THREE-DIMENSIONAL TELEVISION SYSTEM BASED ON SPATIAL IMAGING METHOD USING INTEGRAL PHOTOGRAPHY

Jun Arai

NHK (Japan Broadcasting Corporation)

ABSTRACT

A three-dimensional (3-D) television based on the spatial imaging method using integral photography (integral 3-D television) has three main advantages; displaying 3-D images without having to use special 3-D glasses, being viewable without restricting viewer posture, having a system applicable to live broadcasting service. I describe the principle of an integral 3-D television and introduce an experimental integral 3-D television with 33-mega-pixel imaging devices. In this experimental television, a lens array is configured of 400 lenses in the horizontal direction and 250 lenses in the vertical direction. The results from capture and display experiments confirm 3-D video reproduction that changes seamlessly with viewing position.

Index Terms— three-dimensional television, spatial imaging, integral photography

1. INTRODUCTION

The three-dimensional (3-D) image technique, which uses 3-D glasses, provides high-quality 3-D images in a simple manner. Therefore, it is widely used in movie theaters, and some broadcasting operators have started to distribute stereo 3-D programmes. For more comfortable viewing, however, we feel that 3-D display not requiring 3-D glasses is preferable.

A 3-D television based on the spatial imaging method using integral photography (IP) (integral 3-D television) has three main advantages; displaying 3-D images without having to use special 3-D glasses, being viewable without restricting viewer posture, and having a system applicable to live broadcasting service. Considering that 3-D video broadcasting is used in many different places, a system that allows viewers to watch a program from any posture without wearing 3-D glasses is preferable. A system that allows live presentation is essential for a broadcasting service. Since integral 3-D television satisfies these requirements, we are developing a 3-D television system based on the integral method.

In Section 2, I briefly describe the 3-D display current techniques. In Section 3, I explain the principle and display characteristics of integral 3-D television in terms of resolution and viewing area. In Section 4, I introduce an experimental integral 3-D television with 33-mega-pixel imaging devices.

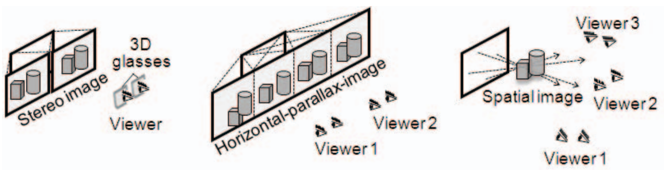
2. THREE-DIMENSIONAL DISPLAY

The difference in a scene as viewed by the left and right eyes is a major factor in producing a 3-D effect in vision. This difference is called binocular parallax. Accordingly, 3-D display requires at least two images, one for the left eye and one for the right. These two images must be accurately separated and presented to the left and right eyes.

Three-dimensional glasses separate the left and right images by using polarizing filters or an optical shutter (Fig. 1(a)). Because this approach is technically simple and produces good image quality, it is widely used in movie theaters and other such venues. Nevertheless, only one pair of left and right images is displayed, so even if the viewer moves, the 3-D video that is seen does not change. Another factor in our sensing of the 3-D effect is motion parallax. Motion parallax is the difference in viewable images created by movement of the viewpoint. Generally, methods that require 3-D glasses cannot provide motion parallax.

One 3-D display method that can present motion parallax is the parallax panoramagram (Fig. 1(b)). In this method, the subject is photographed from multiple positions in the horizontal direction (four, for example) rather than using a single left-right image pair. When the image is displayed, a slit array or lenticular lens controls the directions from which the multiple images can be seen. The different images can only be presented in the horizontal direction, so viewers must keep their two eyes level to experience the 3-D effect. This constraint is the same for 3-D glasses.

IP (Fig. 1(c)) and holography are techniques for achieving 3-D display without using 3-D glasses and without restricting viewer posture. IP in particular offers the advantage of using natural light to photograph a subject and display a 3-D image. Therefore, we are developing integral 3-D television on the basis of IP.



