Repetitive Delivery Scheme for Left and Right Views in Service-Compatible 3D Video Service

Kugjin Yun, Won-Sik Cheong, Jinyoung Lee, Kyuheon Kim, Gwangsoon Lee, and Namho Hur

This paper introduces a novel repetitive delivery scheme for the left and right views in service-compatible (SC) 3D video that provides full backward compatibility to a legacy DTV system while retaining HD 3D visual quality without additional bandwidth or a codec over the legacy broadcasting channel. The proposed SC delivery scheme transmits individual view sequences of a 3D video in interlaced form, that is, a left-view sequence of a 3DTV program to be used repeatedly is transmitted first and stored locally, and the right-view sequence of the 3D program is then transmitted. This paper specifically describes the signaling, synchronization, and storage format methods used to validate the proposed SC delivery scheme. The experiment results show that the proposed SC delivery scheme can be effectively applied for an SC 3DTV service without degrading the DTV quality using only legacy DTV platforms.

Keywords: Stereoscopic, 3DTV, service-compatible, legacy channel, repetitive delivery.

Manuscript received Aug. 30, 2013; revised Dec. 13, 2013; accepted Dec. 30, 2013.

This research was supported by the ICT Standardization program of MSIP (Ministry of Science, ICT & Future Planning), Rep. of Korea, under the title of [2013-PK10-23, Development of standard for Terrestrial and Cable 3DTV broadcasting service], and also under the C-ITRC (Convergence Information Technology Research Center) support program (NIPA-2013-H0301-13-1006) supervised by the NIPA (National IT Industry Promotion Agency).

Kugjin Yun (phone: +82 42 860 1615, kjyun@etri.re.kr), Won-Sik Cheong (wscheong@etri.re.kr), Jinyoung Lee (jinlee@etri.re.kr), Gwangsoon Lee (gslee@etri.re.kr), and Namho Hur (namho@etri.re.kr) are with the Broadcasting & Telecommunications Media Research Laboratory, ETRI, Daejeon, Rep. of Korea.

Kyuheon Kim (Kyuheonkim@khu.ac.kr) is with the Department of Communications and Signal Processing Applications, Kyung Hee University, Yongin, Rep. of Korea.

I. Introduction

Existing 3DTV broadcasting technologies such as MPEG, ATSC, DVB, and TTA have been developed for standard use [1], [2]. The 3DTV service scheme included in these standards can be classified as frame-compatible (FC) or service-compatible (SC), depending on the backward compatibility with the legacy DTV system.

An FC scheme is a key technology that was adopted in first-generation 3DTV broadcasting services [3], [4]. It multiplexes the left- and right-view frames into a single frame configuration as that in a legacy DTV service, which allows the service provider to reuse the legacy DTV production and distribution systems. However, it requires subsampling and decimation of the source images to make them squeezable into one-half the size of an HD frame, and a degraded image quality is hence inevitable. Moreover, owing to the geometrical layout of the composite frame, the presented 3D video will not be compatible with legacy DTV.

To overcome the above-mentioned problems, SC schemes have been developed. An SC scheme provides not only backward compatibility with legacy DTV but also better image quality than an FC scheme provides. The SC hybrid coded (SCHC) scheme was introduced and applied for 3DTV trial service in the Republic of Korea [5]-[7]. In SCHC, left- and right-view sequences are encoded independently using MPEG-2 video and H.264 [8] and are transmitted through various broadcasting channels, such as terrestrial, cable, or satellite, while maintaining not only backward compatibility but also HD quality for both views. Since the scheme uses independent encoders to encode left- and right-view sequences, the left-view sequence encoded by MPEG-2 video is presented in legacy DTV. However, for terrestrial 3DTV broadcasting,

http://dx.doi.org/10.4218/etrij.14.2113.0043

owing to the fact that both views are transmitted over a physical channel, an image quality degradation of the left-view sequence is unavoidable since the bandwidth for the view will be reduced from 17 Mbps to 12 Mbps.

Another SC scheme is SC frame packing (SCFP). This scheme packs two 720 p views into a single 1,080 p frame [9], while one of two views remains unaltered. In this case, if the left- and right-view sequences originate in the 720 p format, no decimation is needed and the reconstructed left- and right-view images will preserve their original resolution. The scheme also provides backward compatibility with an H.264-based DTV receiver using the H.264 encoding parameter for cropping a rectangle, which is used to signal the decoder as to which part of the decoded frame must be output to the display. However, this scheme only provides backward compatibility with an H.264-based DTV receiver since the applied cropping rectangle method is not available in MPEG-2 video. In addition, it may degrade the image quality of legacy DTV supporting a 1,920 × 1,080 resolution as it has a 1,280 × 720 resolution per view.

