# FTV (FREE-VIEWPOINT TELEVISION) FOR RAY AND SOUND REPRODUCING IN 3D SPACE

*Masayuki Tanimoto*

Graduate School of Engineering, Nagoya University, Japan

## ABSTRACT

FTV (Free-viewpoint Television) is the ultimate 3DTV that captures all rays in 3D space and reproduce the same rays at different place and time, thus it allows users to view a 3D scene by freely changing the viewpoints. We proposed the concept of FTV and realized FTV by developing various types of ray capture, processing and display technologies. At present, FTV is available on a laptop PC or a mobile player, and FTV with free listening-point audio is also realized. The international standardization of FTV has been conducted in MPEG. The first phase of FTV was MVC (Multi-view Video Coding) and the second phase is 3DV (3D Video).

***Index Terms***— FTV, Free-viewpoint television, ray-space, MVC, 3DV

## 1. INTRODUCTION

We are surrounded by audio-visual information in 3D space. It is the ultimate target of communication and broadcast to transmit and reproduce 3D space with all audio-visual information.

FTV (Free-viewpoint Television) [1]-[4] enables us to view a 3D scene by freely changing our viewpoints as if we were there. FTV is the ultimate 3DTV that transmits the infinite number of views and ranked as the top of media. It is also the best interface between human and environment, and an innovative tool to create new types of content and art.

We proposed the concept of FTV and verified its feasibility with the world's first real-time system including the complete chain of operation from image capture to display [5]. FTV with audio was realized by adding free listening-point function [6].

FTV is based on the ray-space method [7]-[9]. We developed ray capture, processing, and display technologies for FTV. All-around ray-reproducing FTV has been realized by using these technologies [10].

The international standardization of FTV has been conducted in MPEG. The first phase of FTV was MVC (Multi-view Video Coding). MVC was completed in May 2009 and has been adopted by Blu-ray 3D. The second phase of FTV is 3DV (3D Video). 3DV is a standard that targets serving a variety of 3D displays.

## 2. PRINCIPLE OF FTV

FTV transmits a finite number of views captured by multi-camera and the other views at non-camera positions are generated. This free-viewpoint image generation is performed by integration and interpolation of rays. This process is performed systematically in the ray-space. We define two types of ray-space. One is orthogonal ray-space and another is spherical ray-space. The orthogonal ray-space is obtained by linear camera arrangement and the spherical ray-space is obtained by circular camera arrangement.

For the linear camera arrangement, the orthogonal ray-space is constructed by placing captured camera views upright and parallel, as shown in Fig. 1, and filling the vacancy between views. An example of filled ray-space is shown in Fig. 2. As seen in this figure, the horizontal cross-section of ray-space has a line structure. The line structure is used for the ray interpolation. A free viewpoint image is generated by cutting the ray-space vertically with a planar knife at a position determined by the viewpoint. Several parallel knifes are used to cut the ray-space to generate view images for a 3D display.

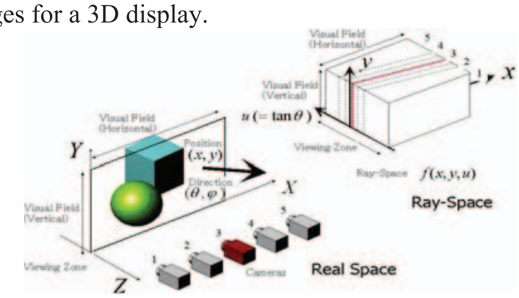


**Fig. 1.** Acquisition of orthogonal ray-space.

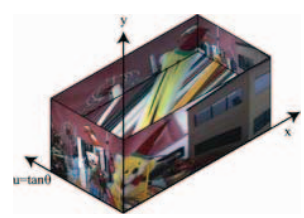


**Fig. 2.** Orthogonal ray-space and its horizontal cross-section.

## 3. FTV SYSTEM

### 3.1. Configuration of FTV System

At the sender side of the FTV system, 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.

