Design and Implementation of 4D Real-Sense Media Broadcasting Service System for Smart Home

Jae-Kwan Yun', Mi-Ryong Park', Hyun-Woo Oh', Mi-Kyung Han', Jong-Hyun Jang', and Kwang-Roh Park[2],

[1]Meta-Verse&Emotion Technology Research Team, 218 Gajeong-no, Yuseong-gu, Daejeon, Republic of Korea
[2] Green Computing Research Department, 218 Gajeong-no, Yuseong-gu, Daejeon, Republic of Korea
{jkyun, mrpark, hyunwoo, mkhan, jangjh, krpark} @etri.re.kr

Abstract. In this paper, we explained 4D real-sense media broadcasting service system which can give users a lot of special effects (e.g., wind, heat, light, vibration, scent effect, etc.) depending on the 3D contents. Previous 3D media broadcasting service system does not support sensory effects transmission and synchronization when it sends 3D contents. And, also does not follow ISO/IEC 23005-1-6 real-sense media presenting standards. Therefore, we developed 4D real-sense media broadcasting service system that can play 4D real-sense media with multiple real-sense devices to give realistic effects according to SEM. In this paper, we described the system architecture of the 4D real-sense media broadcasting service system and the aggregator, acquiring time methods from an aggregator, time synchronization method between 3D contents and SEM, and implemented system and evaluation result.

Keywords: 4D real-sense media, aggregator, synchronization, SEM

1 Introduction

In recent years, wide range of 3D content has been produced. As a result, request to the 4D real-sense media that can stimulate five senses of human gradually becomes an increasing demand [1]. 4D real-sense media refers to the media that can represent real-sense effects by using real-sense devices like electronics fan, vibration chairs for users to feel displayed effects of 3D contents physically. Depending on the prevalence of smart TV, researches for using 4D real-sense media utilized on the theater, experience hall to home has been actively studied on [2], [3].

In addition, the international standard group MPEG-V (Virtual) in MPEG (Moving Picture Experts Group) makes a lot of efforts to express real world's effects to virtual world (e.g., games, movies) or virtual world's effects to real world. Because methods that express previous 3D contents are standardized even if playback devices are different. But, there was difficult distribution problem of 4D real-sense media; reprogramming of controlling motion chair was necessary if different vendor's motion chairs are used. To solve these problems, real-sense effect definition, avatar or virtual object type definition, device control interlocking and sensor information format that can be expressed in real and virtual world was standardized [4].

This paper is organized as follows. In section 2 concepts of the proposed 4D real-sense media broadcasting service system are explained, in section 3 design issues of the system are discussed, in section 4 results of the implementation and evaluation of the system are shown, and the conclusion of this paper is given in section 5.

2 Concept of the Proposed System

Previous broadcasting service system used the method of real-time streaming with MPEG-2 TS(Transmission Stream) format or transmitting modified images by an authoring tools to transmission server after capturing 3D images with a camera. To play 4D real-sense media by using previous playback system, authoring tool was needed to insert real-sense effects and transmission server to transmit its streams. It was difficult to present real-sense effects of recorded spots because inserted real-sense effects was made by editor's guess independent of the real environments.

Therefore, in this paper, we use a broadcasting system, NVS (Network Video Server) to transmit an encoded 3D image recorded with 3D camera to make the best use of 3DTV to solve these problems. And, we also transmit gathered environmental information of recorded spots by using aggregator which has multiple embedded sensors to 4D real-sense broadcasting server in the form of MPEG-2 TS. Finally, HS (Home Server) in smart home receives transmitted images, and then sends 3D images, sensed environment information to 3DTV and real-sense playback devices respectively to represent them. Fig. 1 shows the system concept of the 4D real-sense media broadcasting service system.