The SC non-real-time (SCNRT) scheme is the most recently introduced SC scheme [10]. A file containing an additional view (for example, the right view) is downloaded using ATSC NRT technology before the base view (for example, the left view) is broadcast in real time over a legacy terrestrial broadcasting channel. Even though this scheme can provide 1,080i HD quality 3D, it requires an additional virtual channel and system architecture, such as an NRT encoder, for NRT file delivery or download [11]. In addition, it may degrade the legacy DTV quality by assigning a separate virtual channel bandwidth for NRT file delivery, which includes an additional codec for better compression of the additional view.

Consequently, these SC schemes have been developed to provide high-quality 3D with backward compatibility with a legacy DTV receiver over a legacy terrestrial broadcasting channel. However, the limitation of these SC schemes is that they cannot maintain 1,080i HD quality and require an additional codec or system architecture for successful 3DTV service.

In this paper, we propose a novel SC scheme that fully resolves the previously mentioned problems and efficiently transmits stereoscopic view sequences over a legacy DTV channel, in which the individual view sequences for various rebroadcasting 3D programs are repeatedly transmitted. Furthermore, this paper proposes a new method including the signaling, synchronization, and storage format to validate the proposed SC delivery scheme.

There are three main differences between the proposed SC scheme and a legacy SC delivery scheme. The first difference is the use of the bandwidth for 3D video transmission. In SCHC and SCNRT, division of the legacy DTV bandwidth is required since an additional view sequence encoded by H.264 has to be multiplexed and transmitted with the base view. On the other hand, the proposed SC delivery scheme fully utilizes the allocated video bandwidth for the broadcasting channel. The second difference is the use of an additional codec. While SCHC and SCNRT use two different codecs (MPEG-2 video and H.264) to encode stereoscopic view sequences, the proposed scheme uses only one video codec, which is used in legacy DTV. The third difference is that the proposed SC delivery scheme can transmit stereoscopic view sequences on all legacy DTV platforms while guaranteeing the legacy DTV quality up to 1,080i HD, whereas the SCFP scheme supports a maximum of 720 p and can be used only with an H.264-based DTV platform.

The remainder of this paper is structured as follows. Section II presents the proposed SC delivery scheme. Section III shows the experiment results, which are followed by some concluding remarks in section IV.

II. Repetitive Delivery Scheme for Stereoscopic Views

1. Overview

One of the important concerns in providing a viable 3D video service is to feasibly and simply transmit stereoscopic view sequences over a legacy channel without degrading the quality while preserving the allocated bandwidth, backward compatibility, and legacy DTV system. In the proposed SC delivery scheme, stereoscopic view sequences are transmitted over a legacy DTV channel in an interlaced form. This is similar to the concept of general rebroadcasting programs, as shown in Fig. 1.

The 3DTV encoder transmits Program A, Program B (left view), Program C, and Program B (right view) at the scheduled program service time. Programs A and C are normal DTV programs, whereas Program B is a 3D program. When the 3DTV encoder transmits Program B, it does not send full 3D video. Instead, it interlaces the left- and right-view sequences as a separate program in between other programs. When the 3DTV receiver receives Program B (left view), if it is signaled as a 3DTV program, it stores the program in local storage. Initially, however, Program B (left view) can be displayed in 2D only since there is no matched right-view sequence. When the 3DTV receiver receives a right-view sequence of Program B at a later time, it loads the stored left-view sequence and presents the program in 3D. Similarly, if the 3DTV receiver has a previously stored right-view sequence, it can display Program B in 3D when the left-view sequence is received.