(a) Stereogram (b) Panoramagram (c) Spatial imaging (Integral 3-D television)

Figure 1. Three-dimensional display technology

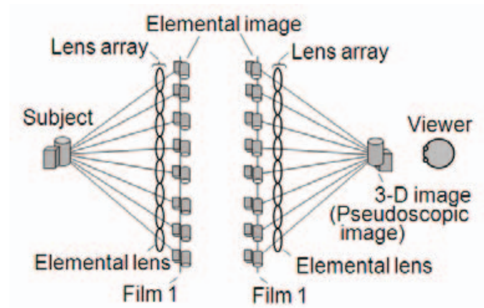
3. INTEGRAL THREE-DIMENSIONAL TELEVISION

3.1. Principles

The French physicist Lippmann proposed IP as a 3-D photographic technique [1]. Integral 3-D television is based on IP and enables shooting a subject and displaying an image in real time. IP uses a lens array composed of many convex lenses and film to capture images (Fig. 2). In the capturing stage (Fig. 2(a)), each convex lens creates an image of the subject, so the number of images captured by the film is the same as the number of lenses. We call these convex lenses “elemental lenses" and the images of the subject produced by the elemental lenses “elemental images.” In displaying the image, a lens array is placed in front of the film on which the elemental images are recorded (Fig. 2 (b)). The light rays from each elemental image then pass through an elemental lens and return in the opposite direction of the incoming light rays during the capturing stage, so the light emanating from the subject is reproduced. The result is that the viewer can see a 3-D image without having to wear 3-D glasses. Natural light is used both when capturing the subject and when displaying the 3-D image.

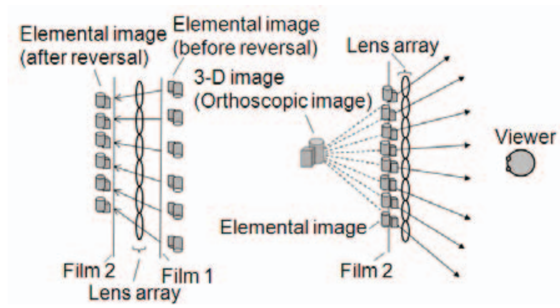
With the configuration shown in Fig. 2, the depth of the reproduced 3-D image is reversed relative to the subject, creating what is called a pseudoscopic image. In Fig. 2(a), for example, the cylinder is in front of the block as seen from the film, but in Fig. 2(b), the cylinder appears to be further away than the block from the viewer. The pseudoscopic image effect can be prevented by reversing the point symmetry of each elemental image. This reversal process can be accomplished by two-step capturing. First, the elemental image is recorded on film 1 (Fig. 2(a)). Next, the elemental image is transferred to film 2 via a lens array (Fig. 3(a)). When a lens array is placed in front of the film from the second capturing stage (Fig. 3 (b)), a 3-D image with correct depth display is produced.

Two problems of achieving real-time capturing and display in IP arise from the use of film and the two-step capturing required to eliminate pseudoscopic images. The first problem can be averted by replacing the film with a



(a) Capturing stage (b) Display stage

Figure 2. Principles of integral photography



(a) Reversal of elemental images (b) Display stage

Figure 3. Display of orthoscopic image

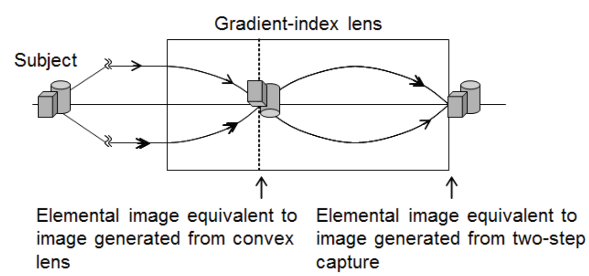


Figure 4. Image formation using gradient-index lens

CMOS or other such electronic capturing device and an LCD panel or other such display device. The second problem can be averted by using a gradient-index lens or a concave lens to generate an elemental image that is equivalent to the elemental image produced by two-step capturing. The refractive index of a gradient-index lens varies parabolically in the radial direction from the center of the lens, as illustrated in Fig. 4 [2].

Thus, integral 3-D television can be achieved by replacing film with electronic imaging devices and using a gradient-index or concave lens for the elemental lens for capturing.

3.2. Resolution of displayed 3-D image

A 3-D image generated by integral 3-D television can be thought of as a stack of flat images superimposed in the depth-wise direction. Imagine an elemental image projected by an elemental lens onto a position where a 3-D image is reconstructed (Fig. 5). Assume that the maximum frequency of stripes per radian of the projected image is α, which is called the "projection spatial frequency." When a 3-D image is viewed from a distance L from the lens array (Fig. 6), the maximum frequency of stripes per radian is β, which is called the "viewing spatial frequency." With a 3-D image, therefore, α is affected by aberrations due to focusing errors of the elemental lenses in addition to pixel pitch and the diffraction limits of the elemental lenses. If the pixel pitch of the display device is p, the spatial frequency α_p of the elemental image projected by each elemental lens is expressed as

$$\alpha_p = |g|/2p, \quad (1)$$

where g denotes the distance from the lens array to the display device. Assume that the diffraction limit spatial frequency of the elemental lens is α_d and the limiting spatial frequency after considering aberrations of the elemental lens is α_e. The limiting spatial frequency of the elemental image projected by the elemental image in this case, i.e., α, is expressed as

$$\alpha = \min [\alpha_p, \alpha_d, \alpha_e]. \quad (2)$$

The elemental images projected at the thus-determined α are combined to generate the 3-D image.

