

International Standardization of FTV

Masayuki Tanimoto

Nagoya University and Nagoya Industrial Science Research Institute

Nagoya, Japan

tanimoto@nagoya-u.jp

Abstract – FTV (Free-viewpoint Television) is visual media that transmits all ray information of a 3D space and enables immersive 3D viewing. The international standardization of FTV has been conducted in MPEG. The first phase of FTV is multiview video coding (MVC), and the second phase is 3D video (3DV). The third phase of FTV is MPEG-FTV, which targets revolutionized viewing of 3D scenes via super multiview, free navigation, and 360-degree 3D. After the success of exploration experiments and Call for Evidence, MPEG-FTV moved MPEG Immersive project (MPEG-I), where it is in charge of video part as MPEG-I Visual. MPEG-I will create standards for immersive audio-visual services.

1 INTRODUCTION

FTV (Free-viewpoint Television) [1]-[8] is visual media that transmits all ray information of a 3D space. FTV was developed based on ray-space representation [9]-[12].

FTV is the ultimate 3DTV, with an infinite number of views, and ranks at the top of visual media. FTV enables users to view a 3D scene by freely changing the viewpoint, as we do naturally in the real world. FTV is a natural interface between humans and the environment. It is also an immersive media that enables a realistic VR experience and revolutionizes 3D viewing.

FTV was proposed to the Moving Picture Experts Group (MPEG) in 2001 [13]. Since then, the MPEG has been developing various FTV standards. Multiview video coding (MVC) [14] is the first phase of FTV and enables efficient coding of multiple camera views. 3D video (3DV) [15] is the second phase of FTV and enables viewing adaptations and display adaptations for multiview 3D displays. MPEG started the third phase of FTV [16] in August 2013. This is MPEG-FTV, which targets immersive viewing of 3D scenes via super multiview, free navigation, and 360-degree 3D (360 3D) video. MPEG-FTV moved to MPEG-Immersive project (MPEG-I) [17] in January 2017 and it has been in charge of video part as MPEG-I Visual.

In this paper, international standardization of FTV is described.

2 HISTORY OF FTV STANDARDIZATION IN MPEG

The MPEG has been developing FTV standards since 2001. The history of FTV standardization in MPEG is shown in Fig. 1. In 2001, FTV was proposed to the MPEG, and the 3D audio visual (3DAV) activity started. In 3DAV activity, many topics, such as omnidirectional video, FTV, stereoscopic video, and 3DTV with depth information, were discussed. According to the results of the call for comments from the industry, discussion converged on FTV and MVC [14] starting in March 2004.

MVC is the first phase of FTV and targeted the coding of multiple videos. The MVC activity moved to the Joint Video Team (JVT) of the MPEG and International Telecommunication Union Telecommunication Standardization Sector (ITU-T) for further standardization processes in July 2006. MVC was completed in March 2009. MVC does not have a function for view generation.

The MPEG started 3DV as the second phase of FTV in April 2007. 3DV is a standard for multiview 3D displays [15]. View generation was introduced into 3DV to increase the number of views for multiview 3D displays. The 3DV activity moved to the Joint Collaborative Team (JCT)-3V for further standardization processes in July 2012, and 3DV was completed in June 2016.

In August 2013, the MPEG started the third phase of FTV, MPEG-FTV [16], which targets immersive 3D viewing by enhancing the function of view generation. MPEG-FTV moved to MPEG-I [17] for further standardization in January 2017.

