Journal of the Optical Society of Korea, Vol 5, No. 2, June 2001, pp. 49-54

Development of a Compact 3-D HDTV Camera with Zoom Lens

H. Yamanoue*, M. Okui, F. Okano, and I. Yuyama
NHK Sci. & Tech. Res. Labs, JAPAN

(Received December 5, 2000)

Research on shooting conditions of 3D program production for natural 3D images has been continued. In the study, it has been shown that orthostereoscopic conditions bring about no inconsistency between depth information from perspective of the lenses and that from binocular parallax. A newly developed 3D camera is based on orthostereoscopic conditions, which result in compactness of the camera (weight 8 kg). At the same time, the new camera has a zooming function and is valuable in many ways, especially sport broadcasting. In this paper, we give an outline of the newly developed 3D HDTV camera and the results of subjective evaluation tests on psychological effects of the images shot by the camera. These tests show that the images shot by this camera are more powerful and comfortable to view than those shot by existing 3D cameras.

OCIS codes : 110.2970.

I. INTRODUCTION

Broadcasting technology is making rapid progress toward the 21st century on the back of advances in digital applications. Amid this background, we are focusing on three-dimensional (3D) HDTV as a promising new broadcasting service [1]. This 3D HDTV uses two HDTV cameras, one for the right eye and the other for the left eye, to shoot an object (Fig. 1) ; pictures thus shot produce stereoscopic effects. This method is called a binocular approach. Unlike real stereoscopic vision seen in a natural setting, these 3D pictures can be unnatural and difficult to watch because of a number of shooting conditions such as the distance between the right and left cameras, the focal length of the lens, and the direction of the lens axes [2]. Typical distortions in 3D images are the puppet theater effect (the object looks unnaturally small) [3], and the cardboard effect (the object looks unnaturally flat, like a thin plate) [4]. To properly display these stereoscopic pictures by 3D television, we need to make sure that the images look natural in terms of relative size and depth sensation; a natural appearance is the key both for creating a sensation of reality and power and ensuring that the viewer can enjoy the pictures without eye strain. Through television program production, we have studied the distortions and other problems of stereoscopic pictures by changing the conditions of shooting and viewing. These studies have revealed the following: For a stereoscopic image to naturally represent the size and depth of an object, it is important to match the depth sensation from the perspective of a 2D image projected onto the retina and a sensation of depth from binocular disparity. And also in order to achieve this matching, orthostereoscopic conditions are effective [5]. To satisfy the orthostereoscopic conditions, the distance between the right and left cameras must be kept the same as the average distance between the right and left pupils of the viewer, while

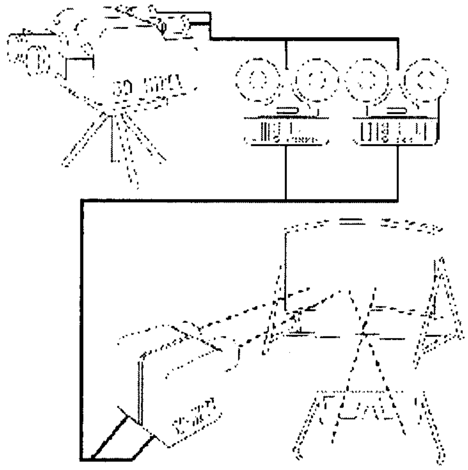

FIG. 1. 3D Hi-vision.

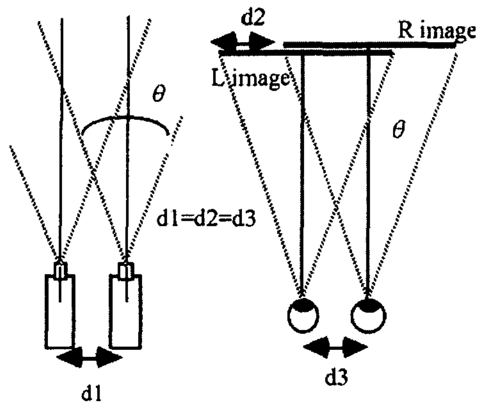

FIG. 2. Orthostereoscopic conditions.

keeping the lens axes of these two cameras parallel during the shooting.

