FTV (Free viewpoint TV) and Creation of Ray-Based Image Engineering

Masayuki Tanimoto [1], Non-member

ABSTRACT

Free viewpoint TV (FTV) enables us to view a distant 3D world by freely changing our viewpoints as if we were there. FTV will bring an epochal change in the history of television since this function has not yet been achieved by conventional TV technology. We propose the concept of FTV and verified its feasibility using the world's first real-time system including the complete chain of operation from image capture to display. The international standardization of FTV technologies is now underway. FTV is a ray-based system rather than a pixel-based system. We are creating ray-based image engineering through the development of FTV.

1. INTRODUCTION

Television realized the human dream of seeing a distant world in real time and has served as the most important visual information technology to date. Now, television can provide us scenes overcoming distance not only in space but also in time by introducing storage devices into television. However, TV shows us the same scene even if we change our viewpoint in front of the display. This is quite different from what we experience in the real world. With TV, users can get only a single view of a 3D world. The view is determined not by users but by a camera placed in the 3D world. Although many important technologies have been developed, this function of TV has never changed.

We have developed a new type of television named FTV (Free viewpoint TV) [1]-[7]. FTV is an innovative visual media that enables us to view a 3D scene by freely changing our viewpoints as if we were there. FTV will bring an epochal change in the history of visual media since such a function has not yet been achieved by conventional media technology. FTV is based on the ray-space method [8] - [11].

The most essential element of visual systems is not the pixel but the ray. FTV is not a conventional pixel-based system but a ray-based system. We have been developing ray capture, processing, and display technologies for FTV [4].

We proposed the concept of FTV and verified its feasibility with the world's first real-time experiment [12], in which a real, moving object was viewed using a virtual bird's-eye camera whose position was moved freely in a space and was controlled by a 3D mouse.

FTV opens a new frontier in the field of signal processing since its process such as coding and image generation is performed in the ray-space, a new domain with higher dimensionality than possible with conventional TV. A new user interface is also needed for FTV to make full use of 3D information.

As the ultimate 3D-TV, FTV needs higher performance in all related devices, equipment and systems, as well as accelerated development of the electronics industry. As the next-generation TV, it will find applications in the fields of communication, entertainment, advertising, education, medicine, art, archives and so on. As part of the social information infrastructure, it can increase the security of public facilities, roads, vehicles, schools, factories, etc.

The FTV Committee has been organized to promote the international standardization of FTV in the Japan Electronics and Information Technology Industries Association (JEITA) under the support of the Ministry of Economy, Trade and Industry (METI). About 20 organizations participate in this activity, including members of industry, carriers, broadcasters, and content providers. We proposed FTV to MPEG (Moving Picture Experts Group) [13] and have been making contributions to its development. It is considered the most challenging scenario in the 3DAV (3D Audio Visual) work of MPEG, and the standardization of MVC (Multi-view Video Coding) has started [14].

MPEG and ITU-T (International Telecommunication Union Telecommunication Standardization Sector) have started the standardization of the entire FTV [15] because the FTV system cannot be achieved by multi-view video coding alone. This activity is supported by the SCOPE-I program of the Ministry of Internal Affairs and Communications (MIC).

In this paper, the principle, basic technologies and advanced ray technologies of FTV are presented. The progress of international standardization of FTV is also described.

2. RAY-SPACE REPRESENTATION

We developed FTV based on ray-space representation [8] - [11]. In ray-space representation, one ray

Manuscript received on January 20, 2008 ; revised on January 20, 2008.

[1] The author is with Graduate School of Engineering, Nagoya University Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan, E-mail: tanimototentative3@yahoo.co.jp and tanimoto@nuee.nagoya-u.ac.jp

in the 3D real space is represented by one point in the ray space. The ray space is a virtual space. However, it is directly connected to the real space. The ray space is generated easily by collecting multi-view images while giving consideration to the camera parameters.

Let (x, y, z) be three space coordinates and θ, ϕ be the parameters of direction. A ray going through space can be uniquely parameterized by its location (x, y, z) and its direction (θ, ϕ); in other words, a ray can be mapped to a point in this 5D, ray-parameter space. In this ray-parameter space, we introduce the function f, whose value corresponds to the intensity of a specified ray. Thus, all the intensity data of rays can be expressed by

$$f(x, y, z; \theta, \phi)$$
$$-\pi \leq \theta < \pi, -\pi/2 \leq \phi < \pi/2 \quad (1)$$

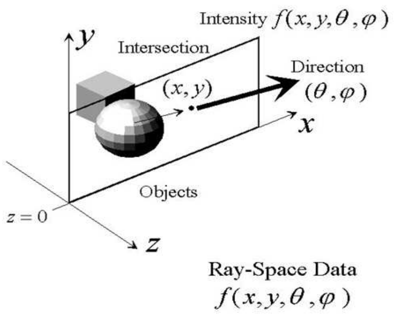


Fig.1: *Definition of orthogonal ray-space*

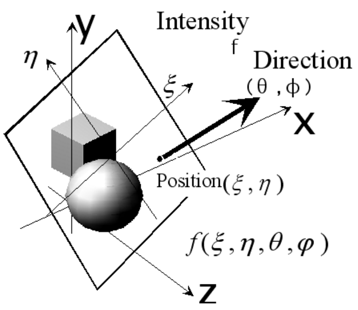


Fig.2: *Definition of spherical ray-space*

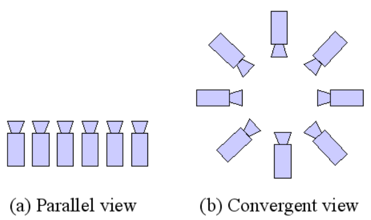


Fig.3: *Camera Arrangements for FTV*

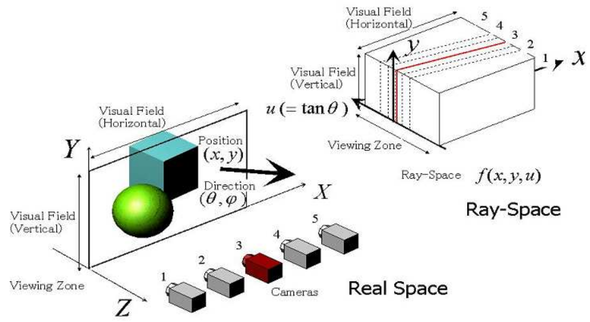


Fig.4: *Acquisition of orthogonal ray-space*

Although the 5D ray-space mentioned above includes all information viewed from any viewpoint, it is highly redundant due to the straight traveling paths of the rays. Thus, when we treat rays that arrive at a reference plane, we can reduce the dimension of the parameter space to 4D.

