Overview of Three-Dimensional Image Technology

Many digital broadcast receivers on the market today have three-dimensional (3D) functions to provide an environment for viewers to enjoy stereoscopic images in comfort at home. There is also a boom in 3D content, including movies. The 3D displays of today use a binocular stereoscopic method that requires the viewer to wear special glasses and can cause fatigue. We hope that the future holds another form of 3D TV that will enable us to view 3D images without special glasses and can accommodate a variety of viewing positions. At NHK, we are conducting research on three-dimensional imaging methods such as the integral method, holography, and multi-view images. In this paper, we give a broad overview of three-dimensional image technology and talk about the current state and prospects of research.

1. Conditions for natural viewing of 3D TV

Television is a ubiquitous and essential part of everyday life, but it still requires thought concerning of its viewability and safety. The two-viewpoint stereoscopic method has limitations in that wearing the glasses are a nuisance and it is not possible to see the 3D images while lying down. There are also problems with ocular fatigue, raising concern about the possible effects on people, particularly children, viewing 3D TV for long periods of time. For the 3D TV of the future, we must resolve these issues with the two-viewpoint stereoscopic method in order to create more natural 3D images. In particular, a future system should meet the following requirements:

(1) Binocular parallax must be achieved without glasses
(2) Motion parallax must be reproduced in the horizontal and vertical directions
(3)Binocular convergence and accommodation must match

From the viewpoint of broadcasting systems, the following two items are also requirements:

(4) Imaging must be possible under natural light
(5) Imaging of distant subjects must be possible

At NHK, we are conducting research into the integral method[1] and the holographic method[2]. In principle, the integral method satisfies conditions (1) to (5), but the holographic method does not satisfy condition (4). However, the holographic method can be made to satisfy all conditions by combining it with something like 3D imaging technology with the integral method.

2. Overview of stereoscopic image technology

Enabling someone to interpret a scene that he or she is seeing in a stereoscopic manner depends to a large extent on the characteristics of human vision. Below is a description of the cues that enable people to see and sense objects stereoscopically and the main stereoscopic imaging methods.

2.1 Stereoscopic vision cues

(1) Binocular convergence

When a person gazes at a target, the left and right eyeballs turn and the sight lines of the two eyes intersect on the target. This state is shown in Fig. 1(a). This intersection of the sight lines is called convergence, and the point at which they intersect is called the convergence point.

(2) Binocular parallax

When a person is gazing at the target A in Fig. 1(b), the positions of target B on the retinas of one eye will be different than on the other eye. This difference in the positions of the images observed by the left and right eyes is called binocular parallax. At a comparatively close distance, binocular parallax has the greatest effect in creating the stereoscopic effect.

(3) Accommodation

The crystalline lenses of the eyeballs change thickness in accordance with the distance to the target object to adjust the focus, as shown in Fig. 1(c). The behavior of the muscles that change this thickness and the blurriness of the optical image on the retina become cues to the stereoscopic effect. The position at which the image comes into focus is called the accommodation point.