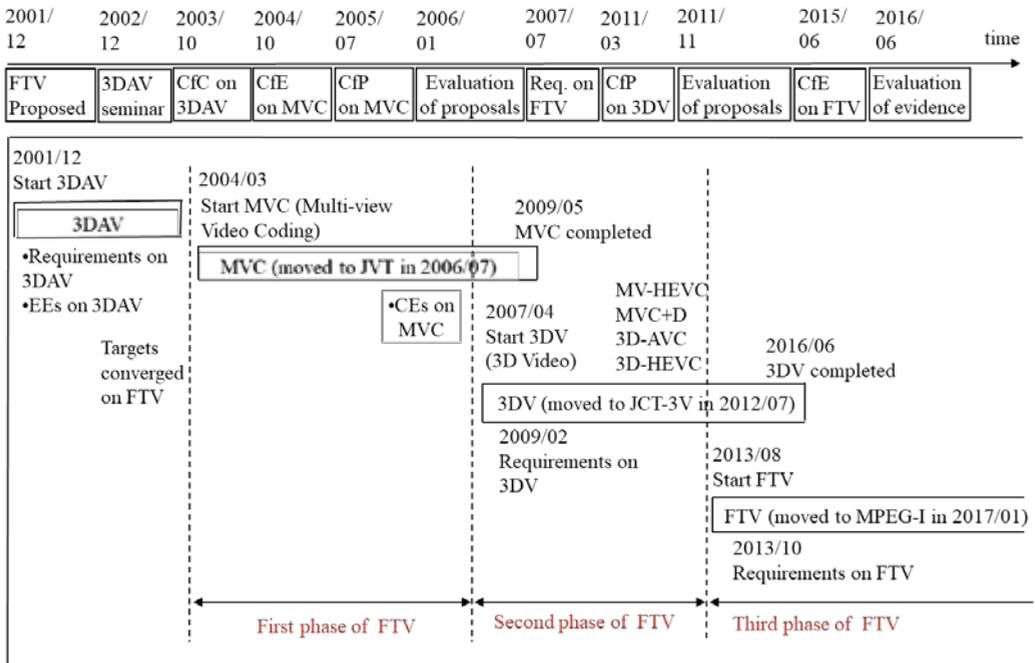


Fig. 1 History of FTV standardization in MPEG.

3 FTV FIRST PHASE: MVC STANDARD

The framework of MVC is shown in Fig. 2. MVC targets efficient coding of multiview video. In MVC, the number of input views is the same as output views. The view-generation function of FTV is not included in MVC.

Multiview video data have a high correlation among views. This redundancy can be removed by interview predication. It can also be done by using a motion compensation method that is widely used to remove temporal redundancy in conventional video coding. MVC applies motion compensation-like prediction to not only time and but also view directions.

MVC was standardized as the extension of H.264/MPEG4-AVC [18]. The MVC standard was adopted by Blu-ray 3D.

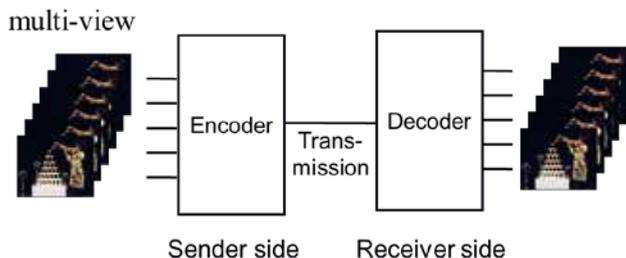


Fig. 2 Framework of MVC.

4 FTV SECOND PHASE: 3DV STANDARDS

The framework of 3DV is shown in Figure 3. View synthesis was introduced into 3DV, which sends a small number of views and generates a large number of views at the receiver for multiview displays. A multiview and multi-depth set is jointly compressed and sent to the receiver, and intermediate views are synthesized from views with the assistance of depth information at the receiver. 3DV enables display adaptation and viewing adaptation [19].

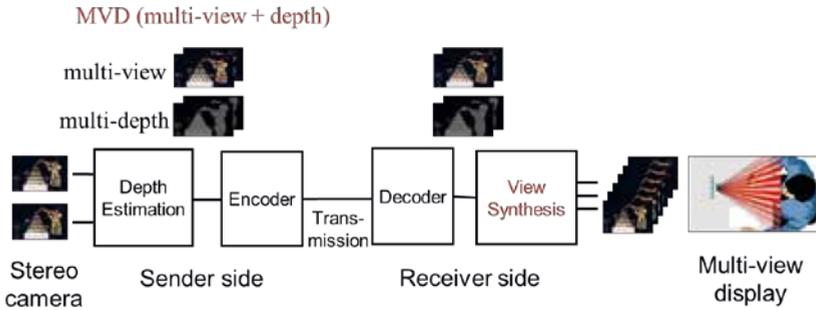


Fig. 3 Framework of 3DV.

