Three-dimensional television based on spatial imaging

Fumio Okano
NHK Science and Technical research Laboratories

ABSTRACT

Research into three-dimensional imaging has attracted growing interest, with the progress of electronic technology in recent years. Our laboratory continues to study stereoscopic imaging, and its results have been applied to 3D-HDTV, further aiming to establish a technology for constructing spatial images as an advanced form of 3D television. This report discusses two main types for spatial imaging: integral method and holography.

1. INTRODUCTION

Expectations are mounting for the application of three-dimensional images in broadcasting, communications and many other areas, and yet practical applications have shown little progress for some reasons. One such reason is that almost 3D imaging systems fail to satisfy the following basic functions at the same time:

(1)Binocular disparity can be obtained without special glasses.

(2)The convergence point matches the eye's accommodation point.

(3) Moving parallaxes that allow an observer to see different images corresponding to different positions (horizontally, vertically).

These functions would allow observers to see a 3D image like a real object. With expanding research on spatial imaging, we aim to establish a 3D imaging technology that satisfies all these functions. In addition, a wider viewing zone and the ability to capture an object at an infinite distance are also necessary for television.

2. Classification and characteristics of 3D images

Many different 3D methods have been proposed, which can be classified into four major types: binocular, multi-view, volumetric, and spatial imaging. Table 1 shows their basic functions regarding binocular disparity, convergence, accommodation, and motion parallax. As shown in Fig.1, the binocular imaging system (stereoscopic) mainly employs binocular disparity and convergence. Images of an object are captured separately for the right and left eyes and shown to the observer separately, i.e., one image for the right eye and the other image for the left eye. A 3D optical image of the object will not be formed.

As the illustration shows, there is a horizontal offset between the right and left images. The observer's eyes rotate and the gaze moves so as to cancel this offset, and this function is called convergence. Where the eyes' gazes cross is the convergence point. The position where the eye is focused (accommodation point) is a little affected by this convergence, but an adjustment is made so that it roughly matches the display plate. As a result, the accommodation point differs from the convergence point, and this offset becomes larger when the binocular disparity is distinctive, causing visual fatigue for observers. To reduce this fatigue, the disparity should be controlled so that both display plate and convergence point are inside the eye's depth of field.

Compared with the binocular type, the multi-view type offers motion parallaxes. Even so, these two types are actually the same in that they both rely on convergence and binocular disparity to create stereoscopic effects. 3D optical images are not formed, and the points of convergence and accommodation do not match, so visual fatigue is not reduced.

Table1: Classification of 3D images

Type	Binocular disparity	Convergence	Accommo-dation	Motion parallax
Binocular	Yes	Yes	No	No
Multi-view	Yes	Yes	No	Yes
Volumetric	Yes	Yes	Yes	No
Spatial imaging	Yes	Yes	Yes	Yes

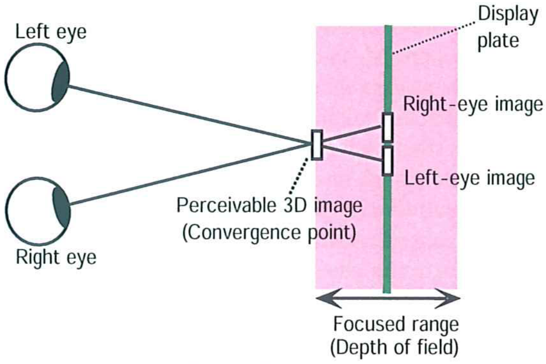

Fig.1: Binocular type

With the volumetric type, the display plate is positioned in the direction of depth to show images. In principle, the eye's accommodation works in this approach, though it is dependent on the pitch of display plates. This method, however, does not produce motion parallaxes, an important function by which the appearance of a 3D image changes depending on the observer's position. Special measures are needed for the imaging system to have this function.

The spatial imaging is a technique for reconstructing light itself from the object. It reproduces an image that looks just like a real object, allowing observers to experience visual 3D effects without special glasses. As this method satisfies all the three functions mentioned above, it should allow the development of an observer-friendly 3D imaging system that does not cause eye fatigue.