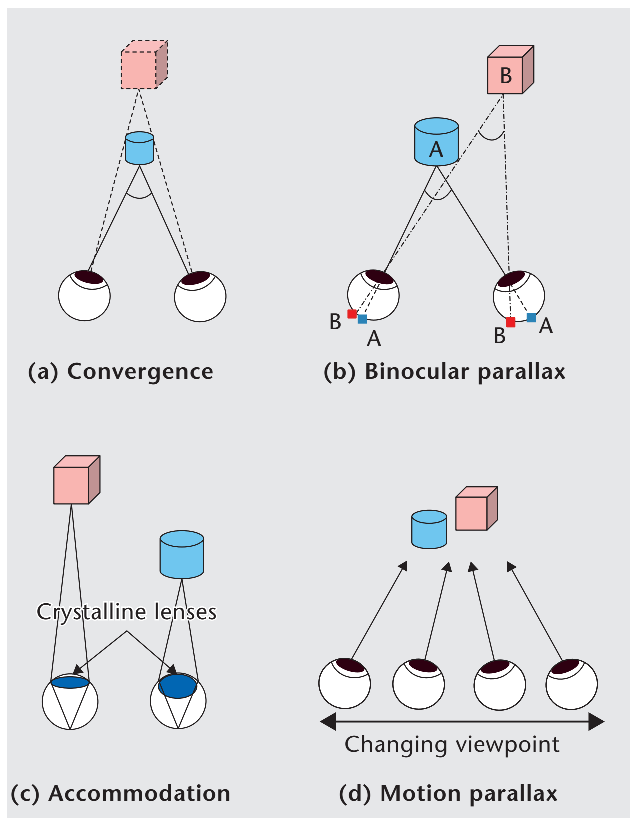

Figure 1: Stereoscopic view cues

(4) Motion parallax

The effect of seeing something that is hidden behind a target object or experiencing the effect of turning a target object that appears to be held in the hand by moving the position of the eyes, as shown in Fig. 1(d), is called motion parallax.

These four phenomena (1) to (4) are based on human visual functions. Other than those, it is known that there are psychological factors based on experience and memory, and stereoscopic effects can be created by using linear perspective, shading, overlapping among objects , and aerial perspective (hazily seen objects seem to be far away); these are methods of depiction used in paintings.

2.2 Various stereoscopic imaging methods

Stereoscopic imaging methods can be divided into the main types shown in Fig. 2. Of these, the spatial imaging methods include the integral method and holographic method as representative examples.

(1) Two-viewpoint method

The two-viewpoint method implements stereoscopic vision by reproducing binocular parallax and convergence. However, the focus of the eyes is accommodated with respect to the display screen, irrespective of the depth-wise position of the stereoscopic image. For that reason, the convergence and accommodation points do not match, which is said to be a cause of ocular fatigue. In addition, since only one pair of images is used, corresponding to the left and right eyes, motion parallax cannot be reproduced.

(2) Multi-viewpoint method

The multi-viewpoint method uses a lenticular lens[*1] or something similar to present the left and right eyes with multi-view images that are arrayed width-ways. The number of viewpoints is often on the order of eight. This enables the reproduction of motion parallax as the viewpoint moves in the horizontal direction.

(3) Integral method (spatial imaging method)

The integral method uses an imaging and display device consisting of a high-definition video system and a lens array to record and reproduce optical spatial images of subjects. Since a spatial image is reproduced, motion parallax in the horizontal and vertical directions can be reproduced. In addition, since the number of viewpoints for the integral method is far larger, e.g., 20 × 20 viewpoints, than the approximately eight viewpoints of the multi-viewpoint method, and conditions similar to those of viewing actual objects can be arranged, the convergence and adjustment points match.

(4) Holographic method (spatial imaging method)

Holography is a method of recording and reproducing the wavefronts of light, which can reproduce the ideal spatial image. During recording, laser beams illuminate the subject and an interference pattern formed by the light reflected from the subject (the object beam) and the original laser light (the reference beam) is recorded. The medium that records the interference pattern is called a hologram and the optical image is reproduced by irradiating the reference light on the hologram. Holograms can also be generated from three-dimensional information about the subject and images created by the integral method.

The various relationships between these methods and stereoscopic vision cues are listed in Table 1. The spatial imaging methods satisfy all of the stereoscopic

