ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 8(S5), 28–32, March 2015* DOI : 10.17485/ijst/2015/v8iS5/61461

A Study on Transmission System for Realistic Media Effect Representation

Seokmo Gu[1] and Yei-Chang Kim[2*]

[1]Department of Technology Management Cooperative Program,Dongguk University, Gyeongju, South Korea
[2]Department of Information and Management, Dongguk University, Gyeongju, South Korea; kimyc@dongguk.ac.kr

Abstract

The emergence of next-generation broadcasting technology affected the rapid growth of a variety of broadcasting convergence industries, the distribution of diverse contents in the realistic media market and the development of transmission technology. IPTV service has also been universalized, becoming a service available to everyone. Although 4K UHD (Ultra High Definition) based realistic media makes it possible to provide better resolution quality than conventional contents, the media faces the limits of transmission technology and encoding technique. To solve the problem, much research is being conducted. Also to increase contents users' presence and immersion, research on transmission of realistic media effects and contents is in progress. MPEG-V standardizes and specifies the metadata structure for realistic media effect representations and realistic media representation effects, and proposes a description method of writing the metadata for realistic media. In this paper, we present transmission system for realistic media effect representations and studied system modeling for increasing user's sense of reality through realistic media devices.

Keywords: MPEG-V, Realistic Media, Sensory Effect Metadata, Transmission System

1. Introduction

The development of next-generation broadcasting technology, it has been possible to produce diverse types of broadcasting service, different from conventional broadcasting service, and transmit them to users. At present, commercial broadcasting service uses terrestrial, satellite, cable, and IPTV transmission media, offering HD (High Definition) quality service to most regions. However, the broadcasting service based on terrestrial has the limits of bandwidth and receivable range, so that it is difficult to transmit high definition media and a variety of information. To solve the problem, much research has been conducted. As of now dual stream an ATSC (Advanced Television Systems Committee) standard is used to transmit Stereoscopic 3D contents at a specific time and the test broadcasting for 4K-UHD quality is in progress[1].

Unlike terrestrial broadcasting service, cable broadcasting service is able to freely use a transmission bandwidth and thus the test broadcasting for Stereoscopic 3D contents and 4K-UHD quality based on Side-by-Side transmission is already completed, and now the cable broadcasting service is in progress to achieve its goal of offering commercial service[2]. In fact, conventional cable broadcasting service provides nothing but visual and auditory elements and simple contents information. So it tries to improve its visual element from HD level to 4K-UHD level, and increase its auditory element from 5.1 channels up to 22.2 channels. Although the transmission technology for contents information is served, it has difficulty in transmitting a lot of information. Furthermore, unless a receiver supports the transmission technology, it is impossible to watch the information.

Realistic media contents are next-generation media contents which are aimed at increasing users' immersion and presence through the five sense representation effects that are not provided by conventional broadcasting service[3]. Realistic media are based on 4K resolution

*Author for correspondence

images and 10.2 channel audio and additionally deliver the representation effects of the five senses that humans feel. With regard to 4K-level media, constant research on transmission via the next-generation codec HEVC (High Efficiency Video Coding) is being conducted[4]. As of now, for real-time transmission, 4K-resolution image is divided into four equal parts and then compressed by H.264 before transmission[5]. In fact, HEVC time complexity of encoding is very high. So the current system is unable to make a transmission in real time and provides a test broadcasting service through non real-time transmission[6].

The five sense representation effects of realistic media have never been tried by conventional broadcasting systems, but are offered to users in limited space like a theater by its internal transmission system. As such, a method using an internal transmission system is used in circumstances like a theater or an experience hall, but a broadcasting system using conventional transmission technique has difficulty in making a transmission to remote users. To solve the problem, more diverse research needs to be conducted. In this paper, we present the quality-of-service based transmission system modeling in which the five sense representation effects of realistic media are transmitted to remote users by conventional broadcasting systems realistic media effect representations.

2. Related Work

MPEG (Moving Picture Experts Group), an international organization that was formed to set international standards of media contents, defines an interface standard of connecting the virtual world and the real world on the basis of MPEG-V standard[7]. The realistic media effects defined by MPEG-V are diverse physical effects including wind, vibration, light, and temperature. More specifically, part 3 of the seven parts of MPEG-V defined the effects[8].

