

Technique for synchronizing two asynchronous images for 3-D display using SPA by optical time marking

T. Hasegawa, T. Akiyama, H. Unno, K. Uehira, H. Kasuga, and K. Yanaka

Kanagawa Institute of Technology
1030 Shimo-ogino, Atsugi-shi, Kanagawa, 243-0292, Japan

Abstract

We have studied a technique of displaying real-time 3-D images captured by Web cameras on the stereoscopic display of a personal computer using screen pixel access. Since it uses the image data in a video memory, it can be applied regardless of format of motion images obtained from Web cameras. However, since the images for the left and right eyes are sent independently through the Internet, they are not synchronized. If there is a great deal of delay between them, the observer cannot experience stereoscopic vision when looking at moving objects.

This paper proposes a technique to prevent this delay by optically marking the time on the images when they are captured by Web cameras, and it also describes an experiment that confirms its feasibility and results that confirm its basic operation.

Introduction

High-quality moving images such as those used in commercial broadcasting have recently been easily delivered over the Internet. As a result, Internet users can see high-quality moving images on their PC displays. The images delivered through the Internet that have recently become popular are those captured by Web cameras. Stereoscopic displays have also recently been produced for consumer use. Since 3-D image data can be captured by just placing two Web cameras side-by-side, input/output devices are now available that can be used to easily construct an Internet 3-D image system. However, it is still not easy for PC users to obtain an interface to create 3-D images from data captured by the two Web cameras in real time on a PC and display it on a 3-D display for moving images, because there are numerous formats for motion 3-D images and various kinds of 3-D displays. We have proposed an interface that can be used regardless of format for motion 3-D images by screen pixel access (SPA).¹ It can create 3-D motion images without the need for special hardware.

However, since the images for the left and right eyes are sent independently through the Internet, they are not synchronized. Therefore, if there is a great deal of delay between them, the observer cannot experience stereoscopic vision when looking at moving objects. This paper proposes a technique of preventing this delay by optically stamping the time on the images when they are captured by Web cameras. It also describes an experiment to confirm the feasibility of the proposed technique and presents the results.

Synchronizing Images for 3-D Display Using SPA by Optical Time Marking

3-D display using SPA

SPA is a method that uses image data displayed on the screen of a PC as the originals to create images for certain purposes. Since the image data that are displayed on the screen are in the video memory, we can use the data for any purpose if we access this area. This does not need the use of any special hardware. We used this technique to create a 3-D image from two Web camera images. The images from the two Web cameras were displayed in the viewer on the display. The image data were in the video memory of the PC after it received them from the Web cameras through the Internet. Figure 1 is a conceptual illustration of 3-D image composition achieved by SPA. L and R indicate the image data area in the video memory for the two Web cameras for the left and right eyes. The two Web-camera images are read from data area L and R in the video memory, composed into one image as a 3-D image, and then written in the video memory again. The S in Fig. 1 indicates the data area for this composed stereoscopic image. A 3-D image can be seen if the PC being used has a 3-D display.

Although all image data in a video memory are usually all displayed on the screen, we do not need to display the two original Web camera images. Therefore, we used a large virtual screen for

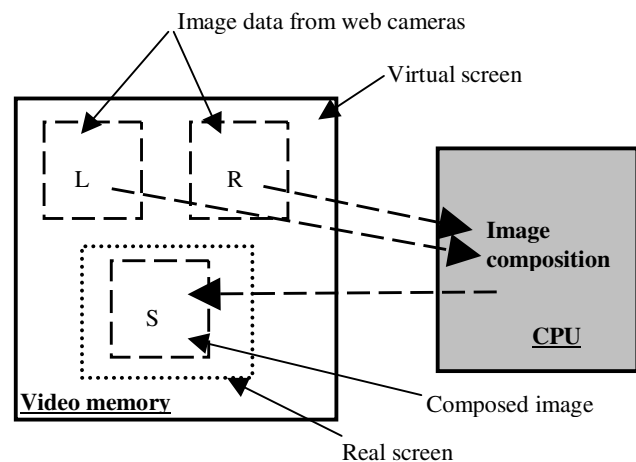


Figure 1. 3-D Image composition by screen pixel access. L and R indicate image data area for two Web cameras for left and right eye images. S indicates data area for composed image, which is stereoscopic.