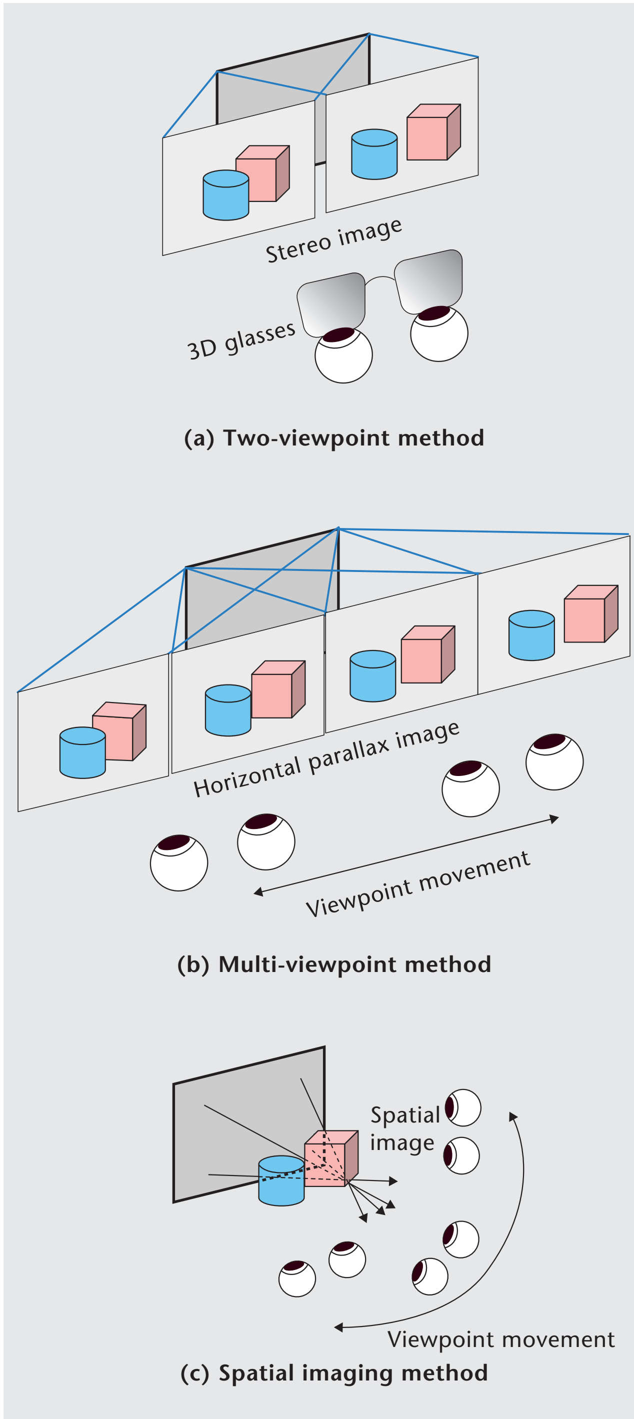

Figure 2: Different stereoscopic vision methods

[*1] A lens array consisting of a large number of long, hemi-spherical lenses arranged in a line.

Table 1: Stereoscopic vision methods and stereoscopic view cues

Stereoscopic vision method	Stereoscopic view cues			
	Binocular parallax	Convergence	Accommodation	Motion parallax
Two-viewpoint method	○	○	×	×
Multi-viewpoint method	○	○	×	△(horizontal only)
Spatial image reproduction method	○	○	○	○

vision cues listed in the table, and we consider they will enable reproduction of a natural stereoscopic image with little fatigue, provided that a video system of sufficient resolution can be devised.

3. Research projects aimed at application to broadcasting

Although we say that the spatial imaging methods are ideal in principle, they have a number of issues as broadcast applications. Note that the holographic method is based on the prerequisite of an imaging means such as the integral method, so we will confine the discussion to the issues faced by the integral method.

A schematic view of a 3D TV system based on the integral method is shown in Fig. 3. The imaging system consists of a gradient index lens array, a camera, a depth control lens, and a condenser lens. The depth control lens forms an optical image of the subject in the vicinity of the lens array. Each tiny lens of the gradient index