To represent realistic effects, MPEG-V Part 3 used XML[9]. SEDL (Sensory Effect Description Language) is defined to represent realistic effects, and SEV (Sensory Effect Vocabulary) is used to define actual effects. By defining Class of realistic effects, SEDL defined the information commonly used in all effects. The reason of the separation of SEDL from SEV is to actively extend realistic effects in the future. In other words, when the realistic effects to be added in the future are extended in SEV, they will be able to be applied immediately. The realistic effect metadata generated on the basis of SEDL is defined as SEM (Sensory Effect Metadata) which has the structure as shown in Figure 1. SEM can be used to transmit realistic representation effects via various realistic media contents and to operate and control many different kinds of realistic effect representation devices.

Therefore, users receive the media contents including realistic representation effects which make them feel presence and immersion, rather than the simple contents including visual and auditory elements.

SEM is comprised of the uppermost attributes auto Extraction and Description Metadata, including one of Declarations, Group Of Effects, Effect, and Reference Effect. Auto Extraction presents whether realistic representation effects are automatically extracted from video and audio characteristics. Description Metadata provides the writing information of realistic effect metadata and the reference information to be used in the whole written document.

Declarations is aimed at being simply referred when the five sense effects frequently used in realistic effect contents are previously declared for use, and is used to provide the information necessary to represent the five sense effects with the use of parameters. Group Of Effects is used to include time information in order for a single effect to be produced in multiple representation devices. A system uses the time information to synchronize realistic media contents with realistic representation effect time. Also, a single effect needs to have at least two definitions, and a concrete definition of a representation effect should be described with an additional definition. Effect is used to define a single representation effect and includes time information in itself.

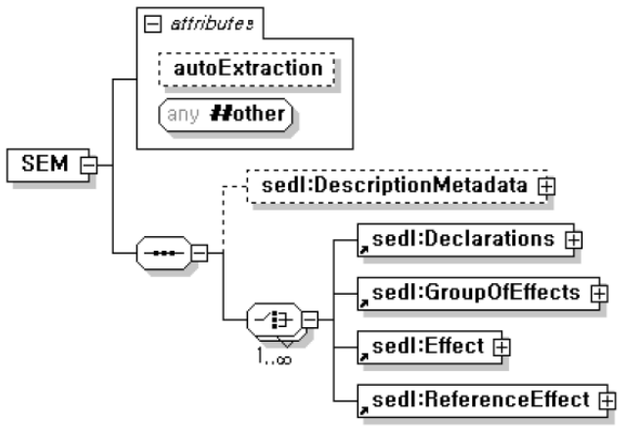

Figure 1. The structure of SEM.

SEV is a XML schema in which each realistic representation effect is defined, and is used together with SEDL. In the case of SEV, the intention of a writer who writes realistic representation effects is not changed by the characteristics of a user's device which receives realistic media contents. The SEV metadata which presents a writer's intended effect in a single or multiple ways are differently presented. SEV metadata are converted into a command to control realistic representation effects and at this time can be presented differently depending on the structure and performance of a realistic representation device. In this sense, SEV metadata are not the data written to control the realistic representation device used, but one written to conceptually present representation effects.

3. System Modeling

The broadcasting system for realistic effect representations goes through the process of editing and saving realistic media contents, the process of acquiring realistic effects and the process of editing the acquired information in various types, and the process of converting to fit them into a realistic effect representation device. The system for realistic effect representations is shown in Figure 2. To serving 4K-UHD contents uses various codecs to process the images captured by a camera supporting 4K resolution and utilizes the highly efficient, next-generation codec HEVC to transmit them on the internet. In the case of 4K resolution contents, they are too big to be operated (contents of a large size) in real time.

Therefore, saved in the distributed file system based storage space and have the replication structure guaranteeing fault tolerance[10]. For fast operation of 4K contents, various interfaces of guaranteeing an I/O speed with high reliability and availability are provided depending on purposes. With the current technology, it is hard to access a 4K resolution image and then insert realistic representation effects into it. For the reason, a method of going through various procedures is used. In the case of 4K resolution image, there is no tool of editing them in real time. So a 4K resolution image is converted into an image with low resolution for real-time edition. In the process, by using EDL (Edit Decision List) information, an editor is able to edit an original image in various ways. Also after making an edition while looking at the converted low resolution image, an editor transfers EDL information as an edition tool to make an original image automatically edited. Realistic representation effects should make it possible to use an edition tool for edition depending on the conditions of contents and then to save in a separate file. And in the case of non 4K-UHD and in the case where an image needs to be transmitted in real time, realistic representation effect data should be written to meet time information when an image is edited.