The viewing spatial frequency β, which is the spatial frequency when the generated 3-D image is viewed from the position of the observer, is also affected by the depth-wise position of the 3-D image and is expressed as

$$\beta = \alpha (L - z)/|z|, \quad (3)$$

where L denotes the viewing distance (distance from the lens array to the viewer), z denotes the 3-D image distance (distance from the lens array to the 3-D image), the right side from the lens array takes positive values, and the left side takes negative values, as shown in Fig. 6.

When a 3-D image is displayed using the integral method, it is necessary to consider the pitch of the elemental lenses that make up the lens array. The distance from the display device that displays the elemental image by using the integral method to the lens array is arranged in such a manner as to approximately coincide with the focal distance of the elemental lenses, as shown in Fig. 7. In this state, parallel light rays are output from each elemental lens, the bundle of parallel light rays from each elemental lens are superimposed, and the 3-D image is generated, as shown in Fig. 8. As a result, when the 3-D image is viewed from the position of the viewer, the 3-D image is sampled at the pitch of the elemental lenses. From the resultant geometric optics, the maximum spatial frequency of the 3-D image is constrained to the Nyquist frequency determined by the pitch p_L of the elemental lenses, as expressed by

$$\beta_n = L/2p_L, \quad (4)$$

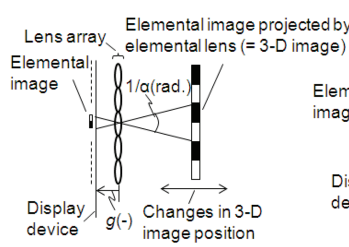


Figure 5. Projection spatial frequency.

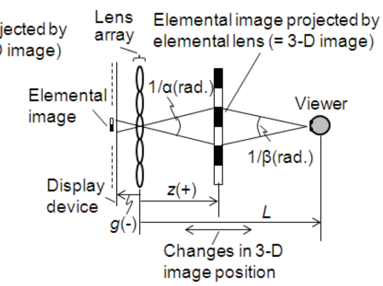


Figure 6. Viewing spatial frequency.

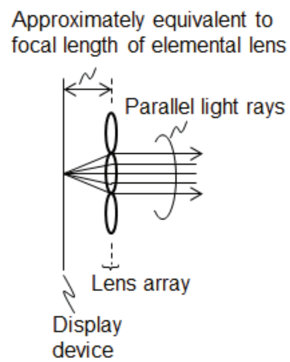


Figure 7. Output of parallel light rays.

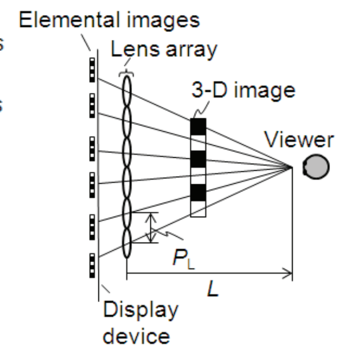


Figure 8. Maximum resolution.

where this spatial frequency β_n is called the "maximum spatial frequency."

From the above the limit of the spatial frequency for a 3-D image generated at any arbitrary depth is constrained to a low value in comparison with the maximum spatial frequency β_n determined by the elemental lens pitch and the viewing spatial frequency β calculated from Eq. (3), as expressed by

$$\gamma = \min [\beta_n, \beta], \quad (5)$$

where the spatial frequency γ expressed in Eq.(5) is called the "limiting spatial frequency [3][4]."

3.3. Viewing area of displayed image

A 3-D image that depends on the position of the viewer is visible with integral 3-D television, whether the viewer is to the right or left or above or below. It should be noted, however, that the range (viewing area) within which the viewer can move is limited to the region within which the light from a certain elemental image is output by the corresponding elemental lens. If the shape of the elemental image is circular, for example, the viewing area is conical.