In conventional 3D television program production, stereoscopic effects are produced by widening the distance between the right-left cameras or by moving the point where the cameras' light beams converge back and forth. This is a time-consuming process as these procedures must be repeated in each scene until naturalness is obtained in the 3D picture. If we can keep the lens axes parallel and fix the distance of camera separation between them, thus eliminating the need for adjusting the lens axis, 3D television programs could be produced much more efficiently. The 3D television camera can also be made lighter and smaller as the mechanism for lens axis adjustment is no longer needed. The 3D HDTV camera system we have developed embodies such improvements. A conventional 3D HDTV camera basically uses a lens with a fixed focal length, since ordinary zoom lenses in the right and left cameras move independently from each other, causing the lens axes to fluctuate during zoom strokes and making differences between the focal lengthes of the L and R lenses. This fluctuation in the lens axes produces geometrical distortions such as horizontal or vertical gaps between the right and left images. The new camera system with two lens axes kept parallel has a common zoom mechanism, and a compensation mechanism for the fluctuation on the lens axes is built inside the optical block. These improvements make zooming very smooth; the depth sensation changes smoothly while the camera zooms in and out, and the created 3D pictures are realistic and natural looking. This zoom mechanism is ideal for shooting sports programs where there is a lot of movement, improving the range of visual expression and making program production more efficient.

This paper outlines this new 3D HDTV camera system, and reports the results of a subjective evaluation of stereoscopic pictures shot with this camera.

II. OUTLINE OF THE NEW 3D HDTV CAMERA SYSTEM

1. Orthostereoscopic conditions

In order to view a stereoscopic image with natural depth amd size sensation as pointed out above, there should be coherence between the depth sensation from lens perspective and that from binocular disparity. To satisfy this requirement, we can effectively use orthostereoscopic conditions, based on the way the new 3D HDTV camera has been developed. Fig. 2 shows the orthostereoscopic conditions.

- The lens axes of the camera lenses are parallel.
- The distances between the two cameras, between the right and left images, and between the pupils are the same.
- The field angle of the camera lenses and the view angle of the image are the same.

2. Compactness and light weight

With conventional 3D cameras, it is necessary to change the camera separation or move the point at which the lens axes of the cameras converge back and forth. Fig. 3 (L) shows a conventional 3D camera system and its lens axis adjustment mechanism. We

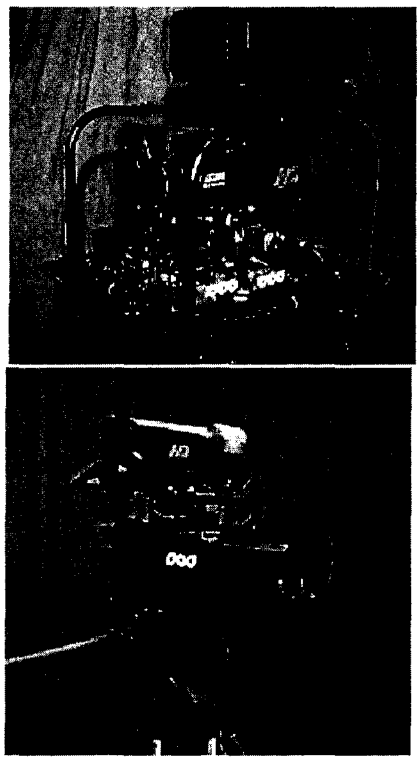

FIG. 3. Conventional camera & New camera.

can see (center of the photo) that the mechanism for adjusting the position of the right-left cameras and their lens axes takes up much more of the system than the camera heads. In this case, the system as a whole weighs 25 kg while the right-left camera heads weigh only about 4 kg. Fig. 3 (R) shows the new 3D HDTV camera. As described earlier, the system has no mechanism for adjusting the lens axes or camera positions, and the right and left cameras share the same casing, resulting in a substantial weight reduction (the whole system weighs about 8 kg including the lenses). Further, it is small enough to be carried on the shoulder without difficulty.

3. Positions of lens axis and CCD center

The distance between the right and left images during viewing (d2 in Fig. 2) is one of the orthostereoscopic conditions. To satisfy this condition, the electrical center of the CCD and the lens axis are kept horizontally apart by a distance of d4 (Fig. 4) within the optical block of the camera head. It should be noted that d2 in Fig. 2 and d4 in Fig. 4 are theoretically same [6]. As a 120-inch projection screen is used for this new camera system, the 2/3-inch CCDs are moved 1.2 % horizontally to satisfy this condition (this projection screen size of 120 inches is the most common in 3D HDTV theaters).