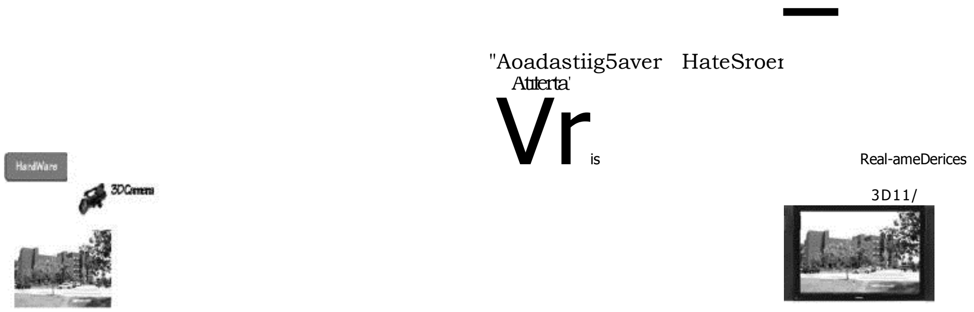

Fig. 1. System concepts of 4D real-sense media broadcasting service system.

3 Design of the Proposed System

3.1 The entire 4D real-sense media broadcasting service system architecture

This proposed 4D real-sense media broadcasting service system can receive more than one media so that it can display 3D contents to 3DTV. In other words, by using more than two cameras, we can obtain multiple audio/videos and encode them into one MPEG-4 file by using NVS. The reason why we used MPEG-4 format to encode 3D images is that it is an international standard to store tracks and metadata to one file. And, when transmitting these media to HS, it is easily converted to MPEG-2 TS format that is an international standard for multimedia transmitting.

NVS can receive SD class of media by using a component cables or HD class of media by connecting HDMI cable, and transfers encoded 4D real-sense media to 4D broadcasting server on wired/wireless network. 4D broadcasting server actually is responsible for transmitting 4D real-sense media to HS and converts MPEG-4 format 4D real-sense media to MPEG-4 over MPEG-2 TS format for transmitting. In addition, while cameras are capturing media, aggregator collects sensory information such as temperature, intensity of illumination, humidity, acceleration, angular speed, amount of gases, wind velocity, location information. And, it describes them with XML format and transfers them to HS periodically.

4D broadcasting server evaluates task scheduling process by synchronizing collected SI (Sensory Information)s with play times of the captured media. Generally, it takes a long time to encode audio/video, but SI takes short time. So, generation time of SI must be delayed for the gap of media encoding time. Media encoding time can differ from according to media encoders. And, 4D broadcasting server multicasts on IP network with UDP/IP protocol. HS receives 4D real-sense media by using UDP/IP receiver and then removes TS headers with MPEG-2 TS demux. Encoded track and SEMs come out after MPEG-2 TS headers are removed, these tracks are transmitted to audio/video devices connected to HS again by TS relay program. Fig. 2 shows architecture of the proposed 4D real-sense media broadcasting service system.

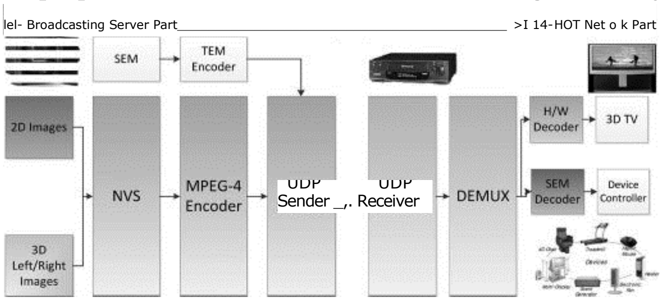

Fig. 2. System concepts of 4D real-sense media broadcasting service system.

3.2 Media Time Scheduling Algorithm

4D broadcasting server compares time of stream transmitted from NVS with SEM's time of start element transmitted ahead of NVS and decides actual order of encoding time of MPEG-2 TS data. It transmits after inserting SEM to MPEG - 2 TS according to the time of filmed image. First, we checked *payload unit start indicator* information within MPEG - 2 TS' header and looked for case that is set by **1.** MPEG-

2 TS' payloads need to be analyzed; there is PES header in case that this value is set to 1. After receiving PTS (Presentation Time Stamp) values from PES (Packetized Elementary Stream), we need to translate PTS to micro-second to compare with the time value of the start element of aggregator (pts=pts*100/9 shown in fig. 3). And then, compare start time in SEM transmitted from aggregator with translated pts. In this process, time correction is needed because there are time gap between images' playing time transmitted from NVS and start time of aggregator's SEM. When correcting time of images and SEM within 10ms, we need to correct time of SEM as much as possible because users can feel bad feelings if there are delays for images. Corrected SEM and images are transferred to HS with MPEG-2 TS format.