3. Integral method

The integral method, a kind of spatial imaging type, uses a lens array comprising tiny elemental lenses for both image pickup and display (see Fig.2). This method was first proposed in 1908 as a photographing technique[1]. It almost treats light as light rays to reconstruct a space using the same rays as those emitted by the object. The light wave itself cannot be reconstructed, but the integral method can reconstruct the light rays from the object. Since the human eye receives light mostly as energy, the observer can watch an image just like the object. As this shows, the method has the characteristics of spatial imaging.

A lens array placed in front of an object captures many small images that each correspond to the tiny elemental lenses on the array. During reproduction, the same lens array used in the capturing is placed before the display device to convey the signals from the image pickup stage to the display stage. As the light rays from the display device travel the same ray path, the same light rays as those from the object can be reconstructed. As a result, a 3D image of the object can be formed exactly where the object was once located. However, since the 3D images are psudoscopic with reversed depth in this basic process, any means should be introduced to avoid the problem.

The total number of pixels required for the image pickup device and the display device is given by the product of N_m and N_e, where N_m is the number of elemental lenses and N_e the number of pixels constituting an elemental image. Since the number of elemental lenses matches that of television pixels, the total number of pixels is equivalent to the number of elemental images multiplied by N_e, which is almost impractically large. To construct a practical system, the size of the elemental lens, and its arrangement, should be inconspicuous to the observer. It depends on the viewing distance, but is generally about the same size as the holes of the shadow mask in a cathode ray tube. Reconstructed images away from the display plate suffer a significant degradation in resolution due to the elemental lens's diffraction. To form 3D images over a wide range of depth, elemental lenses each measuring 0.1mm in diameter are too small; they should be 0.5 – 1.0mm in diameter. If 50 pixels are allotted to an elemental lens with that diameter, each pixel is then 10 – 20μm in size.

Recently, we have developed an experimental system for capturing and displaying moving images in real time based on an extremely high resolution video technology with 2,000 scanning lines[2]. Fig. 3 shows an experimental real-time integral method, called Integral Television. A depth control lens, a GRIN lens array, a converging lens, and an EHR (Extremely High Resolution) camera are

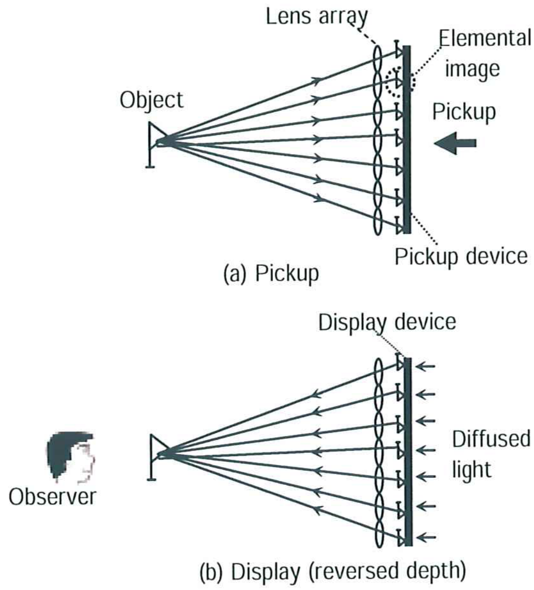

Fig. 2: Basic principle behind integral method

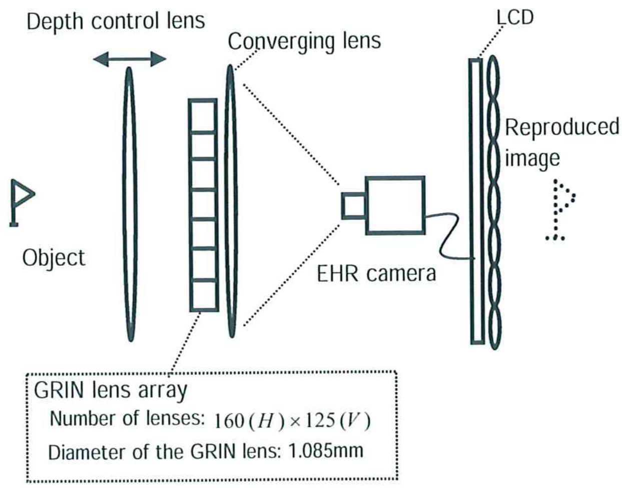

Fig.3: Integral Television

introduced for capturing. The depth control lens forms the image of the object around the lens array, and the image is captured by the GRIN lens array. Many elemental lenses, GRIN lenses, form elemental images near the output plate of the array. An elemental GRIN lens acts as a specific lens that forms an erect image for the far object to avoid pseudoscopic 3D images as mentioned above. The converging lens [3], which is set close to the GRIN lens array, leads the light rays from elemental GRIN lenses to the EHR camera. The converging lens is suitable for using the light rays efficiently. In principle, the camera is focused on the elemental images. In the actual cases, objects are located at different positions in the depth direction. The focusing should be set to avoid resolution degradation of 3D images in the different depth direction.