### 3.2. Capture

We constructed a real-time FTV system [11], [12], where a 1D-arc multi-camera is used for capturing. 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" [13] as shown in Fig. 3 was developed to capture larger spaces. 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.

Furthermore, we have developed an all-around dense ray capture system and an efficient ray capture method that can capture several rays per 1 pixel [14].

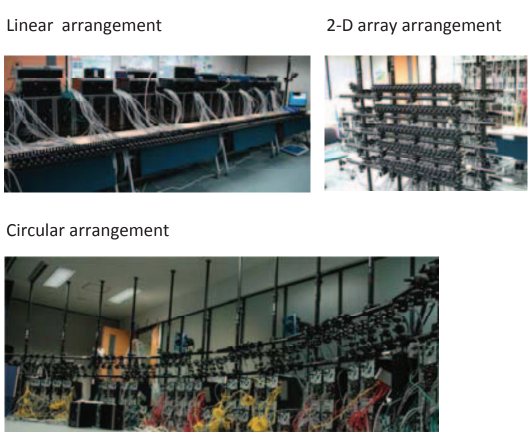


**Fig. 3.** 100-camera system.

### 3.3 Correction

The geometric correction [15], [16] and color correction [17] of multi-camera images are performed by measuring the correspondence points of images. This measurement is made once the cameras are set.

### 3.4 View Generation

Ray interpolation for view generation is made by detecting depth information pixel by pixel from the multi-view video. We proposed several ray interpolation schemes of the ray-space [18]-[21].

In the beginning, free-viewpoint images were generated by a PC cluster. Then, FTV was implemented on a laptop PC and a mobile player.

Synthesized view contains artifacts if there is error in depth estimation. Considering the fact that synthesis artifacts can be measured at the location of the reference views, thus we can eliminate similar error in the synthesized location. Based on this idea, we have developed three methods which can be applied during free-viewpoint rendering to suppress errors [22]-[24]. Fig. 4 shows view synthesis by reliability-based optimization and improvement of free-viewpoint images.

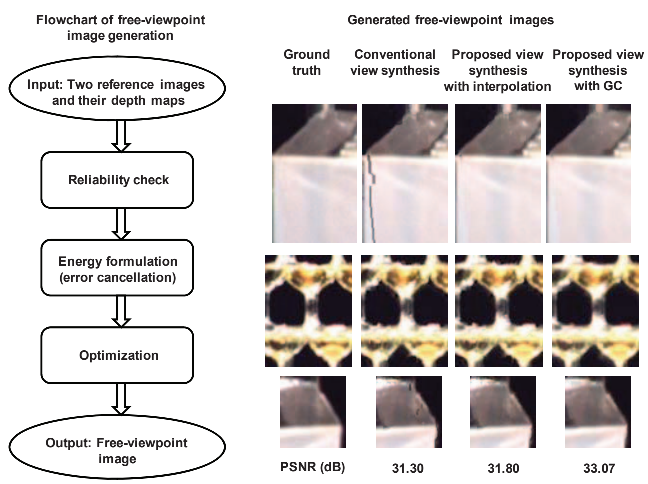


**Fig. 4.** View synthesis by reliability-based optimization and improvement of free-viewpoint images.

### 3.5 User Interface

Various types of user interface as shown in Fig. 5 were developed for FTV.

The first type of user interface shows views only at the viewpoints given by the user. User's viewpoint can be given using an eye/head-tracking system, a remote controller, or a touch panel. We realized this type of user interface on conventional 2D, stereoscopic and autostereoscopic 3D displays.

The second type of user interface provides all views so that users can see any views by changing their locations in front of the interface. Seelinder [25] belongs to the second type. It is a 360-degree, ray-producing display that allows multiple viewers to see 3D FTV images.