We use two types of ray-space for FTV. One is the orthogonal ray-space, where a ray is expressed by the intersection of the ray and the reference plane and the ray's direction as shown in Fig. 1. Another is the spherical ray-space, where the reference plane is set to be normal to the ray as shown in Fig. 2. The orthogonal ray-space is used for FTV with a linear camera arrangement, whereas the spherical ray-space is used for FTV with a circular camera arrangement. The linear camera arrangement is used for parallel view and the circular camera arrangement is used for convergent view as shown in Fig. 3.

Both the orthogonal ray-space and the spherical ray-space are 4D and 5D, including time. If we place cameras within a limited region, the obtained rays are limited, and the ray-space constructed from these rays is a subspace of the ray-space. For example, if we place cameras in a line or in a circle, we have only one part of the data of the whole ray-space. In such cases, we define a smaller ray-space.

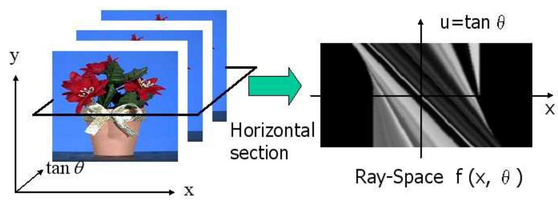


Fig.5: *Example of orthogonal ray-space*

For the linear camera arrangement, the ray-space is constructed by placing many camera images upright and parallel, as shown in Fig. 4, forming the FTV signal in this case. The FTV signal consists of many camera images, and the horizontal cross-section has a line structure as shown in Fig. 5. The line structure of the ray-space is used for ray-space interpolation and compression. Vertical cross-sections of

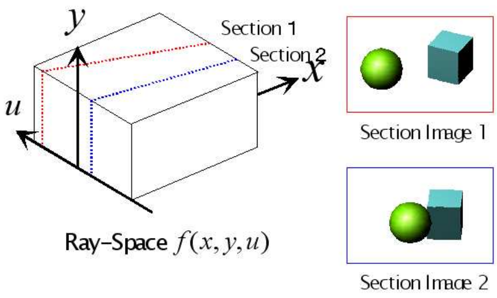


Fig.6: *Generation of view images*

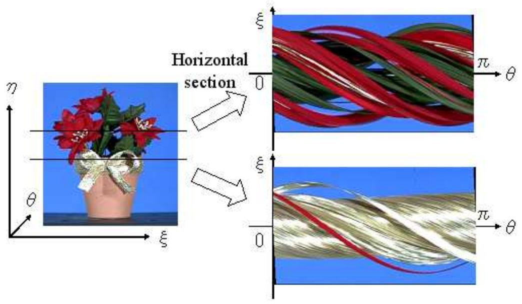


Fig.7: *Example of spherical ray-space*

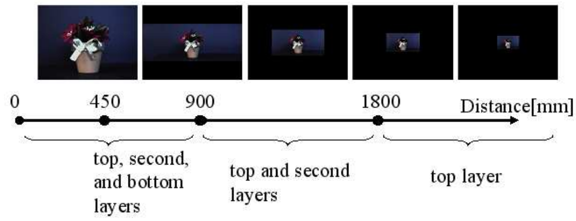


Fig.9: *Free viewpoint image generation at various distances using hierarchical ray-space*

the ray-space give view images at the corresponding viewpoints as shown in Fig. 6.

For the circular camera arrangement, the spherical ray-space is constructed from many camera images, and its horizontal cross-section has a sinusoidal structure as shown in Fig. 7. The sinusoidal structure of the ray-space is also used for ray-space interpolation and compression.

The hierarchical ray-space [10] is defined for scalable expression of 3D scene. Fig. 8 shows the concept of the hierarchical ray-space. Fig. 9 shows free viewpoint image generation at various distances using the hierarchical ray-space.

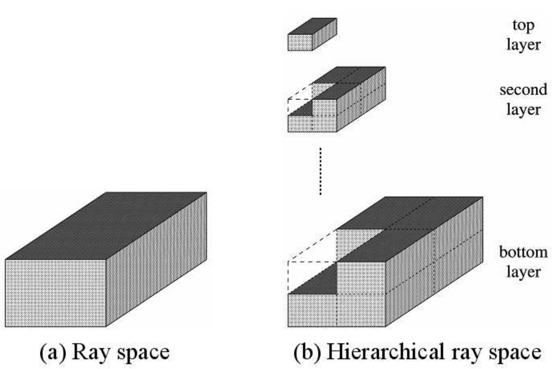


Fig.8: *Concept of hierarchical ray-space*

3. FTV SYSTEM

3.1 Configuration of FTV System

Fig. 10 shows the configuration of the FTV system. At the sender side, a 3D scene is captured by multiple cameras. The captured images contain the misalignment and luminance differences of the cameras. They must be corrected to construct the ray-space. The corrected images are compressed for transmission and storage by the MVC (Multi-view Video Coding) encoder.

At the receiver side, reconstructed images are obtained by the MVC decoder. The ray-space is constructed by arranging the reconstructed images and interpolating them. Free-viewpoint images are generated by cutting the ray-space vertically and are displayed on a 2D/3D display.

The function of FTV was successfully demonstrated by generating photo-realistic, free-viewpoint images of the moving scene in real time. Each part of the process shown in Fig. 10 is explained in greater detail below.

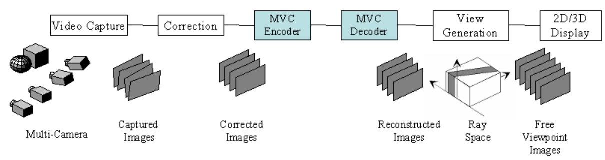


Fig.10: *Configuration of FTV system*

3.2 Capture

We constructed a 1D-arc capturing system shown in Fig. 11 for a real-time FTV system covering the complete chain of operation from video capture to display [16]. It consists of 16 cameras, 16 clients and 1 server. Each client has one camera and all clients are connected to the server with Gigabit Ethernet.

A "100-camera system" has been developed to capture larger space by Nagoya University (Intelligent Media Integration COE and Tanimoto Laboratory) [17]. The system consists of one host-server PC and 100 client PCs (called 'nodes') that are equipped with JAI PULNiX TM-1400CL cameras. The interface between camera and PC is Camera-Link. The host PC generates a synchronization signal and distributes it to all of the nodes. This system is capable of capturing not only high-resolution video with 30 fps but also analog signals of up to 96 kHz. The specification of the 100-camera system is listed in table 1.

The camera setting is flexible as shown in Fig. 12. MPEG test sequences "Rena" and "Akko & Kayo" shown in Fig. 13 were taken by camera arrangements (a) and (c), respectively.

Fig.11: *1D-arc capturing system*

Table 1: *Specification of 100-camera system*

Image resolution	1392(H)x1040(V)
Frame rate	29.4118 [fps]
Color	Bayer matrix
Synchronization	Less than 1 [us]
Sampling rate of A/D	96 [kS/s] maximum
Maximum number of nodes	No limit. (128 max for one sync output)

(a) linear arrangement

(b) circular arrangement

(c) 2D-array arrangement

Fig.12: *100-camera system*

3.3 Correction

The geometric correction [18], [19] and color correction [20] of multi-camera images are performed by measuring the correspondence points of images. This measurement is made once the cameras are set.