4. Problems of 3D zooming and countermeasures

Zooming is one of the major features of this new 3D HDTV camera system. Several problems are inherent in 3D zooming, but this system successfully deals with them as explained below.

There are two major problems that may occur during 3D shooting. One of them is the difference between the right and left images as the right and left zoom lenses do not have exactly the same precision of zooming, focusing, and iris control. The other is the fluctuation of the lens axes by zoom strokes.

1) Control precision of the lens

This disparity in control precision between the right and left zoom lenses is mainly attributable to the use of two control motors, one for the right lens and the other for the left lens. This new camera system, taking advantage of the fact that the two lens axes are kept parallel, uses a single motor for controlling the right and left lenses. This mechanism enables very precise controlling of the right-left lenses (focus, iris, size adjustment) as the object is zoomed in and out.

2) Light axis movement

The fluctuation of the lens axes during zooming is thought to be caused mainly by the following three factors: distortion of the lens itself, too much play between the lens and the camera structure, and because the lens axis does not coincide with the CCD center. (Fig. 5). We examined these three factors and found that the third factor, wrong positioning of the lens axis and the CCD center, has the significant influence. In particular, if vertical positioning is wrong between the right and left lenses (Fig. 5), there occurs a vertical gap between right and left images, making it very difficult to see the 3D images. This new camera system has a mechanism for adjusting the vertical position as well as one for horizontal adjustment as shown in Fig. 4. These mechanisms suppress geometric distortions caused by lens axis fluctuations below a level discernible by the human eye.

3) Distortion by zooming in reconstructed space

The 2/3-inch zoom lens used in this system has a focal length ranging from 12 to 96 mm (the maximum focal length is eight times the shortest focal length), with its field angle varying from 43.6 degrees to 5.7 degrees. To satisfy the orthostereoscopic conditions shown in Fig. 2, we need to change the viewing distance in relation to the focal length (field angle) of the lens. When the focal length is 12 mm, for instance, the orthostereoscopic conditions can be met by setting the viewing distance at 3.3 meters (field angle 43.6 degrees, using a 120-inch screen). If it is 96 mm, the object will have to be placed 26.7 meters from the camera (field angle 5.7 degrees). It is practically impossible to change the viewing distances in this way

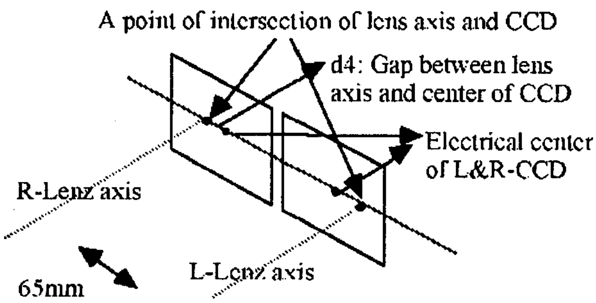

FIG. 4. Gap between lens axis and center of CCD (d4).

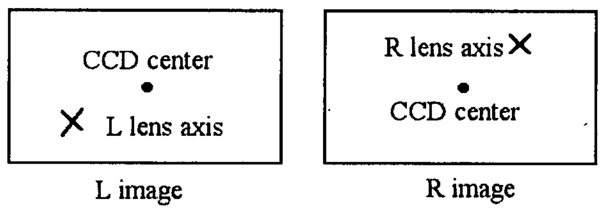

FIG. 5. CCD center and lens axis.

during zooming. If the viewing distance is short and the focal length of the lens is long, the produced image will inevitably look like a flat plate (cardboard effect) as it is compressed depthwise.

III. 3D VISUAL EFFECTS OF THE NEW 3D CAMERA

We set a number of evaluation items and studied the factors influencing them to examine the visual effects of 3D images created by this new camera system. Some of the 3D pictures shot on location in Africa, which was done to check the applicability of this camera system as well, were used as samples for evaluation. For comparison, images on Stereoscopic Test Materials were used [7].