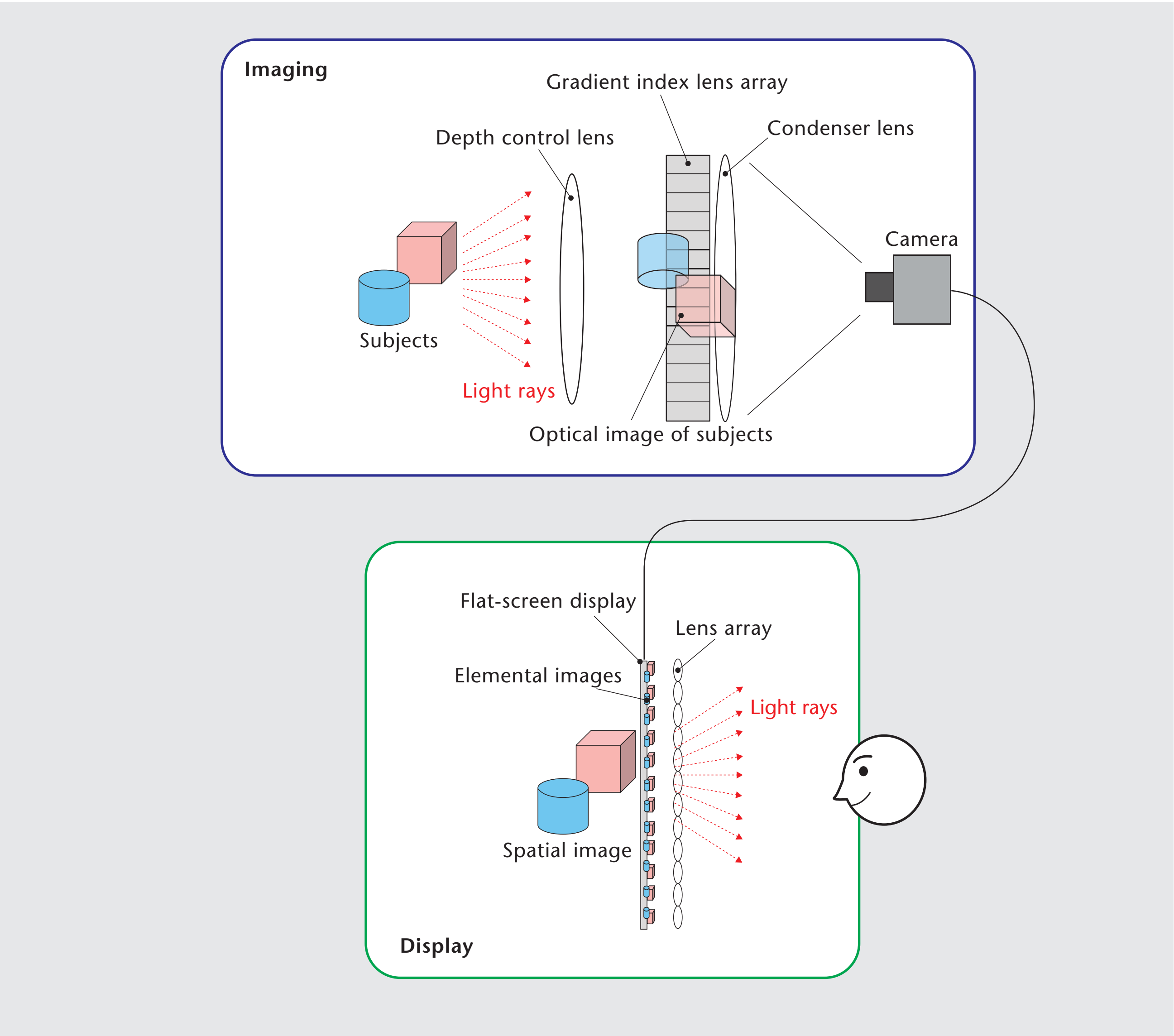

Figure 3: 3D TV system implemented by integral method

lens array forms an image of that optical image on its right-edge surface. The tiny images (called "elemental images") corresponding to each lens of the array can be combined and acquired by using the camera to image the end surfaces through the condenser lens. Bundles of light rays that spread in different directions from the subject are recorded in these elemental images. The display system consists of a flat-screen display with a lens array arranged over its front surface. The bundles of light rays are reproduced by displaying the group of elemental images on a flat-screen display, and thereby reconstructing a spatial image of the subject.