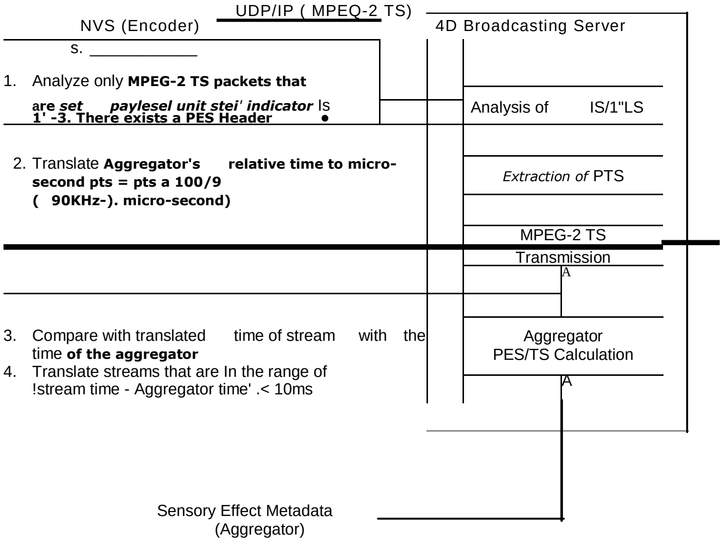

Fig. 3. Media synchronization method in 4D broadcasting server part.

As shown in fig. 4, HS demuxes transmitted MPEG-2 TS. If there is 0x103 of PID in MPEG-2 TS, SEM can be separated from PES after extracting PES from MPEG-2 TS because this PES has SEM inside. The separated SEM is transmitted to a parsing/mapping module of HS to be analyzed, and mapped to a corresponding device type, control type and control value. Real-sense devices are controlled according to the playback time of images. And TS relay program transmits encoded images to H/W Video/Audio Decoder connected with HS.

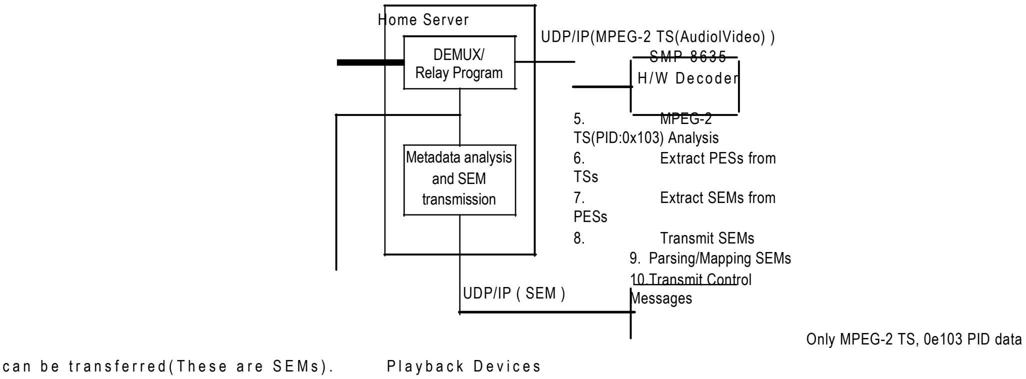

Fig. 4. Media/playback devices synchronization in home network part.

4 Implementation and evaluation of the proposed system

Fig. 5 shows the implemented system of aggregator. Aggregator is connected serial interface for console and debugging with host machine. Host machine can check the aggregator's status whether complete loading the kernel modules or not and sets up the system time by *rtc_set* utility at aggregator. Aggregator is always gathering the whole of the sensor's current status, accumulates the gathered values above the threshold, makes SEM, and sends SEM to 4D broadcasting server.

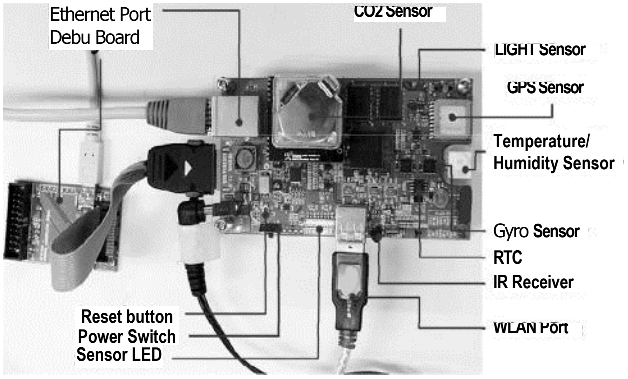

Fig. 5. Implemented aggregator with multiple real-sense sensors.