The FTV reference model, as shown in Fig. 4, was defined to develop the 3DV standard [20], and 3D warping is used for view synthesis of 3DV. View synthesis by 3D warping is sensitive to error in depth information. Nagoya University provided Depth Estimation Reference Software (DERS) [21] and View Synthesis Reference Software (VSRS) [22], as shown in Fig. 4. It also provided various test sequences such as pantomime, champagne_tower, kendo and balloons.

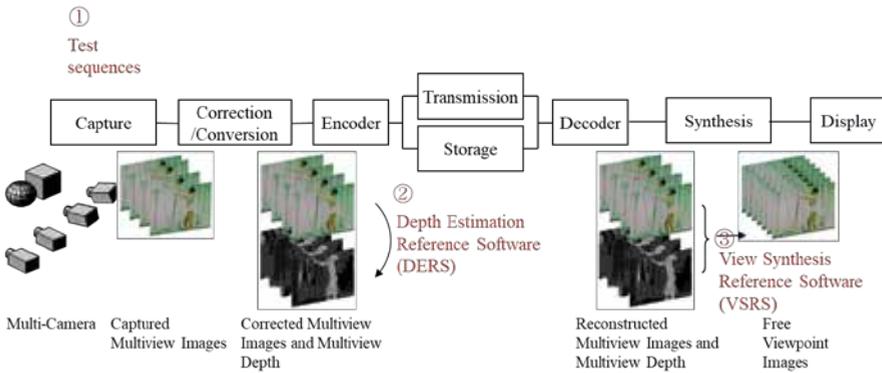


Fig. 4 FTV reference model and Nagoya University's contribution.

The data format of 3DV is Multiview plus Depth (MVD). Coding standards such as MVC+D, 3D-AVC, MV-HEVC, and 3D-HEVC were developed [23]. Here, MVC+D is a depth-extension of MVC, 3D-AVC is AVC-based MVD joint coding, MV-HEVC is HEVC-based MVC, and 3D-HEVC is HEVC-based MVD joint coding.

Global View and Depth (GVD) [24] can be used as an alternative data format. GVD is a compact version of MVD and is obtained by removing the interview redundancy of MVD.

5 FTV THIRD PHASE: MPEG-FTV/MPEG-I

5.1 Motivation and Background

In 2010, the 2022 FIFA World Cup Japan Bid Committee planned to deliver the excitement of the soccer stadium to the world via FTV. It aimed to revolutionize the viewing of the soccer game by super multiview and free navigation. Super multiview realizes very realistic 3D viewing of the scene, and free navigation realizes a walk-through or fly-through experience of the scene. This became a strong motivation for the third phase of FTV.

5.2 Framework of FTV

Based on the above motivations and background, the framework of MPEG-FTV was created, as shown in Fig. 5 [25]. FTV has three types of application scenario [26]. The first is super multiview (SMV) with a high number of views and high density for super multiview displays. The second is a single view with freely changing viewpoint for free navigation (FN) in a wide area. Users can enjoy realistic 3D viewing and walk-through/fly-through experiences in 3D scenes. The third is 360 3D video with a wide FoV.