The video signal of the camera is directly led to a display device. The electronic display device combines a color LCD panel and an array of convex lenses. The lens array has the same delta structure as the GRIN lens array. The gap between the lens array and display device is about the same as the focal length of the elemental convex lens.

Fig. 4 shows a reproduced 3D image taken from the upper, the lower, the left and the right viewpoint, respectively. Different images are obtained according to the viewpoint. The reproduced 3D image is formed with full parallax in real time. Fig. 5(b) shows the images projected on the diffuser set in front of the display. The diffuser is set at the position of the letters I and P, so that these letters are clear and background objects are unclear compared with Fig. 5(a). It is evident that the integral imaging reconstructs optical 3D images like holography.

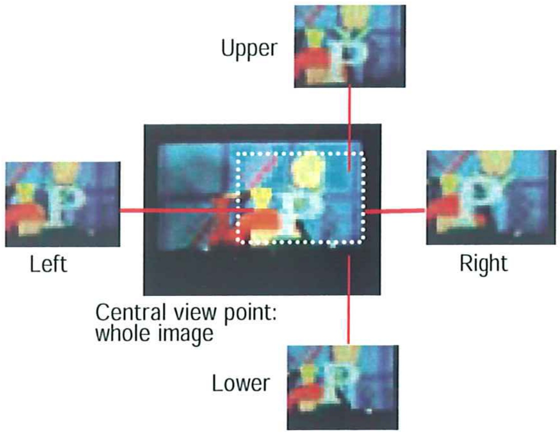

Fig.4: Different viewpoint images by the Integral Television

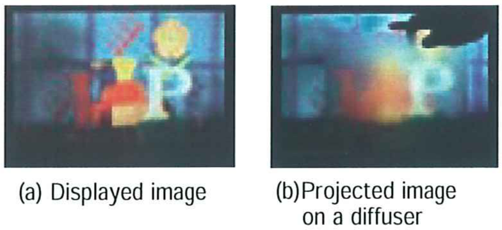

Fig.5: Reproduced image by the Integral Television

4. Holography

4.1 Holographic display

Holography, which records and reproduces interference fringes, is a well known spatial imaging technique. Holography is ideal as it records and reproduces the wave of light. As a means of electronically producing holographic moving images, it needs high-resolution liquid crystal display (LCD) panels and others devices.

The pixel pitch on an LCD needs to be 1μm or less in order to display fine interference fringes, but such narrow fringes are difficult to make with available technology. If the pixel pitch is not adequately fine, the reconstructed images will suffer from the narrow viewing zone, conjugate images and high-order components by diffraction. We proposed a method that ensures a wide viewing zone by eliminating conjugate images and through the effective use of the high-order components[4].

Fig. 6 shows an example of reproduced holographic images using LCD panel. The photos are taken from the left and the right viewpoint, respectively. The viewing zone is 6.5cm from the reproduced image. We can observe the image by two eyes.

Resolution is basically dependent on the size of the display plate if the pixel pitch is adequately fine. This is because resolution is affected by the diffraction of light that comes from the entire display plate. By the integral method, resolution is affected by the diffraction of light from the elemental lens. The display size in holography (from a few inches to a dozen inches) is significantly larger than the size of the elemental lens in the integral method (about 1mm). Because of this, it is generally accepted that holography offers higher resolution than the integral method for an image at a distance from the display plate.

4.2 Conversion from integral imaging to holography

Between the integral method and holography, the integral method appears closer to practical application. On the other hand, despite some fundamental obstacles, holography is also attractive for its ability to reconstruct light wave. The difficulty that uses coherent light must be overcome. Holographic display by photographs can use natural (incoherent) light to reconstruct images. During image pickup, on the other hand, it is necessary to use coherent light, which makes holographic applications in broadcasting extremely difficult. Capturing outside is almost impossible, and capturing of an object at a far distance is impractical.