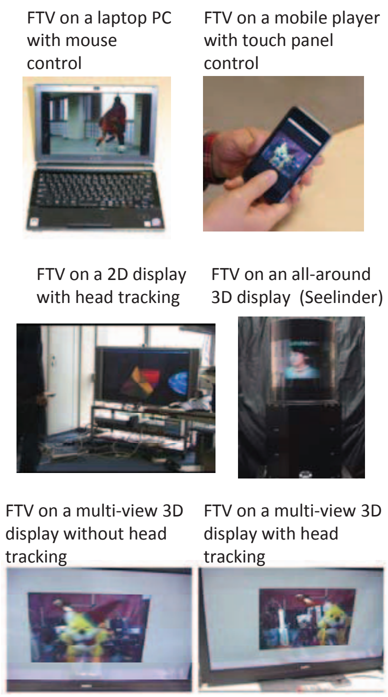


**Fig. 5.** Various types of FTV user interface.

## 4. FTV AUDIO

Realization of free listening-point audio generation in real environment is a challenging issue. Two free listening-point audio generation methods were employed for FTV [5]. The first method is based on sound wave ray-space method [26]. The second method is based on acoustic transfer function (ATF) estimation and blind signal separation (BSS) of sources of sound [27]. Using each of listening-point audio generation methods, we are able to integrate the audio-visual information, and realize FTV with audio.

## 5. INTERNATIONAL STANDARDIZATION OF FTV

We proposed FTV to MPEG in 2001. First, many topics on 3D were discussed. The discussion was converged on FTV in January 2004 and the standardization of the coding part of FTV started as MVC (Multi-view Video Coding) [28]. MVC was the first phase of FTV and completed in March 2009. MVC is based on H.264/MPEG4-AVC and has been adopted by Blu-ray 3D. MVC is used as shown in Fig. 6.

FTV cannot be constructed by coding part alone. We proposed to standardize the entire FTV and MPEG started the standardization activity of FTV and MPEG-FTV was established in April 2007.

In January 2008, MPEG-FTV targeted the standardization of 3DV (3D Video) [29]. 3DV is a standard serving for a variety of 3D displays. 3DV is the second phase of FTV. "Call for Proposals on 3D Video Coding Technology" was issued in March 2011 [30]. 3DV is used as shown in Fig. 7.

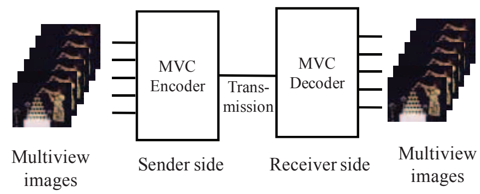


**Fig. 6.** Framework of MVC (Multi-view Video Coding).

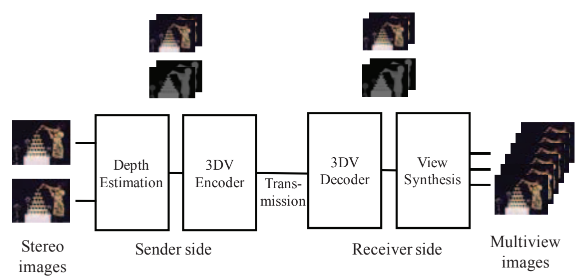


**Fig. 7.** Framework of 3DV (3D Video).

## 6. CONCLUSION

FTV captures all rays in 3D space and reproduce the same rays at different place and time. FTV is the ultimate 3DTV with infinite number of views and ranked as the top of visual media. We realized FTV by developing various types of ray capture, processing and display technologies. At present, FTV is available on a laptop PC or a mobile player. We have also realized all-around ray-reproducing FTV.

FTV enables the realistic viewing and free navigation of 3D scenes. FTV will find many applications in the fields of broadcast, communication, amusement, entertainment, advertising, design, exhibition, education, medicine and so on. FTV became the key concept of 2022 FIFA World Cup bidding to Japan though the bid was not successful. Japan planed to deliver the 3D replica of soccer stadium to all over the world by FTV.

The international standardization of FTV has been conducted in MPEG. The first phase of FTV is MVC and the second phase is 3DV. MVC has been adopted by Blu-ray 3D.

## 7. ACKNOWLEDGMENT

This research was partially supported by Strategic Information and Communications R&D Promotion Programme (SCOPE) of the Ministry of Internal Affairs and Communications, National Institute of Information and Communications Technology, Japan (NICT), and Grant-in-Aid for Scientific Research (B), 22360151.