Figure 2 shows a diagram of a broadcasting service platform

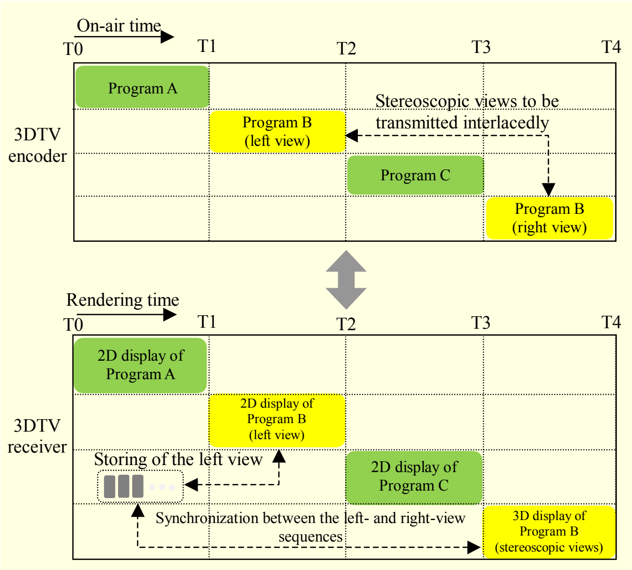


Fig. 1. Concept of repetitive delivery scheme for stereoscopic views.

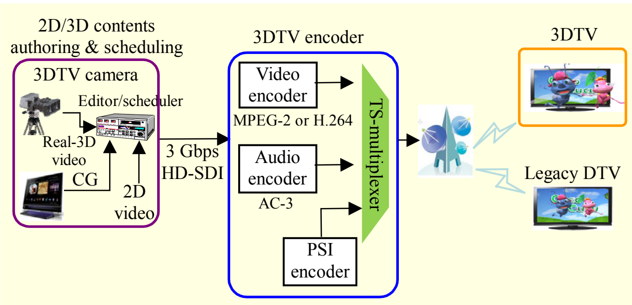


Fig. 2. Diagram of broadcasting service platform based on proposed SC delivery scheme.

based on the proposed SC delivery scheme. This platform is the same as a legacy DTV broadcasting platform. Stereoscopic views are encoded in an interlaced manner using a video encoding method used by the legacy DTV system and transmitted to the transport stream (TS)-multiplexer. The TS-multiplexer generates an MPEG-2 TS [12] including the proposed signaling methods for frame-based synchronization, identification, and association of stereoscopic views to store in the 3DTV receiver. This signaling provides backward compatibility and can be easily updated through a firmware update in the legacy DTV system. That is, the proposed SC delivery scheme has a merit that can be feasibly applied to a legacy DTV broadcasting platform by updating only the legacy TS-multiplexer.

2. Signaling Architecture

To fulfill the proposed SC delivery scheme, a new descriptor and metadata are proposed, as described in Tables 1 and 2. This

Table 1. Stereoscopic_linkage_descriptor.

Syntax
Stereoscopic_linkage_descriptor() {
descriptor_tag
descriptor_length
leftview_flag
filename_length
for (i=0;i<filename_length;i++) {
filename
}
expireDate
}

Table 2. Stereoscopic_pairing_info.

Syntax
Stereoscopic_pairing_info() {
identifier
filename_length
for (i=0;i<filename_length;i++) {
filename
}
frame_number
}

signaling, contained in the program specific information (PSI), is encapsulated into a transport stream to allow the 3DTV receiver to synchronize and store stereoscopic view sequences.

The stereoscopic_linkage_descriptor() in Table 1 is defined for link information between stereoscopic views. In this descriptor, leftview_flag indicates whether the currently broadcasting video elementary stream (ES) is the left view and is used to distinguish individual view sequences when the ES is stored in the 3DTV receiver. In addition, filename_length indicates the length of the filename for individual view sequences to be stored in the 3DTV receiver. The filename is the name to be stored with a newly defined storage format. The term expireDate indicates the expiration time of the stored file in the 3DTV receiver. This descriptor is conveyed in the descriptor loop of the program map table (PMT) for a video ES.

In the proposed SC delivery scheme, the frame-based synchronization between the real-time broadcasting video and the stored file is an important technology for providing a successful 3DTV service. To be specific, each video has a different presentation time stamp (PTS) since the individual view sequences are transmitted in an interlaced form. It therefore requires a new synchronization method.

As shown in Table 2, stereoscopic_pairing_info is used to provide accurate frame-based synchronization, wherein the filename indicates the associated file stored in the 3DTV

http://dx.doi.org/10.4218/etrij.14.2113.0043

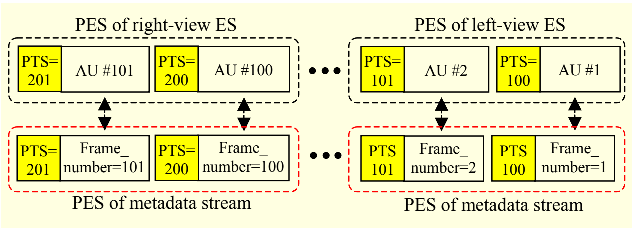


Fig. 3. PES structure of metadata and assignment of frame number for each AU of 3D video ES.

receiver. The term frame_number specifies the frame number assigned for each access unit (AU) of the 3D video ES. This information is conveyed as a metadata stream (stream_type 0x06) in the packetized elementary stream (PES) with the same PTS value of the real-time broadcasting video ES.

