Multiview-image 3-D Viewing System with Depth-Fused 3-D Display

Asami Shiraishi*, Toshiaki Fujii*, Tetsuya Kawamoto*3 and Kenji Mase*4

*1,*4 Graduate School of Information Science, *2 Graduate School of Engineering, Nagoya University

*3 Chukyo TV Broadcasting Co., Ltd, Japan

E-mail: *1 shiraishi@cmc.ss.is.nagoya-u.ac.jp, *2 fujii@nuee.nagoya-u.ac.jp

Abstract—We constructed a new 3-D display for our multi-view video interface Web-PSV using Depth-Fused 3-D method and conducted image processing so that observer can perceive continuous depth in the previous study. In this paper, we added a new functionality where users can change the viewpoint. In the experiment, we succeeded to generate multi-view image content and confirmed continuous depth even if viewpoint of image were switched.

Keywords-Multi-view video; Multi-view video interface; Depth-Fused 3-D display

I. INTRODUCTION

In multi-view video researches, multi-view video contents and viewers are developed for realistic sensation. PegScopeViewer (PSV) is a multi-view video interface that assists users to watch multi-view video contents. PSV enables users to watch the contents from their favorite viewpoints. Furthermore, users can mark a point (Peg-point) to a target on the contents at one scene, then marked target is fixed at center of image even viewpoint of multi-view video is switched. However, this interface can show only 2-D contents, and it’s difficult to generate 3-D contents, therefore a 3-D display which is suitable for PSV is required.

We constructed a new 3-D display for our multi-view video interface using Depth-Fused 3-D (DFD) method. DFD display is one of autostereoscopic display. In the DFD method, by superimposing back and forth two 2-D images of which ratio of intensity were varied based on depth of objects, observer perceives depths and feels as 3-D by optical illusion. In this display, two 2-D images are fused by a half-mirror. Because measuring depth information of objects is required for determinant of suitable intensity of 2-D images, it’s difficult to generate the back and forth images.

In this paper, we studied image processing so that observer can perceive continuous depth without measuring depth and conducted gradational processing to two 2-D images. We extended our previous study where viewpoint of the processed image was fixed. We succeeded to generate multi-view image content and confirmed continuous depth and could feel as 3-D even if viewpoint of image is switched.

II. RELATED WORKS

A. PegScopeViewer (PSV) and Web-based Peg-Scope Viewer (Web-PSV (PSV))

"Pegged to Point Browsing" was first proposed by Tokai et al [1]. It is a view-switching method for browsing a 3D scene from various directions. Assume a scene shown in Figure 1 (a), where there are four cameras and two objects. When viewing the multi-view scene, one of the showing methods of the scene is to provide views of each camera are like as Figure 1 (b), by switching these views. Although we can understand the viewpoint is changing, we cannot pay attention to any objects in the scene and we are confused. Therefore, we developed a re-composing method of the original video to controlled composition of the frame and switching views. Figure 1 (c) or (d) show the examples of our proposed method. We can perceive that we walk through the scene with watching the object from the switching view sequence.
To evaluate the "Pegged to Point Browsing" method, we developed a novel viewing system based on this concept: named "PegScopeViewer". It provides a control interface to viewers, as shown in Figure 2. This is a snapshot of a GUI-based Peg-SCOPE viewer. It contains the standard video play buttons, camera selector, zoom, peg-point list selector, etc. Peg-points, which are given as a list, are manually collected or given by image-processing object tracking. The camera selector chooses the camera position. Camera selection can be performed while the video is playing. PegScopeViewer can provide stabilized targeting to an object even if the original camera arrangement at the filming did not place the object in the center of the frame of each camera shown in Figure 3. Figure 4 illustrates such a viewing result.

We then extended our PegScopeViewer to a networked video viewing system for multiview with a video streaming technology. The system exploits Microsoft’s Smooth Streaming to enable a synchronized hierarchically-encoded streaming. Multiple Smooth Streaming Media Element (SSME) players are evoked on a web client machine to download adequate stream files according to the given performance of the network and the client. The bandwidths for SSMEs are managed so that the highest bandwidth is given to the currently-watching channel while the lower bandwidth are shared by other channels. Thus, a very sophisticated and network-efficient viewing service is realized. The system incorporates a target-centered viewing interface to enhance the multiview video viewing experiences. We named this system the "Web-based Peg-Scope Viewer (Web-PSV)."