Table 1. Specifications for integral 3-D television

Image capture equipment			Image display equipment		
Television camera	Pixel count	7680(H)x4320(V)x4 One each for red, blue, green 1, and green 2.	Projection device	Panel	7680(H)x4320(V)x3 One each for red, blue, and green.
				Time-division multiplexing	Complementary field-offset method
	Capture lens	Focal length 61.59mm		Projection lens	Focal length Approx. 60 mm
Lens array	Number of lenses	400(H)x250(V)		Projection size	Diagonal angle Approx. 26 in.
	Type	Gradient index lens / Concave lens	Lens array	Number of lenses	400(H)x250(V)
	Lens pitch	1.14mm / 1.35mm (Horizontal direction)		Type	Convex lens
				Lens pitch	1.44mm (Horizontal direction)
	Focal length	-2.65mm / -2.745mm		Focal length	2.745mm
	Arrangement	Delta array		Arrangement	Delta array
Converging lens	Focal length	800mm	Viewing angle	28 degrees (Designed value)	

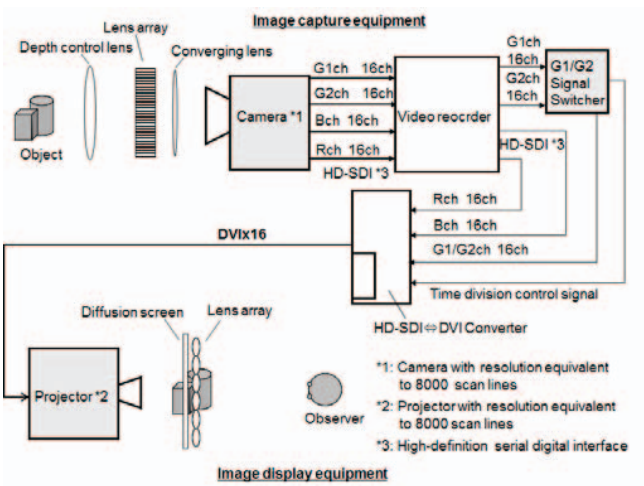


Figure 9. Configuration of an integral 3-D television.

The angle Ω denoted as the expanse of the viewing area is called the “viewing angle”, which is expressed as follows:

$$\Omega \fallingdotseq 2\tan^{-1}(p_L/2|g|). \quad (6)$$

4. EXPERIMENTAL SETUP

We built an integral 3-D television based on a video system with a resolution equivalent to 8000 scan lines by applying the pixel offset technique to 33-mega-pixel imaging devices [5]. The specifications for the integral 3-D television are listed in Table 1. The television camera uses one CMOS sensor for each of the red and blue signals, and two CMOS sensors (G1 and G2) for the green signals. The projection device uses time-division multiplexing of the G1 and G2 signals to display video with a resolution equivalent to 8000 scan lines. The configuration of an integral 3-D television is shown in Fig. 9. A 3-D image reconstructed by the display equipment is shown in Fig. 10. Figure 11 shows an 3-D image when viewed from different positions.

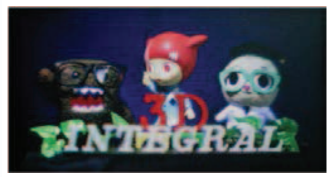
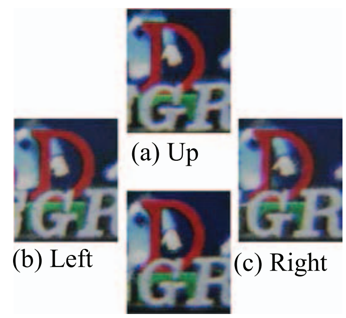


Figure 10. Reconstructed 3-D image.



(d) Down

Figure 11. Reconstructed 3-D images taken from different view positions.

5. CONCLUSION

Integral 3-D television can generate spatial images. I have reviewed the characteristics of 3-D images generated by an integral 3-D television. An experimental integral 3-D television with 33-mega-pixel imaging devices was also introduced. The experimental results confirm that 3-D video reproduction changes seamlessly with viewing position.

6. REFERENCES

[1] M. G. Lippmann, “Épreuves, réversibles donnant la sensation du relief,” J. de Phys, vol. 4, pp. 821–825, 1908.
[2] F. Okano, et al., “Three-dimensional video system based on integral photography,” Opt. Eng. Vol. 38, pp. 1072-1077, 1999.
[3] H. Hoshino, et al., “Analysis of resolution limitation of integral photography,” J. Opt. Soc. Am. A, vol. 15, pp. 2059–2065, 1998.
[4] J. Arai, et al., “Effects of focusing on the resolution characteristics of integral photography,” J. Opt. Soc. Am. A, vol. 20, pp. 996–1004, 2003.
[5] J. Arai, et al., “Integral imaging system with 33 mega-pixel imaging devices using the pixel-offset method,” Proc. SPIE, vol.8167, pp.81670X1-81670X9, 2011

ACKNOWREDGEMENTS

This work was supported in part by the National Institute of Information and Communications Technology.