Figure 3 shows the PES structure of metadata to assign an identical frame number for each AU of the 3D video ES. In detail, each PES including this metadata has the same PTS (as with each PES of the 3D video ES) to assign the frame number per AU. This method provides accurate synchronization between the stored file and real-time broadcasting video ES more robustly.

3. Storage Format

The proposed storage format has a simple packet structure, which improves the memory efficiency by minimizing the header size for fast access of files in the 3DTV receiver, as shown in Fig. 4.

The AU_frame_number is acquired and recorded by the frame_number value of the stereoscopic_pairing_info. The AU_size provides information about the AU_data length, and AU_data includes an actual left-view stream or right-view stream after de-packetizing a video PES. This stored file is represented using the filename value of stereoscopic_linkage_descriptor.

The proposed format minimizes the parsing complexity and provides rapid access and loading of a stored stream in the 3DTV receiver. This is one of the key requirements to synchronize and display between the stored file and the real-time broadcasting video ES. Table 3 shows the comparison

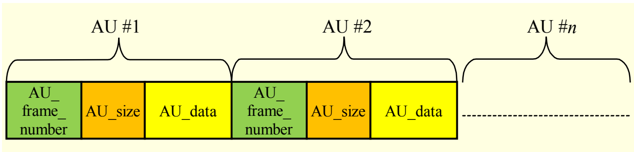


Fig. 4. Proposed storage format structure.

Table 3. Comparison results of access time and storage size.

Storage format	Access time	Storage size	Video
Proposed method	0 ms - 1 ms	24.4 Mbytes	1,920 × 1,080 (300 frames)
MPEG-2 TS	10 ms - 11 ms	25.1 Mbytes	