**B. Depth-Fused 3-D (DFD) Display**

"Depth-fused 3D" or "DFD" display was first developed by Suyama et al [5]. The DFD displays two 2-D images at the front and rear frontal-parallel planes. The observer perceives 3D image by "visual illusion". The perceptual phenomenon is composed of two processes: depth-fusing and continuous-depth perception. In the depth-fusing process, two 2-D images displayed at the front and rear frontal-parallel planes are perceived as a single image at one depth. In the continuous-depth perception process, we can perceive a continuous depth change of the depth-fused image when the luminance ratio between the two 2-D images at each point is continuously changed according to the distance ratio between the object and the two image planes.
The concept of the DFD display is shown in Figure 7. The display is composed of two 2-D images displayed at the front and rear frontal-parallel planes. The two 2-D images are shown on the front plane and rear plane simultaneously. The only difference between them is their luminance distributions, which are calculated according to the depth of each object in the 3-D space. The luminance ratio between the two 2-D images at each point is obtained from the distance ratio between the object and the two image planes. Note that DFD display only requires two 2-D images, however it needs "depth information" of the scene.

The DFD display can be simply composed of 2-D displays as shown in Figure 8. Suyama et al. developed the actual composition of the prototype DFD display as shown in Figure 8, which contains several sets of a half mirror and two 2-D displays. The half-mirror was used for optically constructing the front and rear frontal-parallel planes, but it could be omitted if the 2-D displays were sufficiently transparent. Recent development of the transparent LCD displays seems to be a promising candidate for the DFD display device.

III. MULTIVIEW-IMAGE 3-D VIEWING SYSTEM

We adopted the DFD display as a 3-D display of our WebPSV system. We first constructed a DFD display as shown in Figure 9. Similar to the Suyama’s prototype system, we placed a half mirror at the center. Two 2D color LCD displays were used as the front and read display. The half mirror placed in the center optically fuses two 2-D LCD display images. The two LCD displays and the half mirror have to be carefully aligned so that the perceived images from the observer’s viewpoint can be fused. In the real setup, we registered all of them manually because of the difficulty of registering them according to the ideal calculation. But through the experiment, we confirmed this kind of ad-hoc method works quite well.

As the experimental contents, we captured a wrestling scene, where wide-angle (i.e. all-around in this case) multi-view images can be obtained. Figure 10 shows our camera setup when shooting a wrestling scene. We set 4 cameras around the players and set the convergent point at the center. Each camera’s position, orientation, focal length, brightness are set independently, therefore we adjusted these parameters after we captured the scene.

![Figure 7: Concept of DFD displays [5].](image)

![Figure 8: Composition of DFD prototype system [5].](image)

![Figure 9: Our DFD display: principle and overview. [3]](image)

![Figure 10: Camera setup of the wrestling scene shooting](image)
The most challenging part of this research is how to generate front and back images for the DFD displays. We used only four 2-D cameras and no 3-D information can be utilized. One of the solutions is to use depth-camera, which become available in recent years. Many conventional researches use e.g. Kinect as a depth camera. However, these depth cameras have several limitations: distance of the object, multiple use, etc. In our experimental setup, it is impossible to measure the depth using Kinect because of the distance to the object, it is >10m away from the cameras. Another point is that we cannot use multiple depth cameras that exploit pattern-based active range sensors. We used for cameras, so if we used four depth cameras, each of them conflict each other. Therefore, we cannot use depth camera in this situation. Another solution would be using stereo-base depth camera. But it requires double of cameras and not realistic in this case.

The main drawback using DFD as our 3D display system is that it requires depth information. We took an approach in which a simple 2D-3D conversion is adopted. We first focus on the characteristics of the scene. The scene consists of one ground “plane” and a few persons. Using these a-priori knowledge, we conducted a simple image processing to the front and rear images shown in Figure 11. These a-priori knowledge can be interpreted like this: the scene consists of a slanted plane from the bottom to the top. It can viewed as a simple 2-D to 3-D conversion with 3-D model, where the model is a simple slanted plane. Based on this consideration, we conducted a gradational image processing to both front and back images. Figure 11 shows the processed images. In the front image, the intensity of the bottom part is relatively high, and the intensity gradually decreases according to the vertical axis. On the other hand, in the back image, the intensity of the upper part is relatively high, and it gradually decreases to the lower part.

Figure 12 shows the result image captured from the observer’s viewpoint. The observer can view the fused image generated by DFD display. The observer can switch the viewpoint by pressing a key and enjoy the PegScopeViewer’s function. We conducted simple subjective test and confirmed that this kind of 3-D display is suitable to PSV application. It is also advantageous that no other device such as stereo camera setup and depth camera is required to realize 3-D Web PSV.

![Figure 11: Original image, back image, front image.](image)

**ACKNOWLEDGMENT**

This research is supported by Ministry of Internal Affairs and Communications (MIC) SCOPE (No.111706003).

**REFERENCES**