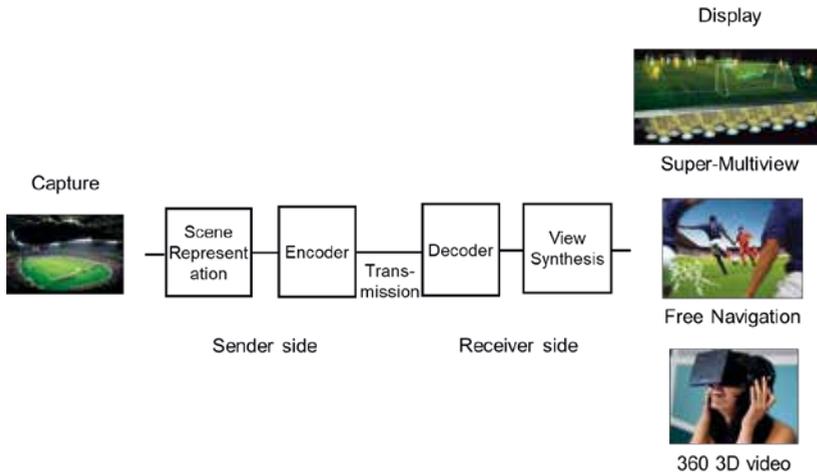


Fig. 5 Framework of MPEG-FTV.

5.3 Call for Evidence

After a series of exploration experiments on FTV [27], the MPEG issued a Call for Evidence (CfE) on FTV [28] in June 2016. CfE is a procedure before a call for proposals to show evidence of a new technology that is better than currently available standards. FTV software used for the CfE is described in [29]. Submissions were collected for SMV and FN application scenarios, as shown in Figs. 6 and 7, respectively, in February 2016. Results evaluated in June 2016 showed clear evidence of the new technology [30].

FTV test material and software developed in MPEG-FTV are summarized in [31] and [32], respectively.

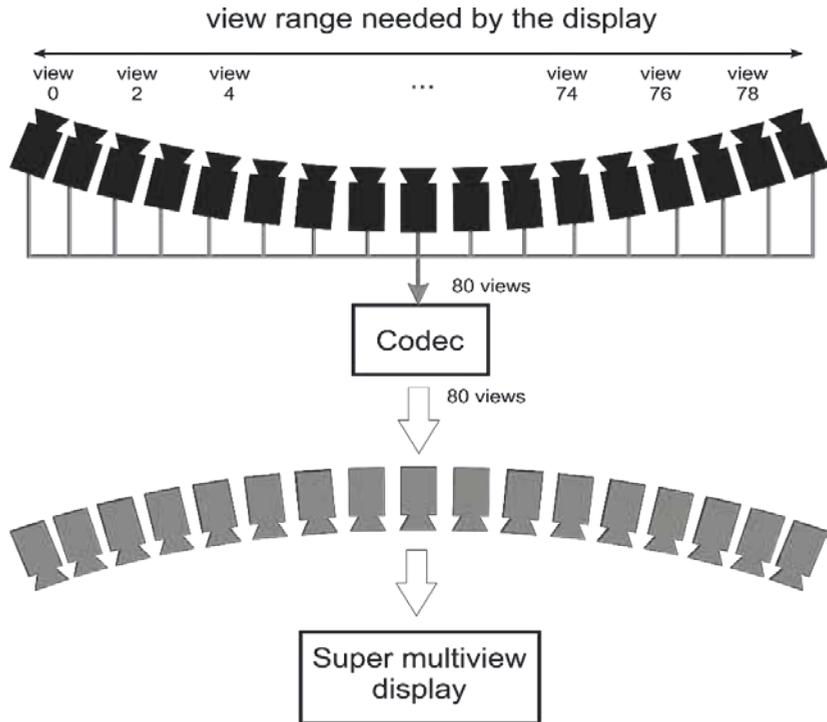


Fig. 6 Super Multiview (SMV) application scenario for the CfE [28].