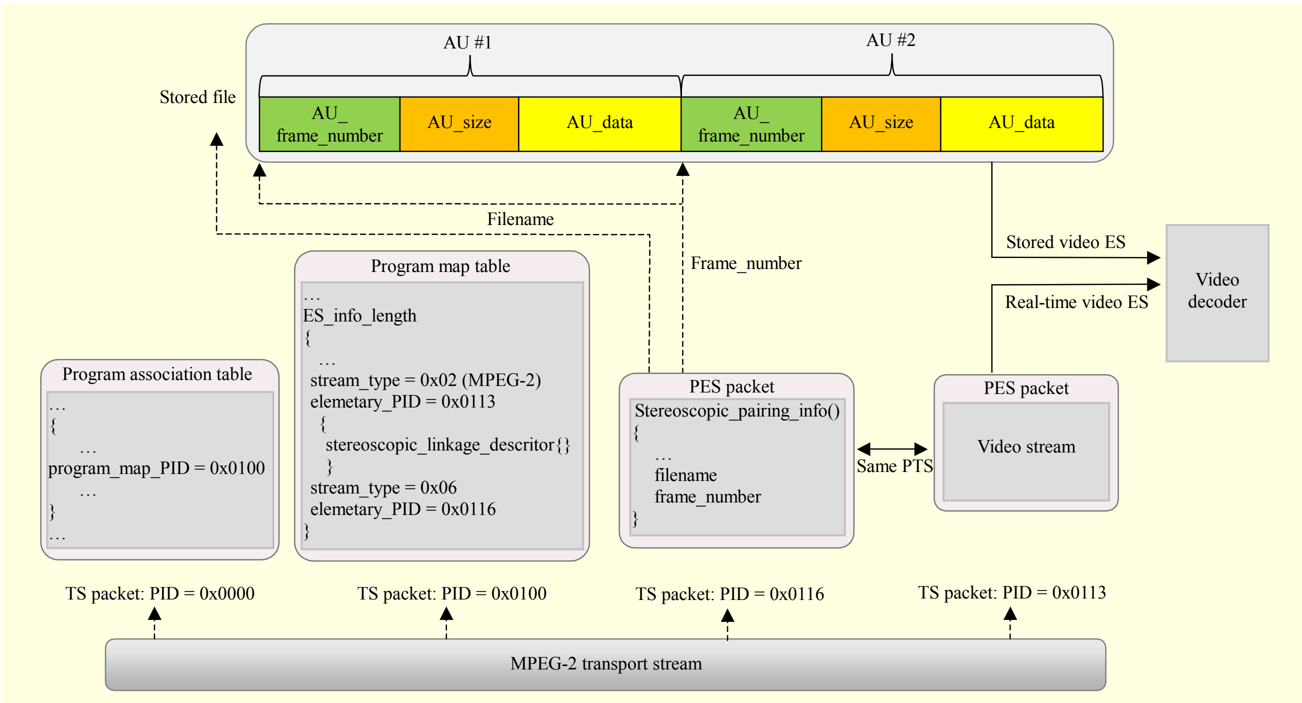


Fig. 5. Process of frame-based synchronization.

Table 4. Test conditions for picture quality evaluation.

SC delivery schemes	Encoding method	Bit rates	Left-view resolution
SCHC	MPEG-2	12 M	1,920 × 1,080
SCNRT	MPEG-2	16 M	1,920 × 1,080
		17 M	1,920 × 1,080
Proposed method	MPEG-2	18 M	1,920 × 1,080

results of the access time and storage size for a stored file, which is encoded by MPEG-2 at 18 Mbps, between the proposed format and MPEG-2 TS, which is used as a storage format in legacy DTV. The access time represents the time in which the first AU is loaded into decoder memory after the parsing of packets.

4. Frame-Based Synchronization

The proposed SC delivery scheme is realized by an arbitrary pairing of real-time broadcasting streams and previously stored files, which do not have the same PTS. It is therefore necessary to provide a novel synchronization method. Figure 5 illustrates the process of frame-based synchronization based on the frame number described in section II.

When the ES for one of the views of the stereoscopic video is being transmitted, the 3DTV receiver recognizes the associated filename and frame number of each AU of the video ES by analyzing the stereoscopic_pairing_information in the metadata stream and comparing the PTS values between the metadata PES and the currently broadcasting video PES. The 3DTV receiver then searches for the previously stored file containing the paired stereoscopic view sequence and synchronizes the file with the currently broadcasting video ES by matching frame_number of the stereoscopic_pairing_infomation, which is assigned for a real-time broadcasting video, and AU_frame_number of the stored file. The proposed synchronization method can be used to robustly perform accurate synchronization between the stored file and real-time broadcasting video ES.

III. Experiment Results

1. Image Quality Evaluation

The delivery of high-quality 3D video, while maintaining legacy HD quality, is an important component for a successful 3DTV service. The primary purpose of this experiment is to verify the preservation of legacy DTV image quality when transmitting a 3D video using the proposed SC delivery scheme.

Table 5. Experiment results for left-view image quality evaluation.

Legacy DTV	SCHC	SCNRT		Proposed method
18 M	12 M	16 M	17 M	18 M
39.2 dB	37.1 dB	38.7 dB	38.9 dB	39.2 dB

The test conditions used to verify the legacy DTV quality between the proposed SC and other SC delivery schemes are described in Table 4. We evaluate the left-view image quality displayed in legacy DTV through the SCHC scheme, SCNRT scheme, and proposed SC delivery scheme. The image quality evaluation is performed using 1,080i HD 3D video sequences under a terrestrial broadcasting environment [13].

The 3D video sequences consist of various types of content, including a documentary, music video, and sports broadcast. The left-view sequence of the SCHC delivery scheme is encoded with MPEG-2 video at 12 Mbps, which was applied to the terrestrial 3DTV trial service in the Republic of Korea. In SCNRT, the left-view sequence is encoded with MPEG-2 video at 16 Mbps and 17 Mbps, considering the required bandwidth for NRT file delivery or download in legacy terrestrial DTV. For the proposed SC delivery scheme, the left-view sequence is encoded with MPEG-2 video at the same coding rate used for legacy terrestrial DTV broadcasting.

The experiment results shown in Table 5 indicate that the peak signal-to-noise ratio of the proposed SC delivery scheme is about 2.1 dB higher than that of SCHC and 0.3 dB to 0.5 dB higher than that of SCNRT.

The image quality degradation of the legacy DTV is unavoidable in SCHC and SCNRT owing to the division of bandwidth to transmit 3D video. Based on this experiment, we confirm that the proposed SC delivery scheme can transmit high-quality 3D video without any degradation in the image quality by maintaining the video coding rate to be the same as that of legacy DTV.