The limiting resolution of the spatial image reproduced by the integral method is determined by the number of lenses in the array. The main issue with this method, therefore, is that the resolution of the stereoscopic image that is reproduced in the depth-wise direction deteriorates as the number of pixels in the elemental images is reduced. In addition, to enable reproduction of a stereoscopic image in a wider region that can be viewed by a number of people side-by-side, the flat-screen display must have a large number of pixels. Thus, the quality of the stereoscopic image depends on the number of pixels in the flat-screen display. We have previously prototyped an integral 3D TV using 8,000-scanline video system[3]. To improve the quality even further, we should increase the resolution or the viewing area by combining a number of display devices, or by making a flat-screen display device with a higher resolution. It is also important to increase the number of pixels per element at the imaging stage, but we must also conduct research from the viewpoint of applying a limited number of pixels with maximum efficiency[4].

A further challenge is zooming. In ordinary TV broadcasts, it is common to zoom into a distant subject while imaging, but it is not easy to implement this effect with the integral method without losing the stereoscopic effect. If, for example, a zoom lens is substituted for the depth control lens in the configuration of Fig. 3, the device would become too large. To implement the effect of zoom, we have developed a method of generating a three-dimensional model of a subject from images taken by a number of cameras, then transforming it into elemental images[5].

Encoding methods also present a major research challenge. The Moving Picture Experts Group (MPEG) has been conducting standardization activities relating to the encoding of content such as multi-view images since 2001. In July 2013, an ad hoc group of specialists, called MPEG free-viewpoint TV (MPEG-FTV), was set up and is scheduled to investigate encoding methods aimed at content such as stereoscopic images with multiple viewpoints. NHK participates in this ad hoc group, and we plan to contribute to its standardization activities.

4. Looking forward

The integral method is a method of reproducing a spatial image of a subject, and we think it will satisfy the requirements for natural 3D TV. We started our research in 1995 and have improved quality of stereoscopic images by migrating the base video system from Hi-Vision to Super Hi-Vision. However, we cannot say that the quality is sufficient for actual usage, and it will be necessary to build on our research in the future. In the process, there will be many projects, such as on increasing the resolution of flat-screen displays, developing a practicable imaging system, and researching encoding methods. It is important to address these issues face-on, but we must also search out approaches that utilize the properties of human vision and sensations to lower the barriers to implementation. On the other hand, while holography is said to be the ideal method, it still requires basic research into numerous fields in order to make a practical system that has sufficient resolution and a large enough viewing region. It will also be necessary to enhance the three-dimensional image technology overall to make it suitable for program production[6] and to apply it to fields other than broadcasting[7]. We are taking a long-term approach to developing 3D TV that will be appreciated by viewers by envisioning the society of the future.

(Yuichi Iwadate)

References

1) F. Okano, J. Arai, H. Hoshino, and I. Yuyama: "Three-dimensional Video System Based on Integral Photography," Opt. Eng., 38, pp. 1072-1077 (1999)
2) NHK STRL R&D: Special Issue on Fundamental Holography Technology, No. 138 (2013)
3) J. Arai, M. Kawakita, T. Yamashita, H. Sasaki, M. Miura, H. Hiura, M. Okui, and F. Okano: "Integral Three - dimensional Television with Video System Using Pixel offset Method,"Optics Express, Vol.21, No.3, pp.3474-3485 (2013)
4) M. Miura, J. Arai, T. Mishina, M. Okui, and F. Okano: "Integral Imaging System with Enlarged Horizontal Viewing Angle,"Proc. SPIE, Vol.8384, pp.83840O.1-83840O.9 (2012)
5) K. Ikeya, K. Hisatomi, M. Katayama, and Y. Iwadate: "Depth Estimation from Three Cameras Using Belief Propagation - 3D Modelling of Sumo Wrestling,"Proc. the 8th European Conference on Visual Media Production (CVMP 2011), pp.118-125 (2011)
6) K. Ikeya, K. Hisatomi, M. Katayama, and Y. Iwadate: "Multi-Viewpoint Robotic Camera System," Proc. the 16th Meeting on Image Recognition and Understanding, OS4-5 (2013)
7) Y. Iwadate and M. Katayama: "Generating Integral Image from 3D Object by Using Oblique Projection," Proc. the 18th International Display Workshops (IDW '11), 3Dp-1, pp.269-272 (2011)