(a) "Rena" (b) "Akko & Kayo"

Fig.13: *MPEG test sequences*

An example of the geometric correction is shown in Fig. 14 [19]. Here, the geometric distortion of a 2 dimensional camera array is corrected by the affine transform. It is seen that the trajectory of correspondence point becomes square by the geometric correction.

An example of color correction is shown in Fig. 15.

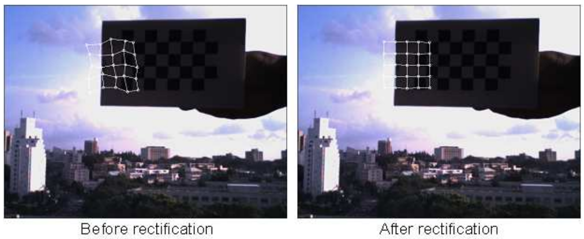


Fig.14: *An example of geometric correction*

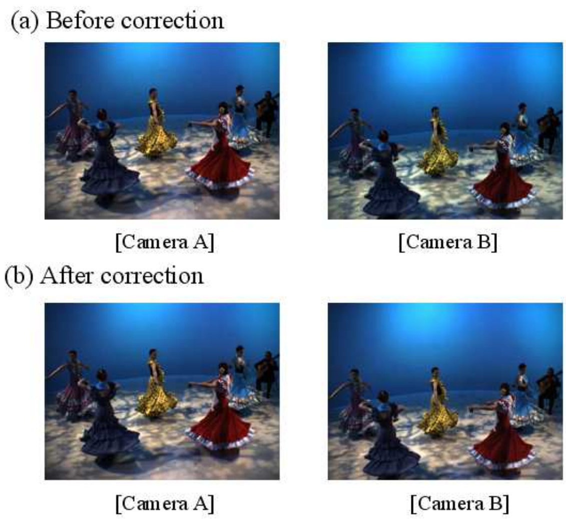


Fig.15: *An example of color correction*

3.4 MVC Encoding and Decoding

An example of time and view variations of multi-view images is shown in Fig. 16. They have high temporal and interview correlations. MVC (Multi-view Video Coding) reduces these correlations [14], [21], [22]. The standardization of multi-camera image compression is progressing with MVC (Multi-view Video Coding) in MPEG. Details are described in Section 5.

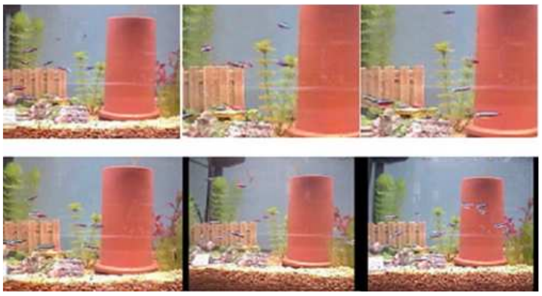


Fig.16: *Time and view variations of multi-view images*

3.5 View Generation

Ray-space is formed by placing the reconstructed images vertically and interpolating them. Free-viewpoint images are generated by making a cross-section of the ray-space.

Examples of the generated free-viewpoint images are shown in Fig. 17. Complicated natural scenes, including sophisticated objects such as small moving fish, bubbles and reflections of light from aquarium glass, are reproduced very well.

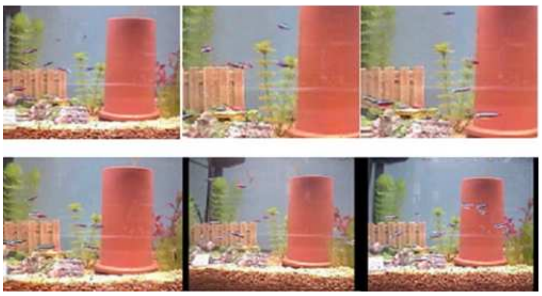


Fig.17: *An example of generated FTV images at various times and viewpoints*

The quality of the generated view images depends on the view interpolation. The ray-space interpolation is achieved by detecting depth information pixel by pixel from the multi-view video. We proposed several interpolation schemes of the ray-space [23]-[28].

The free-viewpoint images were generated by a PC cluster in [16]. Now, they can be generated by a single PC, and FTV on a PC can be accomplished in real time [27].

3.6 2D/3D Display

FTV needs a new user interface to display free-viewpoint images. Two types of display, 3D display and 2D/3D display with a viewpoint controller, are used for FTV as shown in Fig. 18.

Viewpoint control by head-tracking is shown here. Many head-tracking systems have been proposed using magnetic sensors, various optical markers, infrared cameras, retroreflective light from retinas, etc. Our head-tracking system uses only a conventional 2D camera and detects the position of a user's head by image processing. The user doesn't need to attach any markers or sensors.

In the user interface using a 2D display, the location of the user's head is detected with the head-tracking system and the corresponding view image is generated. Then, it is displayed on the 2D display as shown in Fig. 19.

Automultiscopic displays enable a user to see stereoscopic images without special glasses. However, there are two problems: a limited viewing zone and discreteness of motion parallax. Because the width of the viewing zone for each view approximates the interpupillary distance, the view image does not change with the viewer's movement within the zone. On the other hand, when the viewer moves over the zone, the view image changes suddenly.

In the user interface using the automultiscopic display, the function of providing motion parallax is extended by using the head-tracking system. The images fed to the system change according to the movement of the head position to provide small motion parallax, and the view channel for feeding the images is switched for handling large motion. This means that binocular parallax for the eyes is provided by automultiscopic display, while motion parallax is provided by head tracking and changing the image adaptively as shown in Fig. 20.

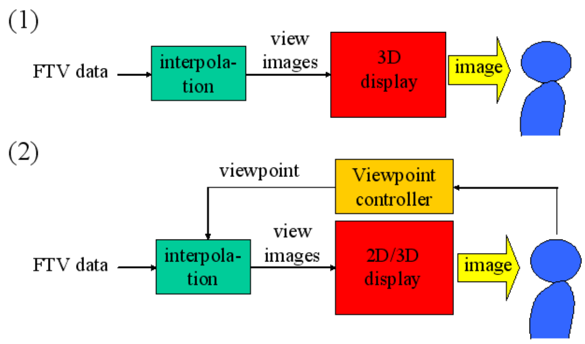


Fig.18: *Display of FTV*

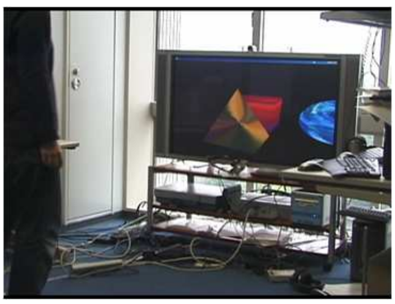


Fig.19: *2D display with eye tracking*

(a)without head tracking (b)with head tracking

Fig.20: *3D display with and without head tracking*

4. CREATION OF RAY-BASED IMAGE ENGINEERING

4.1 Evolution of Image Systems

Fig. 21 shows the evolution of image systems. In the past, image systems such as photography, film and TV were individual systems. At present, they are digitized and can be treated on the same platform as pixel-based systems. These pixel-based systems are developing toward using more pixels. This trend is exemplified by Super High-Definition TV [29]. Although Super HDTV has about 100 times the pixels of SDTV, there is still only one view.