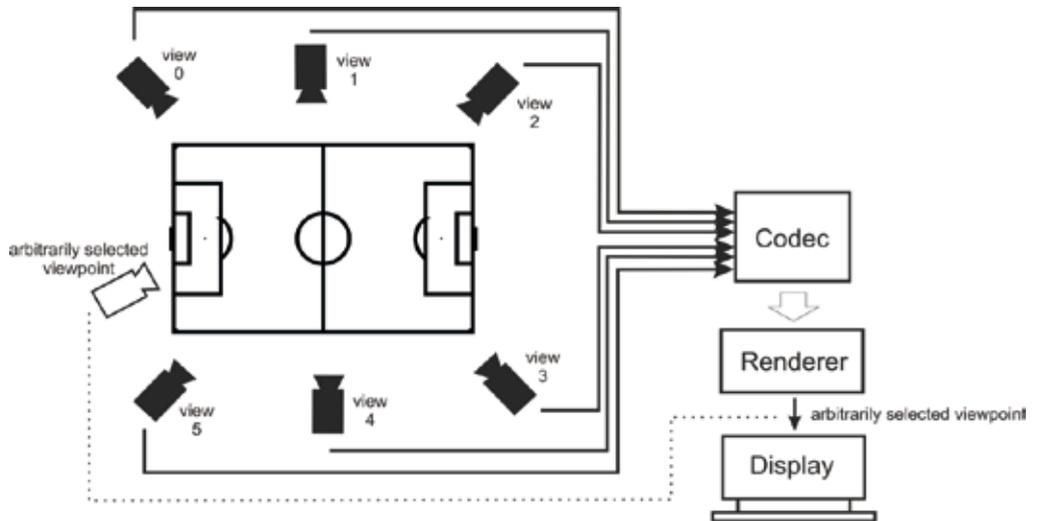


Fig. 7 Free navigation (FN) application scenario for the CfE [28].

5.4 MPEG-I

MPEG-FTV moved to MPEG-I in January 2017. MPEG-I was established by integrating the FTV, light field, point cloud, and 360 video ad hoc groups. MPEG-I will create standards for immersive services. MPEG-FTV is in charge of the video aspects of MPEG-I. All application scenarios, requirements, test material, and software for MPEG-FTV were transferred to MPEG-I.

MPEG-I will use various technologies, such as FTV, light field, point cloud, 360 video, and 3D audio, to build immersive services. Therefore, the MPEG has structured MPEG-I as a suite of standards focusing on specific technologies. The five parts to MPEG-I are as follows [33]:

- Part 1 – Technical Report on Immersive Media
- Part 2 – Application Format for Omnidirectional Media
- Part 3 – Immersive Video
- Part 4 – Immersive Audio
- Part 5 – Point Cloud Compression

MPEG-I standards will be developed according to the stages of immersion shown in Fig. 8 [33]. The stages of immersion are categorized by degrees of freedom (DoF), which denotes the number of independent parameters used to define movement of a viewpoint in 3D space.