An image transcoded into HEVC requires a large transmission bandwidth, since its data size is very big. However, the realistic effect metadata providing realistic representation effects has a relatively smaller size than video and audio data. In the case of TS transmitted over UDP/IP, packet loss can occur according to network circumstances. As for realistic effect metadata, any pack loss

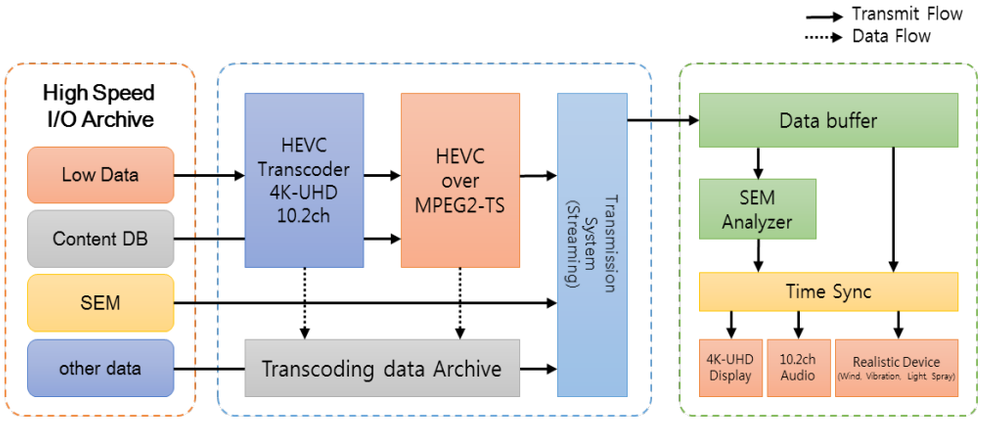

Figure 2. The structure of the transmission system for realistic effect representations.

should not occur. In this sense, such metadata needs to use a protocol guaranteeing reliability. Therefore what is required for realistic effect metadata is to use a separate protocol to transmit the metadata to a user, play contents at the time of receiving TS of contents, and minimize any delay time which may occur to a user.

Video and audio data are decoded by De-Multiplexer, and then are transmitted to each user's device. A sensory analyzer which received realistic effect metadata over a reliable protocol analyzes the structure of realistic effect metadata to send data to each realistic effect representation device. A sensory device control system searches for its connected realistic effect representation devices, and then extracts only representable data to send them to each device and ignores non-representable data. In the decoding process, when data arrives later than the already-input time because of a small network bandwidth and a difficulty in data reception, the sensory analyzer resets the time information on the basis of the last transmitted data and sends the information back to the sensory device control system. When video or audio data is lost, the sensory analyzer analyzes the time information of realistic effect metadata and performs an operation not to transmit data to the realistic media device control system. In so doing, it is possible to provide realistic media contents guaranteeing QoS (Quality of Service).

The sensory analyzer analyzes and classifies the realistic effect metadata received on network, and then sends data to the sensory device control system, as shown in Figure 3. Based on the time information of media contents, a system is divided into a control system for controlling realistic effect motion chair and a control system for controlling stage setting. All realistic representation effect devices but a motion chair is mostly the devices produced for existing stage setting effects. Therefore, the control system for controlling stage setting is also controlled on the basis of the time information of media contents. But, since they are physical devices, there is a margin of error between the time when an operation signal is received and the time when a user feels the effects, and sometimes a screen scene is inconsistent with sensory effects.

Realistic representation effects are produced on the basis of visual and auditory elements so that it is hard to check an error of operation time of a realistic representation device. To solve the problem, it is necessary to predict each device's operation time and provide the least time difference for users. That is a core role that a transmission system for realistic media effect presentations needs to play. Also the system needs to have a signal conversion function to control the interface of each realistic effect representation device. In so doing, the system should have no problem in using the interface of any unspecified devices without any specific conversion process.