To overcome these difficulties, we have developed a method that combines the advantages of the integral method and holography. The idea is to capture images by the integral method, which uses incoherent natural light, and then to convert the images into holograms for display. The conversion is performed by calculation by a computer[5].

Figure 7 shows images reconstructed by the hologram that is converted from the elemental images of the integral method. We can confirm the images of two objects which are located in the different depth.

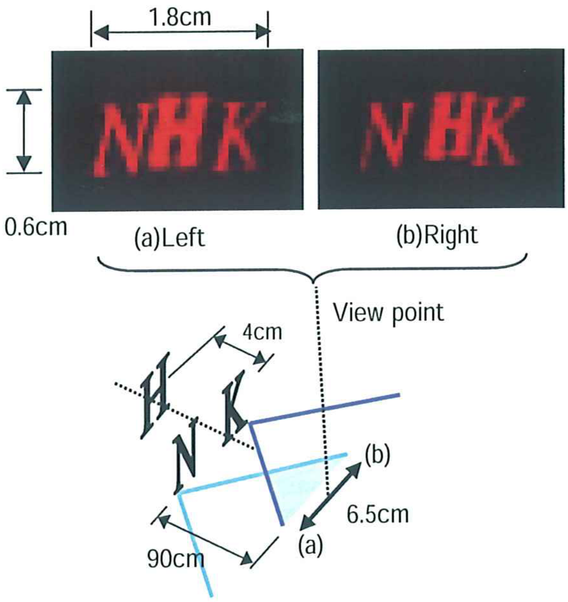

Fig.6: Computer generated holographic image displayed by an LCD. The image is obtained by use of both 0-order reconstructed component and the 1-order one, and eliminate conjugate components are eliminated.

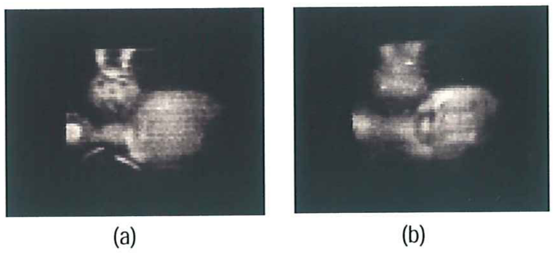

Fig.7: Holographic image converted from the elemental images of the integral method. These photos are focused on the rear object (a) and the front object (b), respectively.

5. Future research

This report outlined research at our laboratory for developing practical 3D television which is shifting from binocular imaging to spatial imaging. The experimental integral system we have made is not yet practical in resolution or viewing zone, and we still cannot confirm the matching of the eye's convergence point and accommodation point. Holography, another 3D imaging method, also poses difficulties including an extremely narrow viewing zone. At the moment, holographic 3D images are so primitive that an observer can barely recognize 3D effects.

On other hand, there has been dramatic improvement in high resolution imaging owing to Ultrahigh-definition ("super Hi-Vision") and other advanced technologies. By incorporating some of the elemental technologies of this extremely high resolution video system, we can expect to significantly improve 3D spatial imaging technology.

REFERENCES

[1] G. Lippmann, "Epreuves reversible donnant la sensation du relief," L. de Phys., Vol. 7, 4th series, pp. 821–825 (1908).

[2]F. Okano, J. Arai, K. Mitani, and M. Okui: " Real-time Integral Imaging based on Extremely High Resolution Video System," Proc.of IEEE, 94,3, pp.490- 501 (2006).

[3]F. Okano, J. Arai, H. Hoshino, and I. Yuyama: "Three-dimensional video system based on integral photography," Opt. Eng., 38 (6), 1072–1077(1999).

[4] T.Mishina, M.Okui, and F.Okano: "Viewing-zone Enlargement Method for Sampled Hologram that Uses High-order Diffraction," Appl.Opt. Vol.41, No8, pp.1489-1499 (2002)

[5] T.Mishina, M.Okui, and F.Okano :"Calculation of holograms from elemental images captured by integral photography," Applied Optics, Vol. 45, Issue 17, pp. 4026-4036 (June 2006)