In the future, the demand for more pixels will be saturated, and more views will be needed. This will result in the evolution from a pixel-based system to a ray-based system. We have been developing FTV according to this scenario. Roughly speaking, we can achieve SD-FTV by using the technologies of HDTV or Super HDTV and balancing the number of pixels and views.

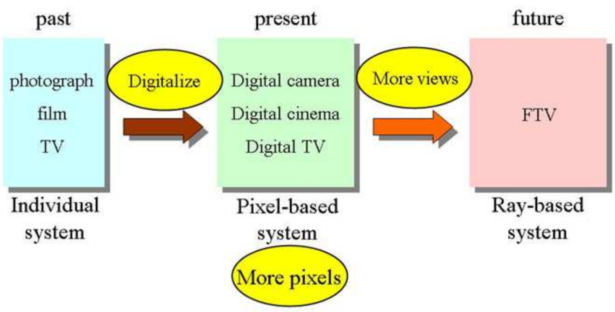


Fig.21: *Evolution of image systems*

4.2 Ray Capture and Display

We are developing ray-reproducing FTV to create ray-based image engineering. Ray-reproducing FTV consists of ray capture, ray processing and ray display.

We developed a ray capturing system [30] shown in Fig. 22. It acquires a dense ray-space without interpolation in real time. In this capturing system, a high-speed camera and a scanning optical system are used instead of multiple cameras. The important feature of this configuration is that the spatial density of a multi-camera setup is converted to a time-density axis, i.e. the frame rate of the camera. This means that we can increase the density of the camera interval equivalently by increasing the frame rate of the camera. The scanning optical system is composed of a double-hyperbolic mirror shell and a galvanometric mirror. The mirror shell produces a real image of an object that is placed at the bottom of the shell. The galvanometric mirror in the real image reflects the image in the camera-axis direction. The reflected image observed from the camera varies according to the angle of the galvanometric mirror. This means that the camera can capture the object from various viewing directions that are determined by the angle of the galvanometric mirror. To capture the time-varying reflected images, we use a high-speed camera with an electronic shutter that is synchronized with the angle of the galvanometric mirror. We capture more than 100 view images within the reciprocation time of the galvanometric mirror. The collection of the view images is then mapped to the ray-space.

However, this system can capture the spherical ray-space with the viewing zone of only 55 degrees.

Then, we have developed a 360-degree ray capturing system as shown in Fig. 23 [31], [32]. This system uses two parabolic mirrors. Incident rays that are parallel to the axis of a parabolic mirror gather at the focus of the parabolic mirror. Hence, rays that come out of an object placed at the focus of the lower parabolic mirror gather at the focus of the upper parabolic mirror. Then, the real image of the object is generated at the focus of the upper parabolic mirror and a rotating aslope mirror scans rays at the focus of the upper parabolic mirror. Finally, the image from the aslope mirror is captured by a high-speed camera. By using this system, we can capture all-around convergent views of an object as shown in Fig. 24.

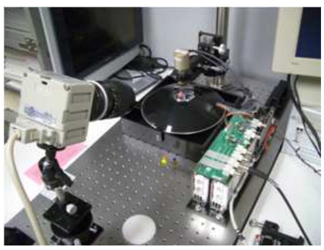


Fig.22: *Mirror-scan ray capturing system*

Fig. 25 shows SeeLINDER [33], a 360-degree, ray-producing display that allows multiple viewers to see 3D FTV images. Structure of the SeeLinder is shown in Fig. 26. It consists of a cylindrical parallax barrier and one-dimensional light-source arrays. Semiconductor light sources such as LEDs are aligned vertically for the one-dimensional light-source arrays. The cylindrical parallax barrier rotates quickly, and the light-source arrays rotate slowly in the opposite direction. If the aperture width of the parallax barrier

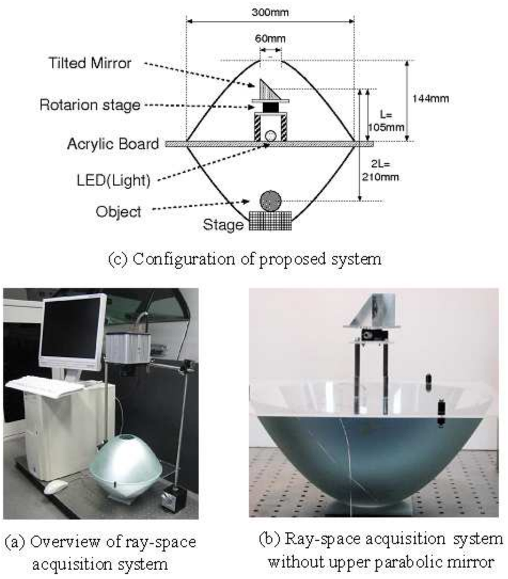


Fig.23: *360-degree mirror-scan ray capturing system*

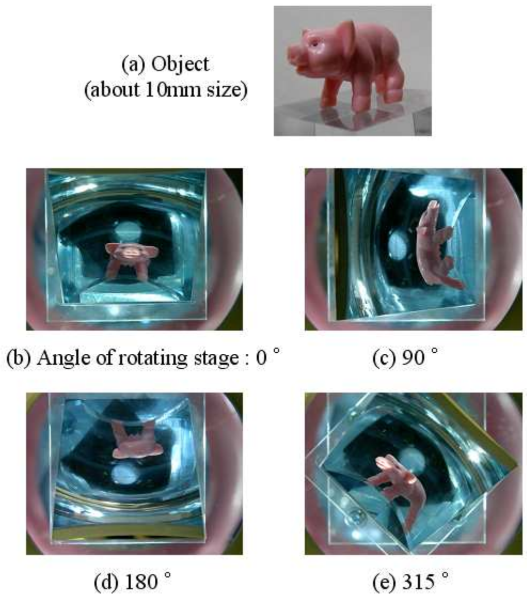


Fig.24: *Object and captured images of 360-degree ray capturing system*

is sufficiently small, the light going through the aperture becomes a thin flux, and its direction is scanned by the movement of the parallax barrier and the light-source arrays. By synchronously changing the intensity of the light sources with the scanning, pixels whose luminance differs for each viewing direction can be displayed. We can see the 3D image naturally, and the images have the strong depth cues of natural binocular disparity. When we move around the display, the image changes corresponding to our viewing position. Therefore, we perceive the objects just as if they were floating in the cylinder.