1. Pictures for evaluation

Moving pictures shot by this camera system, 14 types in all, and one image from the Stereoscopic Test Materials were used for evaluation. They were shown in two ways: as 3D images and as 2D images shot by the left-eye camera. Each moving picture is 15 seconds long per scene. Pictures No.1 through No.5 were shot with the cameras' focal length set at 12 mm; they satisfy the orthostereoscopic conditions under the viewing conditions of this experiment. Pictures No.6 through No.10 were shot with medium to long focal lengths; the field angle of the lens is narrower than the display angle of the screen. All these pictures were shot with the lens's focal distance fixed at each shooting. Pictures No.11 to No.14 were shot while slowly zooming in on the objects. Picture No.15, named "the flower pot," was from the Stereoscopic Test Materials (it has been used on a number of occasions as an evaluation picture, including the evaluation tests of MVP) [8]. It was shot by having the lens axes converge, with other conditions as follows: a distance of 5 meters between the cameras and the point of convergence, that of 65 mm between the right and left cameras, and the lenses with fixed focal point (12 mm focal length) each. Fast-moving objects and rapid camera movements were not employed in this experiment.

2. Evaluation test

The viewing distance was set at 3.3 meters so as to satisfy the orthostereoscopic conditions when the display screen used in the experiment was 120-inch wide and the camera lenses each had the focal length of 12 mm (field angle of 43.6 degrees, Fig. 5). The viewing distance was set the same as the focal depth of the human eye (+/-0.3D: 3.3 m) so as to minimize the effects of contradiction between the accommodation and convergence of the lenses. The pictures were shown in two ways as described earlier, as 3D pictures and as 2D pictures. The experimental results of each sample picture were graded at the following five levels for each of the evaluation items shown below.

- Evaluation items :

Distracted by the frame, Easy to see, Looks like a miniature model, Strong depth sensation, Sensation of reality, Tiring to see, Stands out from the screen surface, Eye strain, Natural (size, depth sensation), Large, Comfortable, Powerful, Flat, lack of depth sensation.

- Categories:

5: Yes, 4: Yes, to some degree, 3: Not sure, 2: Not very much, 1: No

Each of the test participants, sixteen men and women aged 20 to 50, was shown the sample pictures for 15 seconds repeatedly until he or she answered all the questions. Before the test, the visual ability of all of them had been checked using the Stereoscopic Test Materials, and the results confirmed that none had any problem in fusing 3D pictures, moving or stationary.

3. Test results

We then analyzed factors from the subjective evaluation data, extracting two effective factors. The first factor (vertical axis) represents "power and sensation of reality" (40 % contribution ratio) and the second (lateral axis) "ease of viewing" (29 %). Fig. 6 shows the factor scores of each sample picture about "power and sensation of reality" when compared with the 2D pictures; Fig. 7 shows factor scores about "ease of viewing" when compared with the 2D pictures.

4. Power and sensation of reality

It has been recognized that 3D pictures are more powerful and realistic than 2D images, a notion clearly supported by the experimental results this time as

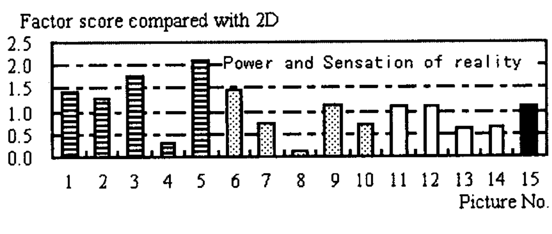

FIG. 6. Factor score about power and sensation of reality.

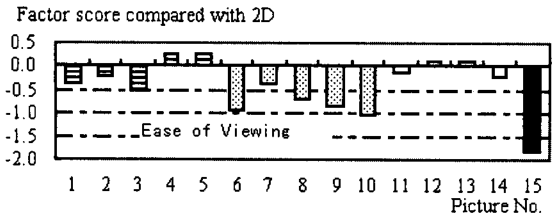

FIG. 7. Factor score about ease of viewing.

shown in Fig. 6. Even so, some pictures, namely No.4 and No.8 did not show as large an increase in the factor "power and sensation of reality" as the other pictures did, though all were shot under the same conditions. This result indicates that, even under the same shooting conditions, stereoscopic effects vary depending on the patterns of the pictures. These two pictures, No.4 and No.8, when shown as 2D images, scored lower on "depth sensation" than the other pictures in the subjective evaluation test. On the other hand, they scored higher on "flatness" than the others. These results mean that pictures No.4 and No.8 lack depth sensation in their 2D picture patterns to begin with.

5. Ease of viewing

Factor scores of the 3D pictures about the ease of viewing, as shown in Fig. 7, are as a whole lower than those of the 2D images. However, the 3D pictures shot by the new camera (No.1-No.14) scored higher than the 3D picture (No.15) shot by the conventional method, especially No.1-No.5 which were shot when the focal length was fixed at 12 mm (the orthostereoscopic conditions for shooting and viewing were met) and No.11-No.14 shot by slow zooming. Some of these 3D pictures even scored better than the 2D counterparts regarding the ease of viewing.