4. Conclusion

With the development of next-generation broadcasting technology, realistic media contents will replace conventional broadcasting contents. However, the current technology is not developed enough to easily operate 4K resolution image based on realistic media, and consequently it takes a lot of time and efforts. Although a method using existing broadcasting technology is used to overcome the current challenges, there are still a lot of

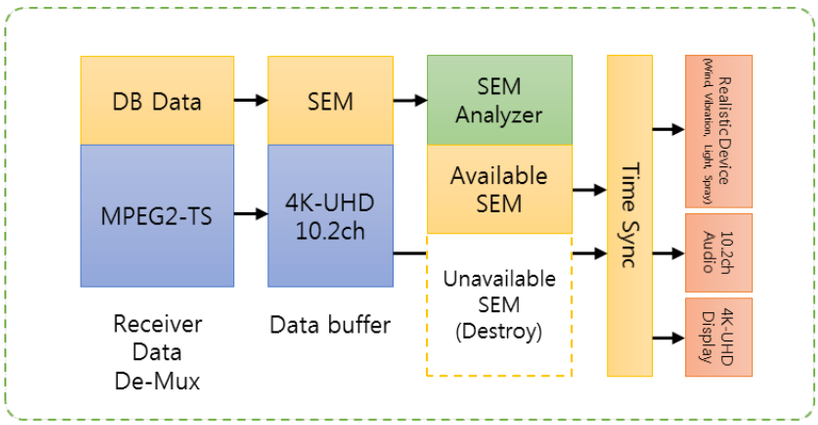

Figure 3. The structure of a realistic effect representation device.

issues to be addressed. According to a report, it takes very long to compress a 4K resolution image for transmission with the next-generation codec HEVC, and as a result it is hard to transmit such an image in real time. Although distributed transcoding technique is used to overcome the challenge, it is still difficult to do so. Furthermore, as a data size becomes large, a storage volume is also on the sharp rise. And an I/O speed of a file system is an important factor in the structure of broadcasting system. As in conventional broadcasting systems, realistic representation effect metadata are so large that research on various encoding techniques has been conducted. However, given the fact that 4K images become relatively too large, a method of sending realistic effect metadata without encoding is more efficient in terms of management and transmission speed.

In this paper, we present modeled the system which uses a protocol guaranteeing reliability to send realistic effect metadata in a separated type. Given that MPEG-V continues to standardize a realistic effect representation method, this researcher will study a transmission system reflecting additional realistic effects and will continue to research a realistic effect representation system applicable to even a small-sized environment.

5. References

1. ATSC, 3D-TV terrestrial broadcasting. Part 2 - service compatible hybrid coding using real-time delivery. ATSC Standard A/104:2012; 2012.
2. Yun K, Lee B, Lee J, Cheong W, Hur N, Kim K. Signaling and multiplexing for service-compatible 3DTV broadcasting. KOSBE. 2011; 16(6):951–5.
3. Yoon K, Choi B, Lee E, Lim T. 4-D broadcasting with MPEG-V. IEEE International Workshop on Multimedia Signal Processing (MMSP); 2010. p. 257–62.
4. Ugur K, Andersson K, Fuldseth A, Bjontegaard G, Endresen LP, Lainema J, Hallapuro A, Ridge J, Rusanovskyy D, Zhang C, Norkin A, Priddle C, Rusert T, Samuelsson J, Sjoberg R, Wu Z. High performance low complexity video coding and the emerging HEVC standard. IEEE Transactions on Circuits and Systems for Video Technology. 2010; 20:1688–97.
5. Schwarz H, Marpe D, Wiegand T. Overview of the scalable video coding extension of the H.264/AVC standard. IEEE Transactions on Circuits and Systems for Video Technology. 2007; 17:1103–20.
6. Gu SM, Seo A, Kim YC. A study on distributed transcoding using dynamic allocation of virtual machines in cloud computing environment. International Conference on Digital Policy and Management; 2013. p. 125–6.
7. Waltl M, Timmerer C, Hellwagner H. A test-bed for quality of multimedia experience evaluation of sensory effects. International Workshop on Quality of Multimedia Experience; 2009. p. 145–50.
8. ISO/IEC FDIS 23005-3 information technology - media context and control - Part 3: sensory information. ISO Publication; 2010 Aug.
9. ISO/IEC FDIS 23005-5 information technology - media context and control - Part 5: data format for interaction devices. ISO Publication; 2010 Aug.
10. Shvachko K, Kuang H, Radia S, Chansler R. The hadoop distributed file system. IEEE 26th Symposium on Mass Storage Systems and Technologies; 2010. p. 1–10.