We are going to connect these two systems directly in real time.

Fig.25: *The SeeLINDER, a 360 degree, ray-reproducing display*

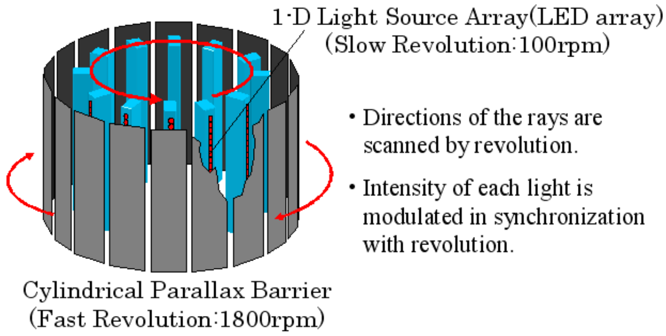


Fig.26: *Structure of the SeeLinder*

Fig. 27 shows the progress made in 3D capturing and display systems. In this figure, the ability of 3D capturing and display is expressed as a factor of the pixel-view product, defined as (number of pixels) ×(number of views). In Fig. 27, (1) denotes the factor of the 100-camera system mentioned earlier. We have also been developing new types of ray-capturing and display systems. Their factors are indicated by (2) and (3); (2) is a mirror-scan ray-capturing system [30] and (3)is the 360-degree, ray-reproducing SeeLINDER display [33].

In Fig. 27, the progress of space-multiplexing displays follows Moore's Law because it is achieved by miniaturization. The factor of the time-multiplexing display is larger than that of the space-multiplexing display. The difference is a result of time-multiplexing technology. The progress of capturing does not follow Moore's Law because it depends on camera resolution and the number of cameras used. Furthermore, the pixel-view product has increased very rapidly year after year in both capture and display. This development strongly supports our scenario.

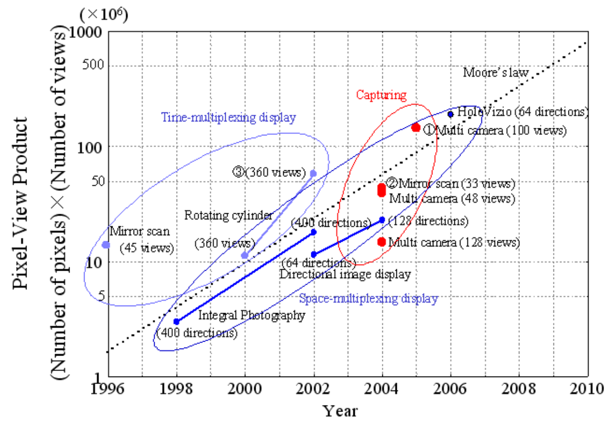


Fig.27: *Progress in increasing pixel-view product for 3D capturing and display*

4.3 Ray Processing

Typical example of orthogonal ray-space with a horizontal cross-section is shown in Fig. 28. The horizontal cross-section has a line structure. The slope of the line corresponds to the depth of object.

The ray-space is a platform of ray processing. Various kinds of signal processing can be done in the ray-space. Vertical cross-sections of the ray-space give real view images at the corresponding viewpoints. Manipulation, division and composition of 3D Scenes are also performed by ray-space processing [26].

Fig. 29 shows an example of the ray-space processing. Bars in the scene of Fig. 29 (a) are eliminated in Fig. 29 (b) by applying non-linear filtering to the ray-space [34]. Composition of 2 scenes shown in Fig. 30 is performed by ray-space processing as shown in Fig. 31 [35].

Images with optical effects are generated by cutting the ray-space with a curved plane as shown in Fig. 32. The shape of the curved plane is determined due to an optical effect to be realized. Artistic images shown in Fig. 33 are generated by cutting the ray-space with more general planes [36], [37].

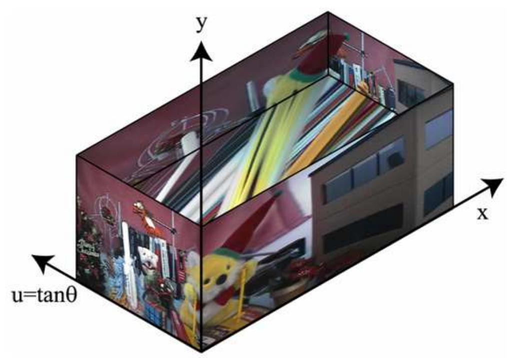
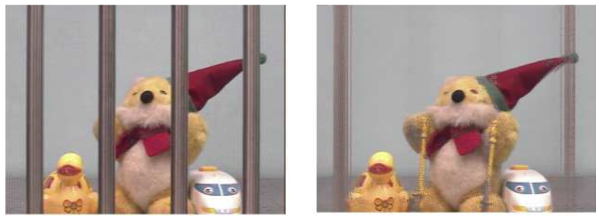


Fig.28: *Typical example of orthogonal ray-space and a horizontal cross-section*



(a)Before filtering (b)After filtering

Fig.29: *An example of ray-space processing: object elimination by non-linear filtering.*

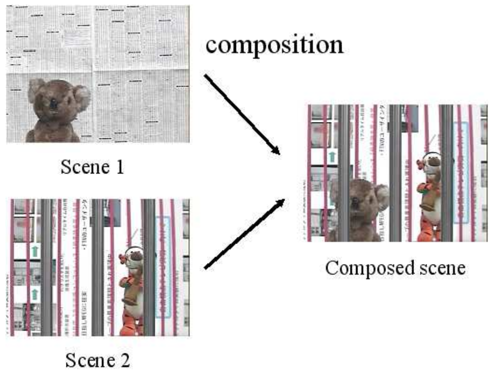


Fig.30: *Scene composition by ray-space processing*

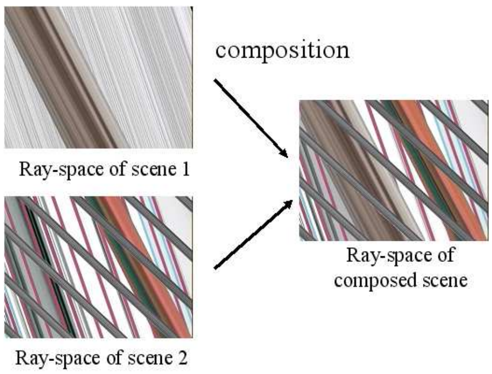


Fig.31: *Ray-space processing for scene composition*

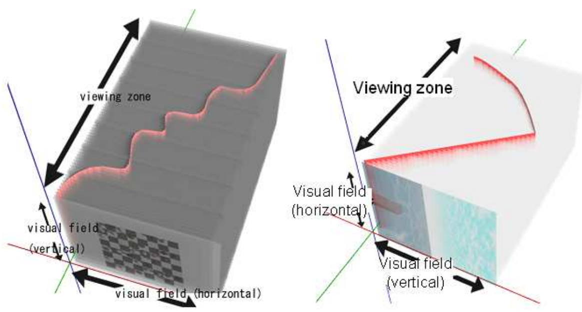


Fig.32: *Cutting ray-space with curved planes for image generation with optical effects*