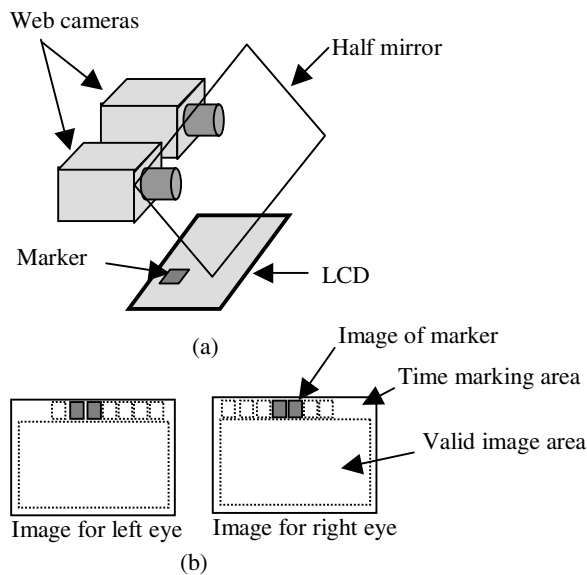


Figure 2. Conceptual illustration of synchronizing technique for 3-D image using SPA by optical time marking. (a) System configuration and (b) images captured by Web cameras.

SPA in this study, i.e., we used a video memory that was larger than that for a real screen. Only the necessary 3-D image was displayed on the real screen.

3-D image composition with SPA using Web camera images can accept any kind of format as long as they can be seen in any browser or viewer on the display of a PC. Therefore, PC users only need to install a program for SPA, without the need for special hardware, and they do not need to be aware of the image format for the Web cameras.

Synchronizing Images by Optical Time Marking

Since the images for the left and right eyes are sent independently through the Internet with the SPA method, they are not synchronized. If there is a great deal of delay between them, the observer cannot experience stereoscopic vision when looking at moving objects. We propose a technique that synchronizes images for 3-D display using SPA by optically marking the time. Figure 2 has a conceptual illustration of this synchronizing technique. Two cameras capture a marker image displayed on the LCD in addition

to the object with a half mirror. The LCD displays the marker changing its position every frame. The background of the marker on the LCD is black. The marker and object are captured together as one image. The image areas of these are separated as shown in Fig. 2. We used the image indicated by “valid image area” as the final image. The image area indicated by “time marking area” was used to identify the frame.

Figure 3 shows the process flow for synchronizing the images. To detect the frame number, the marker is first abstracted from composition made up of marker and object. The frame number is then identified from the marker position because the marker position changes every frame by a constant interval and the frame numbers correspond to the positions of the markers. Since the LCD is not synchronized with the Web cameras, markers may be at two successive positions in the image captured by the Web cameras as shown in Fig. 2. In such cases, we can calculate the frame delay between the LCD and a Web camera to a decimal place using the brightness ratio of the two markers in two successive positions or two markers in the same position in two successive frames. Therefore, the frame number is represented using a number with a decimal. The difference between the frame number of images for the right and left eyes is the frame delay between the two images. To cancel the delay between two images, one or both images are stored temporarily in the frame buffer. As the delay cannot be canceled completely when it is a decimal, frame interpolation is carried out for one of the images so that frame delay becomes an integral number. The two images are then composed into a 3-D image.

Experiments and Results

To evaluate our technique, we constructed an experimental system shown in Fig. 4 and carried out experiments. The two Web cameras captured moving images in frames of 640 x 480 pixels. At the same time, they captured the marker displayed on the LCD. The image area for the marker in the captured image (indicated as the time marking area in Fig. 2 (b)) is at top about 50 pixel rows. The marker is a blue square pattern of 50 x 50 pixels. The time marking area was divided horizontally into ten blocks. Each block corresponds to each marker. Evaluating whether the image of the marker was in a block within a frame was done by subtracting the background image from the image in the block, i.e., the marker image was determined to be in the block when the difference from the background was larger than a threshold. This difference was obtained as a averaged brightness difference of blue component for

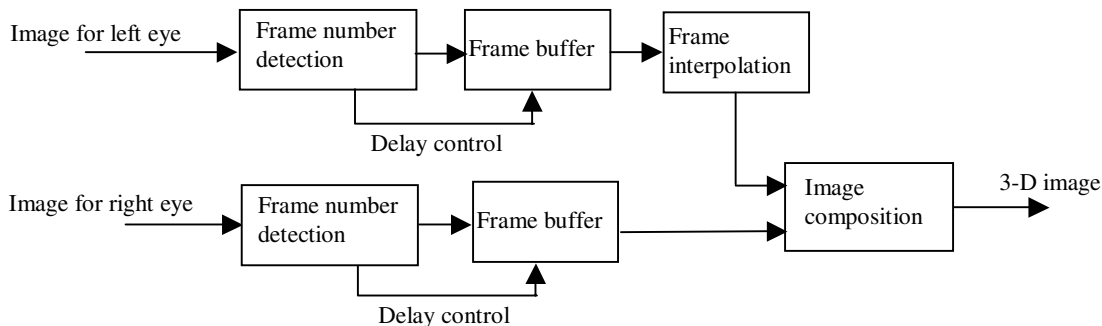


Figure 3. Process flow for synchronizing images.

