
- [9] Roger Y. Tsai: "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and Lenses", IEEE Journal of Robotics and Automation, Vol. RA-3, No.4, 1987.

BIOGRAPHY



Nobuyuki Ozaki is chief project manager in at Toshiba Corporation. He received his PhD in Computer Science from the University of Tokyo in 1996. He joined Toshiba in 1983 and has been with Toshiba since then. He first specialized in development of the process automation System, mainly for paper pulp industry and steel industry. After receiving his PhD in 1996, he changed his background to multimedia technology. His research interests are tele-presence, virtual reality, and computer vision and their implementation studies as products.

Contact information:

3-22 Katamachi
Fuchu-shi
Tokyo, 183-8512 Japan
Phone: +81-423-40-6884
Fax: +81-423-40-6012
Email link: <mailto:nobuyuki.ozaki@toshiba.co.jp>