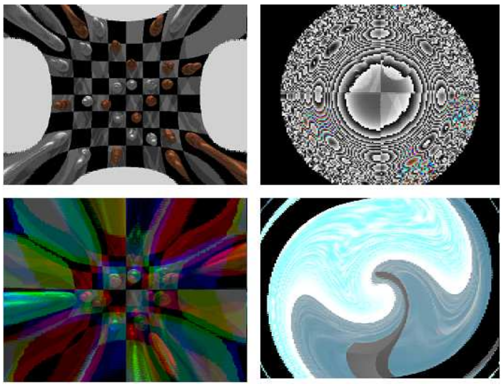


Fig.33: *Examples of artistic images generated by cutting the ray-space with more general planes*

5. INTERNATIONAL STANDARDIZATION

Figure 34 shows standardization activities of 3DAV/MVC/FTV in MPEG. MPEG established AHG on 3DAV (3D Audio Visual) and started 3DAV activities in December 2001. We proposed FTV to MPEG in July 2002 [13]. In the first stage, many 3D topics were discussed. In particular, the following four topics were studied intensively as EEs (Exploration Experiments).

EE1: omnidirectional video
EE2: free viewpoint
-FTV (Ray-Space)
-FVV (model based)
EE3: stereoscopic video
EE4: 3D-TV with depth disparity information

EE2 was established due to our proposal of FTV. The discussion was focused on FTV and its coding part (MVC, Multi-view Video Coding) in January 2004 because FTV got strong support from industry in response to the "Call for Comments on 3DAV" [38].

Succeeding in "Call for Evidence on Multi-view Video Coding" [39] led the start of standardization of MVC in January 2005. "Call for Proposals on Multi-view Video Coding" [40] was issued in July 2005. The background of MVC is described in [14]. The proposals were evaluated in January 2006 [41]. The MVC activity moved to the Joint Video Team (JVT) of MPEG and ITU for further standardization processes in July 2006. The standardization of MVC will be finalized in July 2008 as shown in Table 2. The new standard of MVC is based on H.264/MPEG4-AVC and will be MPEG-4 Part 10, Amd. 4.

The FTV cannot be constructed by coding part alone. We proposed to standardize the entire FTV under the support of SCOPE-I of the Ministry of Internal Affairs and Communications (MIC) [42] and presented the requirements for FTV [43]. The FTV standardization is supported by industry [44]. MPEG established AHG on FTV [45] for the standardization of FTV in April 2007 and issued "Preliminary FTV Model and Requirements" [46].

In October 2007, the scope of MPEG-FTV was extended to cover MVC and SSV (Stereoscopic Video) and a document "Applications and Requirements on FTV" [47] was issued. This document combines the applications and requirements of FTV, SSV and MVC to keep the consistency of the larger scope of FTV project.

As shown in Fig. 35, FTV system can be constructed in various ways, depending on the location of depth search and interpolation. FTV data formats for the 3 cases are shown in Fig. 36. MPEG selected Case B where depth search is performed at sender side and interpolation is performed at receiver side. Depth can be estimated from multi-view video or can be measured [48]. This configuration is suitable for the download/package and broadcast services since processing at sender side is heavy and processing at receiver side is light.

Figure 37 shows the FTV reference model [46] and standardization items. Transmission data format and protocol are discussed at ITU-T (International Telecommunication Union Telecommunication Standardization Sector). At the sender side of the FTV reference model, multi-view images are captured by multiple cameras. The captured images contain the misalignment and color differences of the cameras. They are corrected and the depth of each camera image is obtained. The corrected multi-view images and multi-view depth are compressed for transmission and storage by the encoder. At the receiver side, the multi-view images and multi-view depth are reconstructed by the decoder. Free-viewpoint images are generated by interpolating the reconstructed images using multi-view depth information and displayed on a 2D/3D display.

An example of intermediate view generation by interpolating multi-view images using multi-view depth is shown in Fig. 38 [49]. Compared to the view synthesis by using one view and one depth, the number of pixels which are not filled is greatly reduced. Especially, the occluded area is successfully interpolated by the reliable pixel data.

Please follow the link to subscribe MPEG-FTV. https://mailman.rwth-aachen.de/mailman/listinfo/mpeg-ftv

Table 2: *Schedule of MVC Standardization*

2006/10	WD(Working Draft)
2007/07	CD(Committee Draft)
2008/01	FCD(Final Committee Draft)
2008/07	FDIS (Final Draft International Standards

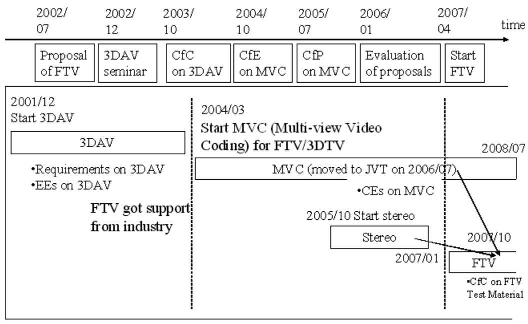


Fig.34: *Examples of artistic images generated by cutting the ray-space with more general planes*

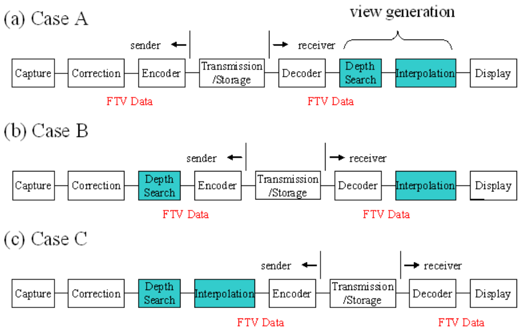


Fig.35: *3 cases of FTV configuration based on the positions of depth search and interpolation*