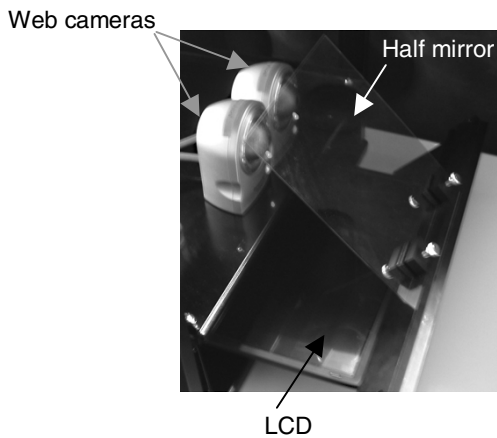


Figure 4. Experimental system

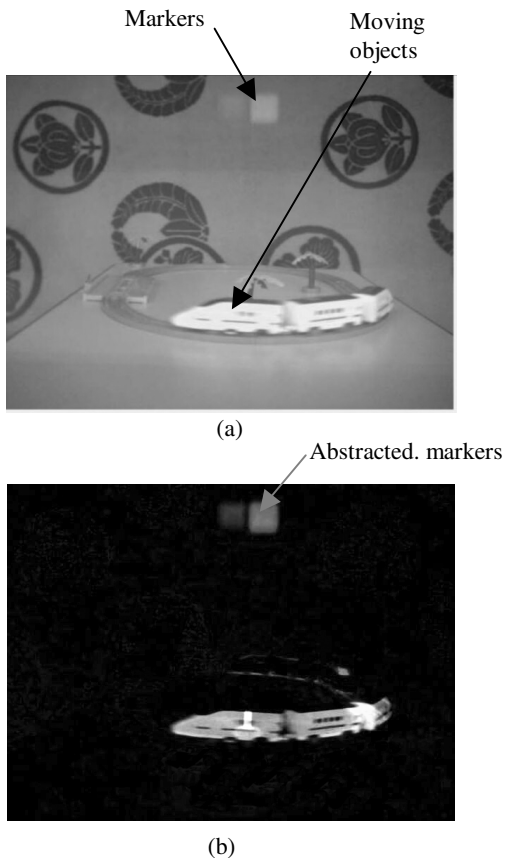


Figure 5. Example test image. (a) Original image, and (b) image after markers were abstracted.

every pixels in a block.

We used a PC with a parallax barrier-based stereoscopic display that had 1024 x 768 pixels. We used a test program developed in a previous study [1] that carried out SPA and displayed 3-D image on the stereoscopic display.

We evaluated whether the frame could be identified with the experimental system. We also evaluated the 3-D image generated by the synchronizing and SPA process and displayed on the stereoscopic display through a subjective test.

Figure 5 shows an example of the original image (a), i.e., an image of the objects and markers and an image in which markers were abstracted (b). We confirmed that the process of synchronizing the two images for the 3-D image worked as expected. As a result, normal 3-D images were observed for images generated by the synchronizing and SPA process and displayed on the stereoscopic display.

Although we confirmed the basic operation of our proposed technique, we still need to clarify details of the conditions under which it can be used as future work.

Summary

We studied a technique of displaying real-time 3-D images captured by Web cameras on the stereoscopic display of a personal computer using screen pixel access. There was a problem with delay between images for the left and right eyes that arose from passing through the network. We therefore proposed a means to compensate for this delay by optically stamping a time marker on the images when they were captured by the Web cameras. Using an experimental system constructed to evaluate our technique, we showed that each frame could be identified by the time marker, and using this identified frame we could display 3-D image generated by screen pixel access without any deterioration. However, we still need to clarify details of the conditions under which it works.

Acknowledgements

This study was partially supported by the Academic Frontier Promotion Program of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- [1] T. Hasegawa, et al.; "3-D Display Technique for Moving Pictures from Web Camera Using Screen Pixel Access", Proc SPIE (2006). To be published.

Biography

Takuya Hasegawa received his B.S. degree in network engineering in 2005 from Kanagawa Institute of Technology. Currently, he is a master course student in the graduate school of Kanagawa Institute of Technology and studying image transmission technology via Internet.