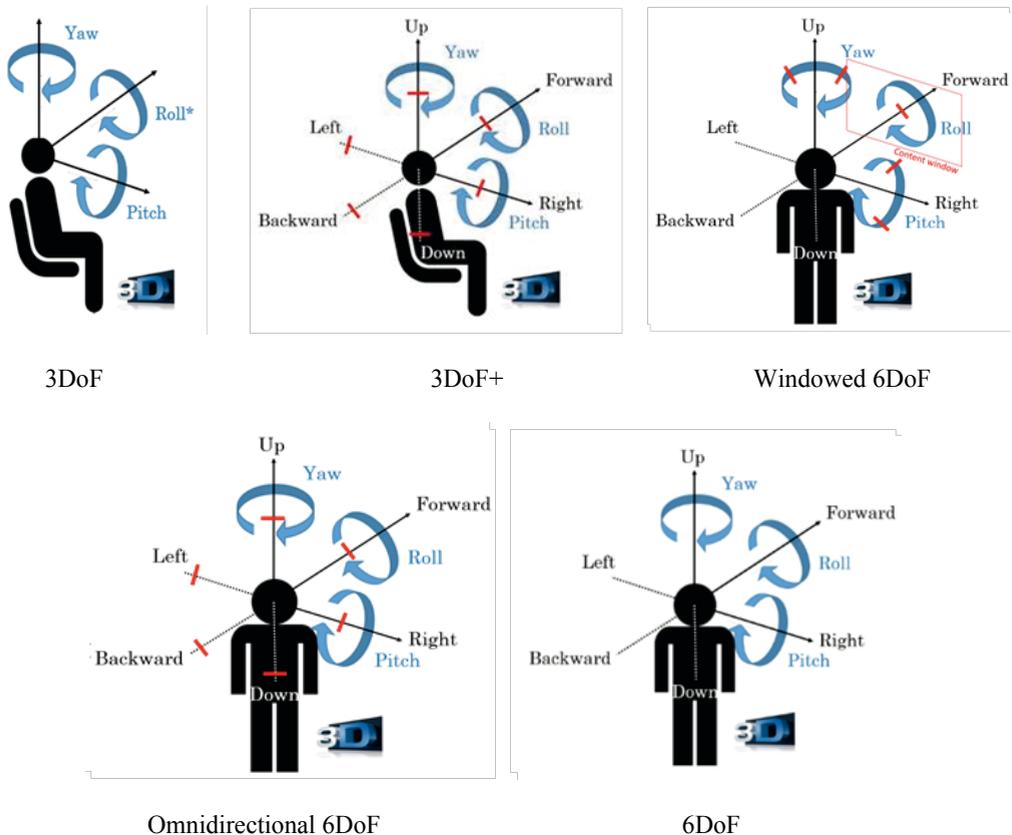


Fig. 8 Stages of Immersion in MPEG-I [33].

For example, 3DoF is three rotational and unlimited movements around the X, Y, and Z axes. 3DoF has a fixed viewpoint and no translational movements along the X, Y, and Z axes. A

typical use case is a user sitting in a chair looking at 3D 360 VR content on an HMD, as shown at the most left in Fig. 8. On the other hand, 6DoF is 3DoF with full translational movements along the X, Y, and Z axes. A typical use case is a user freely walking through 3D 360 VR content displayed on an HMD, as shown at the most right in Fig. 8.

3DoF+, windowed 6DoF, and omnidirectional 6DoF are stages in between. 3DoF+ is 3DoF with additional limited translational movements along the X, Y, and Z axes. Windowed 6DoF denotes 6DoF with constrained rotational movements around the X and Y axes (pitch and yaw, respectively) and constrained translational movements along the Z axis. Omnidirectional 6DoF denotes 6DoF with constrained translational movements along X, Y, and Z axes.

6 CONCLUSION

MPEG has been creating various standards on FTV. In the first phase of FTV, MPEG developed the MVC standard. In the second phase of FTV, MPEG developed the MVC+D, 3D-AVC, MV-HEVC and 3D-HEVC standards. The current third phase of FTV is MPEG-FTV. MPEG-FTV targets revolutionized viewing of 3D scenes via super multiview, free navigation, and 360-degree 3D technologies. MPEG-FTV developed test material, reference software, and evaluation methods for them. After the success of the exploration experiments and Call for Evidence, MPEG-FTV moved to MPEG-I and has been in charge of its video part. MPEG-I will create standards for immersive services based on the stages of immersion.

ACKNOWLEDGEMENTS

This research is partially supported by JSPS KAKENHI Grant Number 16H04370.

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. Tanimoto, M. P. Tehrani, T. Fujii, T. Yendo, "FTV for 3-D Spatial Communication", *Proceedings of the IEEE*, Vol. 100, No. 4, pp. 905-917, April 2012. (invited paper)
- [6] Masayuki Tanimoto, "FTV: Free-viewpoint Television", *Signal Processing : Image Communication*, Vol. 27, Issue 6, pp. 555-570 (June 2012). (invited paper)
- [7] Masayuki Tanimoto, "FTV (Free-viewpoint Television)", *APSIPA Transactions on Signal and Information Processing*, Vol. 1, Issue 1, e4 (14 pages), August 2012. (invited paper)
- [8] Masayuki Tanimoto, "FTV (Free-viewpoint Television) and its International Standardization", *IEIE Transactions on Smart Processing and Computing*, vol. 6, no. 6, pp.415-427, December 2017. (Invited paper)
- [9] T. Fujii, "A Basic Study on Integrated 3-D Visual Communication", Ph.D dissertation in engineering, The University of Tokyo, 1994 (in Japanese).
- [10] T. Fujii, T. Kimoto and M. Tanimoto, "Ray Space Coding for 3D Visual Communication," *Picture Coding Symposium 1996*, pp. 447-451, March 1996.
- [11] 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.
- [12] T. Fujii and M. Tanimoto, "Free-viewpoint Television based on the Ray-Space representation", *Proc. SPIE ITCOM 2002*, pp. 175-189, August 2002.