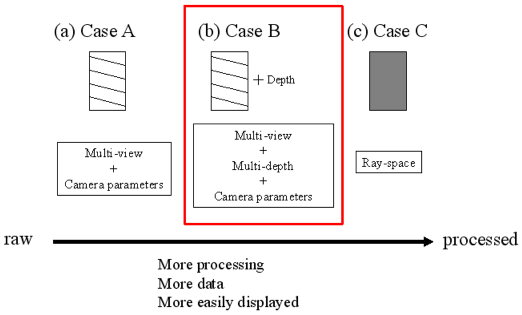


Fig.36: *Candidates for FTV data format*

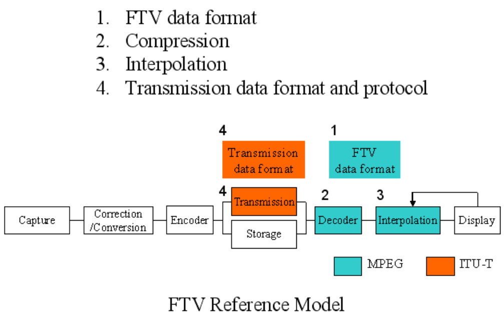


Fig.37: *FTV reference model and standardization items*

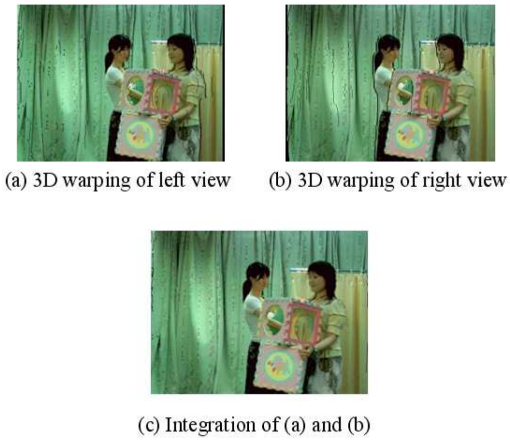


Fig.38: *An example of view generation by FTV reference model*

6. CONCLUSIONS

We have developed FTV that enables users to view a 3D scene by freely changing the viewpoints. FTV is the most advanced visual system and will bring an epochal change in the history of television.

The most essential element of visual systems is ray. FTV is not a conventional pixel-based system but a ray-based system. FTV needs advanced technologies of ray capture, processing and display. We are creating ray-based image engineering through the development of FTV.

The international standardization of FTV is now underway at MPEG and ITU-T. The introduction of FTV is not far because the interactive view generation of FTV has already been realized in real time on a single PC. Therefore, we can enjoy FTV on PCs if the new standard of FTV is available and FTV contents are delivered over internet or with packaged media.

7. ACKNOWLEDGEMENT

We would like to thank JEITA and FTV Committee members for their support to FTV standardization activity. This research is partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (A) 18206044, and the Ministry of Economy, Trade and Industry and the Ministry of Internal Affairs and Communications, Strategic Information and Communications R&D Promotion Programme (SCOPE) 041306003.

References

[1] Masayuki Tanimoto, "Free Viewpoint Television," *The Journal of Three Dimensional Images*, vol.15, no.3, pp.17-22, September 2001 (in Japanese).

[2] Masayuki Tanimoto, "Overview of Free View-

point Television," *Signal Processing: Image Communication*, vol. 21, no.6, pp.454-461, July 2006

[3] Masayuki Tanimoto, "Free Viewpoint Television - FTV," *Picture Coding Symposium 2004*, Special Session 5, December 2004.

[4] Masayuki Tanimoto, "FTV (Free Viewpoint Television) Creating Ray-Based Image Engineering," *Proc. of ICIP2005*, pp. II-25-II-28, September 2005.

[5] M. Tanimoto and T. Fujii, "FTV-Free Viewpoint Television," ISO/IEC JTC1/SC29/WG11, M8595, July 2002.

[6] Masayuki Tanimoto; "Free Viewpoint Television," *OSA Topical Meeting on Digital Holography and Three Dimensional Imaging*, DWD2, June 2007(Invited Paper).

[7] Akira Kubota, Aljoscha Smolic, Marcus Magnor, Masayuki Tanimoto, Tsuhan Chen, and Cha Zhang; "Multiview Imaging and 3TDTV," *IEEE Signal Processing Magazine*, Vol.24, No.6, pp.10-21, November 2007.

[8] T. Fujii, "A Basic Study on Integrated 3-D Visual Communication," *Ph.D dissertation in engineering, The University of Tokyo*, 1994 (in Japanese).

[9] T. Fujii, T. Kimoto and M. Tanimoto, "Ray Space Coding for 3D Visual Communication," *Picture Coding Symposium 1996*, pp. 447-451, March 1996.

[10] M. Tanimoto, A. Nakanishi, T. Fujii and T. Kimoto, "The Hierarchical Ray-Space for Scalable 3-D Image Coding," *Picture Coding Symposium 2001*, pp. 81-84, April 2001.

[11] T. Fujii and M. Tanimoto, "Free-viewpoint Television based on the Ray-Space representation," *Proc. SPIE ITCom 2002*, pp. 175-189, August 2002.

[12] M. Sekitoh, T. Fujii, T. Kimoto and M. Tanimoto, "Bird's Eye View System for ITS," *IEEE, Intelligent Vehicle Symposium*, pp. 119-123, 2001.

[13] M. Tanimoto and T. Fujii, "FTV Free Viewpoint Television", ISO/IEC JTC1/SC29/WG11, M8595, July 2002.

[14] "Introduction to Multi-view Video Coding", ISO/IEC JTC 1/SC 29/WG11, N7328, July 2005.

[15] "AHG on FTV (Free Viewpoint Television)", ISO/IEC JTC1/SC29/WG11 N8947, April 2007.

[16] N. Bangchang, T. Fujii and M. Tanimoto, "Experimental System of Free Viewpoint TeleVision," *Proc. IST/SPIE Symposium on Electronic Imaging*, Vol. 5006-66, pp. 554-563, Jan. 2003.

[17] T. Fujii, K. Mori, K. Takeda, K. Mase, M. Tanimoto and Y. Suenaga, "Multipoint Measuring System for Video and Sound: 100-camera and microphone system", *IEEE 2006 International Conference on Multimedia & Expo (ICME)*, pp. 437-440, July 2006.

[18] K. Matsumoto, T. Yendo, T. Fujii and M. Tanimoto, "Multiple-Image Rectification for FTV," Proc. of 3D Image Conference 2006, P-19, pp. 171-174, July 2006.

[19] Norishige Fukushima, Tomohiro Yendo, Toshiaki Fujii, Masayuki and Tanimoto, "A Novel Rectification Method for Two-Dimensional Camera Array by Parallelizing Locus of Feature Points," *Proc. of IWAIT2008*, B5-1, January 2008.

[20] Kenji Yamamoto,Tomohiro Yendo, Toshiaki Fujii,Masayuki Tanimoto, "Colour Correction for Multiple-camera System by using Correspondences," *The journal of the institute of Image Information and Television Engineers*, vol. 61,no. 2, pp.213-222 ,February 2007

[21] Yun He, Joern Ostermann, Masayuki Tanimoto, Aljoscha Smolic; "Introduction to the Special Section on Multiview Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.17, No.11, pp.1433-1435(November 2007).

[22] Kenji Yamamoto, Masaki Kitahara, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, Shinya Shimizu, Kazuto Kamikura, and Yoshiyuki Yashima; "Multiview Video Coding Using View Interpolation and Color Correction," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.17, No.11, pp.1436-1449(November 2007).