We examine the whole of the demo system which consists of aggregator, 3D camera, broadcasting server, etc. And we gathered wind, light, temperature information from the aggregator according to the 3D images. NVS encodes the received left/right frames which are divided the left and right screen as the side-by-side method. Table 1 shows the encoded audio and video types. 3D video encoded Full-HD size by mixing left and right screen, and audio encoded stereo type.

Table 1. Encoded Audio and Video Types.

VIDEO		AUDIO	
Codec	**MPEG-2 TS(CBR)**	**Codec**	**MPEG-2 Layer 2**
Width,Height	1920x1080	Bit per sample	**16bit/sample**
Frame rate	29.97	Channels	2 channels
Average bitrate	**17,000KBits/sec**	Average bitrate	3841(Bits/sec

To evaluate the timer accuracy, we calculate the average difference between gathering time at the aggregator and the encoded timestamp in the SEM metadata more than 20 times. In the fig. 6, (+) value means encoded SEM timestamp leads to the gathered time of the aggregator, (-) values aggregator generation time leads to SEM timestamp. Fig. 6 shows time error from 800ms to 20ms after calibration.

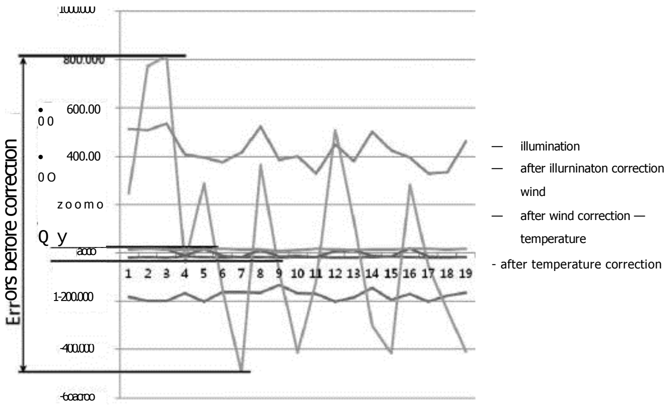

Fig. 6. The results of the evaluation test.

5 Conclusion

Recently, requirements for **4D** real-sense media that give users real-sense effect services like real-sense broadcast, real-sense news, real-sense education arises. To present 4D real-sense media efficiently, there needs methods for acquiring, inserting, time synchronizing, and transferring SIs with captured 3D contents. To do this, in this paper, we explained the system architecture of the 4D real-sense media broadcasting service system, aggregator, acquiring time from aggregator, time synchronization method between captured media and SIs. As shown in the evaluation test, by adapting our algorithms, we reduced time error from 800ms to 20ms after calibration. The future research issue includes interaction related with IPTV or N-Screen services.

Acknowledgments. This work was supported by the Industrial Strategic Technology Development Program funded by the Ministry of Knowledge Economy(MKE, Korea). [10039171, Development of Convergence Media Service Platform for Realistic Experience in Smart Home Environment]

References

1. Park, S.W., Park, N.S., Kim, J.T., Paik, E.H.: Provision of the Expressive Multisensory Adaptation Platform for Heterogeneous Multimedia Devices in the Ubiquitous Home. In: IEEE Transactions on Consumer Electronics, vol. 55, no. 1, pp. 126--131. (2009)
2. Yoon, K.R., Choi, B.S., Lee, E.S., Lim, T.B.: 4-D broadcasting with MPEG-V. In Multimedia Signal Processing (MMSP), pp. 257--262, Saint-Malo. (2010)
3. Yun, J.K., Jang, J.H., Park, K.R., Han, D.W.: Real-Sense Representation Technology Using Multiple Devices Synchronization. In 7th IFIPWG International Workgroup, pp. 343--353, Newport Beach. (2009)
4. Waltl, M., Timmerer, C., Hellwagner, H.: A Test-Bed for Quality of Multimedia Experience Evaluation of Sensory Effects. Proceedings of the First International Workshop on Quality of Multimedia Experience (QoMEX 2009), pp. 145--150, San Diego. (2009).