- [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] “Introduction to 3D Video”, ISO/IEC JTC1/SC29/WG11 N9784, May 2008.
- [16] M. Tanimoto, T. Senoh, S. Naito, S. Shimizu, H. Horimai, M. Domański, A. Vetro, M. Preda and K. Mueller, “Proposal on a New Activity for the Third Phase of FTV,” ISO/IEC JTC1/SC29/WG11 MPEG2013/M30229, Vienna, Austria, July 2013.
- [17] “Working Draft 0.1 of Technical Report on Immersive Media,” ISO/IEC JTC1/SC29/WG11 MPEG2017/N16718, Geneva, Switzerland, January 2017.
- [18] A. Vetro, T. Wiegand, and G. J. Sullivan, “Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard”, Proc. IEEE, vol. 99, no. 4, pp. 626–642, Apr. 2011.
- [19] “Applications and Requirements on 3D Video Coding”, ISO/IEC JTC1/SC29/WG11 N10570, April 2009.
- [20] Yuji Mori, Norishige Fukushima, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, “View Generation with 3D Warping Using Depth Information for FTV,” Signal Processing: Image Communication, Volume 24, Issue 1-2 pp.65-72, January 2009.
- [21] Masayuki Tanimoto, Toshiaki Fujii, Mehrdad Panahpour Tehrani and Menno Wildeboer, “Depth Estimation Reference Software (DERS) 4.0,” ISO/IEC JTC1/SC29/WG11 MPEG2009/M16605, London, UK, June 2009.
- [22] Masayuki Tanimoto, Toshiaki Fujii and Kazuyoshi Suzuki, “View synthesis algorithm in View Synthesis Reference Software 2.0 (VSR2.0),” ISO/IEC JTC1/SC29/WG11 MPEG2009/M16090, Lausanne, Switzerland, February 2009.
- [23] “Standardization Tracks Considered in 3D Video Coding,” ISO/IEC JTC1/SC29/WG11 MPEG, N12434, December 2011.
- [24] Kazuyoshi Suzuki, Masayuki Tanimoto, "Global View Depth Format for FTV Data", Proceedings of IWAIT 2013, 3B-1, pp.198 - 201, Nagoya Japan, January 7-9, 2013.
- [25] “Use Cases and Requirements on Free-viewpoint Television (FTV),” ISO/IEC JTC1/SC29/WG11 MPEG2014/N14178, San Jose, USA, February 2014.
- [26] “Use Cases and Requirements on Free-viewpoint Television (FTV) v.3,” ISO/IEC JTC1/SC29/WG11 MPEG2016/N16130, San Diego, USA, January 2016.
- [27] “Description of Exploration Experiments on Free-viewpoint Television (FTV),” ISO/IEC JTC1/SC29/WG11 MPEG2013/N14105, Geneva, Switzerland, October 2013.
- [28] “Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation,” ISO/IEC JTC1/SC29/WG11 MPEG2015/N15348, Warsaw, Poland, June 2015.
- [29] “FTV software framework,” ISO/IEC JTC1/SC29/WG11 MPEG2015/N15349, Warsaw, Poland, June 2015.
- [30] “Summary of the results of the Call for Evidence on Free-Viewpoint Television: Super-Multiview and Free Navigation,” ISO/IEC JTC1/SC29/WG11 MPEG2016/N16318, Geneva, Switzerland, June 2016.
- [31] “FTV test material summary,” ISO/IEC JTC1/SC29/WG11 MPEG2016/N16521, Chengdu, China, October 2016.
- [32] “FTV software summary,” ISO/IEC JTC1/SC29/WG11 MPEG2016/N16522, Chengdu, China, October 2016.
- [33] “Working Draft 1.0 of TR: Technical Report on Architectures for Immersive Media,” ISO/IEC JTC1/SC29/WG11 MPEG2017/N17067, Torino, Italy, July 2017.