What accounts for this? Naturalness of the reproduced 3D space and suppressed distortion caused by the frame are possible reasons. The factor loading indicates that these two evaluation items exert a large influence on the "Ease of viewing". This naturalness of the reproduced 3D space means whether the size of the object and depth sensation are naturally reconstructed in a 3D picture. Distortion caused by the frame refers to a type of distortion that occurs in a binocular stereoscopic system when part of a 3D picture formed in front of the display area is cut by the frame. This distortion undermines the naturalness of 3D pictures by giving rise to a half-occluded region between the frame and the 3D image.

1) Pictures that satisfy orthostereoscopic conditions

Pictures No.1-5 satisfy the orthostereoscopic conditions, as referred to earlier. These conditions by definition are the most effective in producing a natural-looking 3D image from the object in terms of size and depth sensation.

2) Pictures shot with middle/long focal length lens

Pictures No.6-10 were shot using lenses with medium to long focal lengths. For these images to satisfy the orthostereoscopic conditions, it is necessary to lengthen the viewing distance in relation to the field angle of the lens. When the focal length of the camera lens is 96 mm (field angle of 5.7 degrees), for instance, the viewing distance will have to be 27.6 m (120-inch projection screen is used). If the distance to the object is shorter than this 27.6 m, its image will be formed at a point before the screen, thus increasing the possibility of distortion caused by the frame. As the actual viewing distance was 3.3 m, the created images were all compressed depthwise. This is why these pictures scored worse than those (No.1-5) which satisfied the orthostereoscopic conditions.

3) Pictures shot by conventional method

Picture No.15 was shot by a conventional method, in which lens axes of L R cameras converge. In this picture the distance between the camera and the convergence point is about 5 m. So the image of any object set before this point is formed before the screen, giving rise to the distortion caused by the frame. The need to have the lens axes converge also increases the likelihood of distortion of reconstructed space and the keystone problem. Because of these problems, picture No.15 fared worse in both evaluation items than those images that satisfied the orthostereoscopic conditions.

4) Pictures shot by zooming

These zoom pictures have the following characteristics, attributable to the position of the lens axes and the CCD center. When zooming in, the 3D image of the object on the screen is enlarged and at the same time it begins to stand out. Just the opposite occurs when zooming out; the 3D image on the screen shrinks while it begins to recede (moving away from the viewer). Although this zooming is inevitably accompanied by some depth-wise compression of the reproduced space, the sensation of movement is extremely close to what we naturally experience. This naturalness of movement sensation explains the high scores in the "Ease of viewing" category.

IV. SUMMARY

This paper reported on the newly developed 3D HDTV camera system based on orthostereoscopic conditions, which are effective in producing three-dimensional images that look natural in terms of relative size and depth sensation. The system is compact and lightweight with a zoom mechanism, and the features were fully tested and utilized on location in Africa. Subjective evaluations were conducted to examine the psychological effects produced by the 3D images shot by this camera system. The results show that the pictures produced by the system have enough power and sensation of reality, and that the pictures are less straining to the human eye than conventional 3D images. Further studies will be made on the distribution of binocular disparity, time-spatial characteristics, and other properties characteristic of 3D images. A greater number of sample pictures will also be used for the evaluation.

*Corresponding author : yamanoue@strl.nhk.or.jp.

REFERENCES

[1] H. Ywates and S. Yano, IEICE Tech. report, **HCS 96-1**, 1 (1996).
[2] D. L. MacADAM, SMPTE J. **62**, 271 (1954).
[3] H. Yamanoue, SMPTE J. **106**, 225 (1997).
[4] S. Herman, SMPTE J. **80**, 539 (1971).
[5] H. Yamanoue, M. Nagayama, M. Bitou, and J. Tanada, Proc. of IS&T/SPIE Strereoscopic Displays and Applications IV, 3295-12, 111 (1998).
[6] Stuink, L. M. and T. Iwamoto, Proc. of IROS '90, (1990), 429.
[7] Recommendation ITU-R JWP 10-11Q, Subjective Assessment of Stereoscopic Television Pictures, (1999).
[8] ISO/IEC JTC1 SC29/WG11 N1373, Report on the verification test on multiview profile, Oct. 1996.