[23] A. Nakanishi, T. Fujii, T. Kimoto and M. Tanimoto, "Ray-Space Data Interpolation by Adaptive Filtering using Locus of Corresponding Points on Epipolar Plane Image," *The Journal of the Institute of Image Information and Television Engineers (ITE)*, vol. 56, no. 8, pp. 1321-1327, August 2002.

[24] M. Droese, T. Fujii and M. Tanimoto, "Ray-Space Interpolation based on Filtering in Disparity Domain," *Proc. 3D Conference 2004*, pp. 213-216, Tokyo, Japan, June 2004.

[25] M. Droese, T. Fujii and M. Tanimoto, "Ray-Space Interpolation Constraining Smooth Disparities Based On Loopy Belief Propagation," *Proc. of IWSSIP 2004*, pp. 247-250, Poznan, Poland, Sept. 2004.

[26] N. Fukushima, T. Yendo, T. Fujii and M. Tanimoto, "Real-time arbitrary view interpolation and rendering system using Ray-Space," *Proc. SPIE Three-Dimensional TV*, Video, and Display IV, Vol. 6016, pp. 250-261, Nov. 2005.

[27] N. Fukushima, T. Yendo, T. Fujii and M. Tanimoto, "An Effective Partial Interpolation Method for Ray-Space," *Proc. of 3D Image Conference 2006*, pp. 85-88, July 2006.

[28] N. Fukushima, T. Yendo, T. Fujii and M. Tanimoto, "Free Viewpoint Image Generation Us-

ing Multi-Pass Dynamic Programming," *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems XIV*, Vol.6490, pp. 460-470, Feb. 2007.

[29] M. Sugawara, M. Kanazawa, K. Mitani, H. Shimamoto, T. Yamashita and F. Okano, "Ultrahigh-Definition Video System with 4000 Scanning Lines," *SMPTE Motion Imaging*, 112, pp. 339-346, October/November 2003.

[30] T. Fujii and M. Tanimoto, "Real-Time Ray-Space Acquisition System", *SPIE Electronic Imaging*, vol. 5291, pp. 179-187, Jan. 2004.

[31] Keisuke Manoh, Tomohiro Yendo, Toshiaki Fujii, and Masayuki Tanimoto, "Ray-Space Acquisition System of All-Around Convergent Views using a Rotation Mirror", *Proc. of SPIE*, Vol. 6778, pp.67780C-1-8, September 2007.

[32] Toshiaki Fujii, Tomohiro Yendo, Masayuki Tanimoto; "Ray-Space Transmission System with Real-Time Acquisition and Display", *Proc. of IEEE Lasers and Electro-optics Society Annual Meeting 2007*,pp.78-79, October 2007.

[33] T. Endo, Y. Kajiki, T. Honda and M. Sato, "Cylindrical 3-D Video Display Observable from All Directions," *Proc. of Pacific Graphics 2000*, pp. 300-306, October 2000.

[34] Ryuji Takano, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, "Scene Separation in Ray-Space," *Proc. of IMPS2005*, pp.31-32, November 2005

[35] Ryuji Takano, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, "Scene Separation and Synthesis Processing in Ray-Space," Proc. of IWAIT2007, P6-23, pp.878-883, January 2007.

[36] Naoyuki Chimura, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, "New Visual Arts by Processing Ray-Space," *Proc. of Electronic Imaging & the Visual Arts (EVA) 2007 Florence*, pp.170–175, March 2007.

[37] Naoyuki Chimura, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, "Image Generation with Special Effects by Deforming Ray-Space," *Proc. of NICOGRAPH*, S1-4, November 2007.

[38] "Call for Comments on 3DAV," ISO/IEC JTC1/SC29/WG11 N6051, Oct. 2003.

[39] "Call for Evidence on Multi-view Video Coding," ISO/IEC JTC1/SC29/WG11 N6720, Oct. 2004.

[40] "Call for Proposals on Multi-view Video Coding," ISO/IEC JTC1/SC29/WG11 N7327, July 2005.

[41] "Subjective test results for the CfP on Multi-view Video Coding," ISO/IEC JTC1/SC29/WG11 N7779, January 2006.

[42] M. Tanimoto, T Fujii, S. Sakazawa and H. Kimata,"Proposal on Standardization of Free Viewpoint TV (FTV)," ISO/IEC JTC1/SC29/WG11, M13612 (JVT-T140), July 2006.

[43] M. Tanimoto, T. Fujii, H. Kimata and S. Sakazawa, "Proposal on Requirements for FTV," ISO/IEC JTC1/SC29/WG11, M14417, April 2007.

[44] "Liaison Statement from the 3D Consortium to SC29/WG11 on FTV Standardization," ISO/IEC JTC1/SC29/WG11, M 14533, April 2007.

[45] "AHG on FTV (Free Viewpoint Television)," ISO/IEC JTC1/SC29/WG11 N9170, April 2007.

[46] "Preliminary FTV Model and Requirements," ISO/IEC JTC1/SC29/WG11, N9168, July 2007.

[47] "Applications and Requirements on FTV," ISO/IEC JTC1/SC29/WG11, N9466, October 2007.

[48] M. Kawakita, K. Iizuka, H. Nakamura, I. Mizuno, T. Kurita, T. Aida, Y. Yamanouchi, H. Mitsumine, T. Fukaya, H. Kikuchi, and F. Sato, "High-definition real-time depth-mapping TV camera: HDTV Axi-Vision Camera," *Optics Express*, Vol. 12, Issue 12, pp. 2781–2794, June 2004.

[49] M. Tanimoto, T. Fujii and K. Suzuki, "Experiment of view synthesis using multi-view depth," ISO/IEC JTC1/SC29/WG11, M14889, October 2007.

Masayuki Tanimoto received the B.E., M.E., and Dr.E. degrees in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1970, 1972, and 1976, respectively. His doctoral work was on GaAs devices. From 1974 to 1975, he was with the Institut fur Halbleitertechnik, TH Aachen, Germany. He joined Nagoya University, Nagoya, Japan, in 1976 and started research on visual communication and communication systems. Since 1991, he has been a Professor at Graduate School of Engineering, Nagoya University. His current research interests include image coding, image processing, 3-D images, FTV and ITS.

Dr. Tanimoto was the Vice President of ITE, the Chairperson of Technical Group on Communication Systems of IEICE, the Chairperson of the Steering Committee of Picture Coding Symposium of Japan, IEICE Councilor, ITE Councilor and Tokai Section Chair of IEICE. He was a member of Editorial Board of Image Communication of EURASIP and is a member of the International Steering Committee of the Picture Coding Symposium. He received the Ichimura Award, the TELECOM System Technology Award from The Telecommunications Advancement Foundation, the Niwa-Takayanagi Paper Award from ITE, Outstanding Papers Interactive Sessions Award in IEEE Intelligent Vehicle Symposium, Excellent Paper Award in 3-D Image Conference, IEICE Fellow Award, IEICE Achievement Award, ITE Fellow Award, and the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science, and Technology.