2 Three-dimensional images

We are continuing our research and development of more natural and viewable three-dimensional television that does not require special glasses, with the goal of developing a new form of broadcasting delivering a strong sense of presence.

The integral method, which uses an array of many tiny lenses for capture and display, requires imaging equipment with a large number of pixels to generate high-quality 3D images. We prototyped image capture and display devices in FY 2011 that were based on a video system with an equivalent of 8,000 scan lines. We have been working on the elemental technologies to achieve higher resolution and a wider viewing zone since FY 2012. In FY 2013, we built equipment to capture partitioned viewing zones with multiple cameras as well as compact equipment with the lens array adhering to the image sensor. We conducted a study on expanding the viewing zone by combining a number of small projectors and making a large screen from multiple flat displays.

Our study on the reproduction quality of integral 3D images started in FY 2012. In FY2013, we measured the ocular accommodation response to 3D images and confirmed that the accommodation conforms to the depth of the 3D images. We also conducted subjective evaluations using computer simulations on the relation between display parameters such as the pitch and focal length of lenses in the array and the resolution, and we obtained basic data relating the depth and resolution of 3D images.

Coding of the integral method has been included in a subject of study by the MPEG-Free-viewpoint Television (FTV) ad hoc group, which started its activities in July 2013. We subsequently began to study coding technologies for the integral method and decided to participate in activities of the ad hoc group.

We have been studying multi-viewpoint video technology since FY 2012 as a way of generating integral 3D images from 3D CG models of objects built from images taken by multiple cameras and researching its application to video production. To improve the accuracy of the 3D models, we built a camera array consisting of multiple cameras and infrared cameras in FY 2013. We also conducted experiments on generating integral 3D images from images captured by multi-viewpoint robotic cameras and improved the stability and speed of processing. This technology was successfully used in live sports broadcasting.

2.1 Three-dimensional images

2.1.1 Integral 3D television

We are researching a form of spatial imaging technology, called the integral method, as a way of creating a more natural and viewable 3D television that does not require special glasses. In FY 2013, we studied ways to improve the quality of 3D images and their reproduction.

Increasing 3D image quality

Generating high-quality 3D images using the integral method requires many more pixels than for ordinary 2D imaging, both for capture and display. We studied ways to make 3D images with more pixels by using multiple imaging devices. For capture, we used two sets of an image sensor and a lens array. This structure, however, suffers from a loss of information and a difference in luminance level between image sensors. To solve this problem, we interpolated the missing information and reduced the luminance level difference. Thus, we were able to use a large number of pixels, two image sensors worth, to make 3D images (1). For the display, we prototyped equipment consisting of four projectors (Figure 1). We found that we could expand the viewing zone (the areas in which the 3D image can be viewed) by a factor of 2.5 in the horizontal direction and 2.0 in the vertical direction, by using projective and affine transformations to correct the distortion in the images from each projector (2).

Part of this research was conducted under contract with the Ministry of Internal Affairs and Communications for its project titled, “R&D on spatial information acquisition system using
multiple image sensors.”

**Reproduction quality of 3D images**

We have been studying the quality of 3D images reproduced using the integral method since FY 2012. The integral method reproduces the light rays from an object and forms the equivalent 3D image in the air near the viewer. It is expected to have the advantage of not straining the viewers’ eyes, because the ocular accommodation of viewing such a 3D image is said to be at the same depth as when one observes the actual object.

In FY 2012, we measured the ocular accommodation response when viewing integral 3D images with both eyes in order to verify the relation between the accommodation position and the image’s depth. The results showed that the accommodation conforms to the depth of the 3D image. Meanwhile, the accommodation is directed to the point of regard by moving the right and left eyeballs. To eliminate this factor, we measured the accommodation response in viewing integral 3D images with a single eye in FY 2013. The results confirmed that the accommodation has a tendency to conform to the image’s depth position even with monocular vision (3).

To evaluate the reproduction quality of 3D images against the display parameters of the integral method, we conducted subjective evaluations on image quality versus the depth of the image by using the pixel pitch of an image displayed by a lens array as a parameter. The experiment provided us with enough data to derive the relation between the pixel pitch and the reproduction range of the depth. We plan to conduct evaluations with a wider range of parameters to further study the relationship between pixel pitch and the reproduction range of the depth.

We began a study on coding technologies for the integral method in FY 2013 and decided to participate in the standardization activities of the MPEG-FTV ad hoc group, which studies the integral method’s coding and other subjects.

**References**


---

**2.1.2 Generating 3D content from multi-viewpoint images**

We are studying ways to acquire integral 3D images of objects that are difficult to capture with optical equipment, such as very distant or very large objects.

In FY 2012, we studied how to capture objects with infrared patterns projected on them using a camera array consisting of two infrared cameras and two color cameras (Figure 1) and estimate the depth information. In FY 2013, we improved the stability and accuracy of the depth estimation by extending the two-dimensional smoothing filter it uses to the temporal dimension to make the three-dimensional spatial-temporal smoothing filter (4). We also studied how to generate 3D models by integrating depth information gotten from multiple camera arrays.

We continued our work from FY 2011 on generating 3D models of objects from multi-viewpoint images. In FY 2013, we developed a new method to improve the accuracy of the 3D models by evaluating the reliability of the depth estimations obtained from the multi-viewpoint images and combining only the areas of the images that are free of estimation errors (5). This method reduced the depth estimation errors and enabled the generated integral 3D images to appear to show camera work such as zooming and panning. The multi-viewpoint images in this study were captured by a multi-viewpoint robotic camera (6).

Regarding this robotic camera, we implemented a “weak” camera calibration method that does not require special patterns in FY 2012. In FY 2013, we improved the accuracy of this camera calibration. Time slice video (appearing as if one camera had captured footage of the object while rotating around it) captured by the multi-viewpoint robotic cameras was used during the production of the “Figure Skating 2013, NHK Trophy” program.

**References**