## 8. 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, "Free Viewpoint Television - FTV", Picture Coding Symposium 2004, Special Session 5, December 2004.

[3] Masayuki Tanimoto, "Overview of Free Viewpoint Television," Signal Processing: Image Communication, vol. 21, no. 6, pp. 454-461, July 2006.

[4] Masayuki Tanimoto, Mehrdad Panahpour Tehrani, Toshiaki Fujii, Tomohiro Yendo, "Free-Viewpoint TV", IEEE Signal Processing Magazine, vol.28, no.1, pp.67-76, January 2011.

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

[6] M. Panahpour Tehrani, K. Niwa, N. Fukushima, Y. Hirano, T. Fujii, M. Tanimoto, K. Takeda, K. Mase, A. Ishikawa, S. Sakazawa, A. Koike, "3DAV Integrated System Featuring Free Listening-point and Free Viewpoint Generation", IEEE International Conference on Multimedia signal processing, MMSP, 855-860, Australia, Oct. 2008.

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

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

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

[10] T. Yendo, T. Fujii, M. Panahpour Tehrani and M. Tanimoto: "All-Around Ray-Reproducing 3DTV", IEEE International Workshop on Hot Topics in 3D (Hot 3D), July 2011.

[11] P. Na 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.

[12] P. Na Bangchang, M. Panahpour Tehrani, T. Fujii, M. Tanimoto, "Realtime System of Free Viewpoint Television", The journal of the institute of Image information and Television Engineers (ITE), Vol. 59, No. 8, pp. 63-701, Aug. 2005.

[13] 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.

[14] Masayuki Tanimoto, "FTV and All-Around 3DTV", IEEE Visual Communications and Image Processing 2011, S-02, November 2011.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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, September 2004.

[20] 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.

[21] N. Fukushima, T. Yendo, T. Fujii and M. Tanimoto, "Free Viewpoint Image Generation Using Multi-Pass Dynamic Programming", Proc. SPIE Stereoscopic Displays and Virtual Reality Systems XIV, Vol.6490, pp. 460-470, Feburuary 2007.

[22] H. Furihata, M. Panahpour Tehrani, T. Yendo, .T Fujii, M. Tanimoto, "Novel view synthesis with residual error feedback for FTV", Proc. IS&T/SPIE Electronic Imaging, session on DIBR and FTV, 7524-19 Jan. 2010, San Jose, USA.

[23] L. Yang, T. Yendo, M. Panahpour Tehrani, T. Fujii and M. Tanimoto: "Artifact reduction using reliability reasoning for image generation of FTV", Journal on Visual Communication and Image Representation, 21, 5-6, pp. 542-560, Jul.-Aug. 2010.

[24] L. Yang, T. Yendo, M. Panahpour Tehrani, T. Fujii and M. Tanimoto, "Probabilistic reliability based view synthesis for FTV", IEEE International Conference on Image Processing, ICIP2010, September 2010.

[25] Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, Mehrdad Panahpour Tehrani, "The Seelinder: Cylindrical 3D display viewable from 360 degrees", Journal of visual communication and image representation , Vol. 21, Issues 5-6, pp. 586-594, Jul.-Aug. 2010.

[26] M. Panahpour Tehrani, Y. Hirano, T. Fujii, S. Kajita, K. Takeda, K. Mase, "The Sub-band Sound Wave Ray-Space Representation", IEEE ICASSP , MMSP-p3.11, v541-v544, May 2006.

[27] K. Niwa, T. Nishino, C. Miyajima, K. Takeda, "Blind source separation of musical signals applied to selectable-listening-point audio reconstruction" Jr of ASA, 3pSP20, Hawaii, Nov 2006.

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

[29] "Introduction to 3D Video," ISO/IEC JTC1/SC29/WG11 N9784, May 2008.

[30] "Call for Proposals on 3D Video Coding Technology," ISO/IEC JTC1/SC29/WG11 MPEG, N12036, March 2011.