2. Proposed SC Delivery Evaluation

To verify the usability of the proposed SC delivery scheme, the newly developed 3DTV broadcasting system is applied as shown in Fig. 6. The 3DTV encoder is composed of a scheduler for transmitting 2D or 3D video sequences according to their schedule, an MPEG-2 video and AC-3 encoder for AV encoding, and a TS-multiplexer including a PSI encoder. The 3DTV receiver is composed of a TS-demultiplexer to parse the TS, an MPEG-2 and AC-3 decoder for AC decoding, a file generator to store and load left-view streams or right-view streams, a sync-manager to synchronize between a real-time broadcasting stream and previously stored files, and a renderer

http://dx.doi.org/10.4218/etrij.14.2113.0043

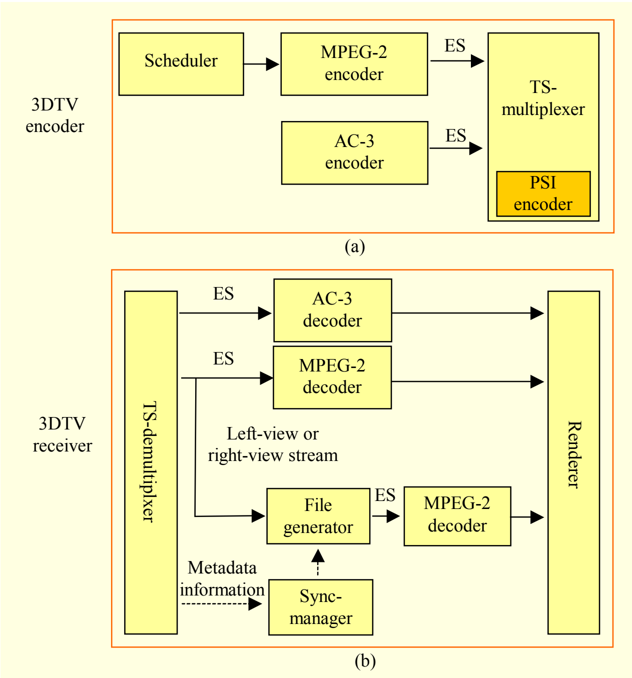


Fig. 6. Structure of developed 3DTV broadcasting system: (a) 3DTV encoder and (b) 3DTV receiver.

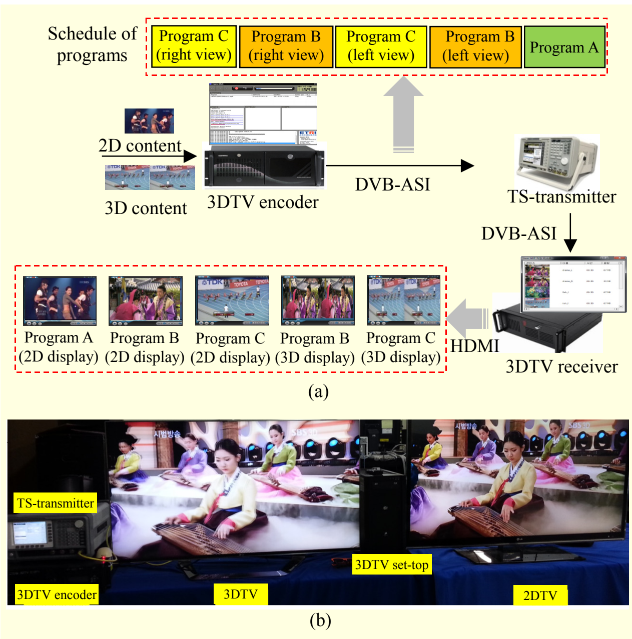


Fig. 7. Experiment results: (a) developed system and 3DTV service using proposed 3D delivery scheme and (b) snapshot of backward compatibility with legacy DTV.

for the presentation.

To verify the 3DTV broadcasting service using the proposed SC delivery scheme, this experiment has progressed under similar conditions as terrestrial DTV broadcasting. The results in Fig. 7(a) verify the 3D video transmission and reception quality using the proposed SC delivery scheme. Figure 7(b) shows the backward compatibility of the developed system with legacy DTV.

Based on the experiment results, the proposed SC delivery scheme provides a simple and efficient 3D video transmission, but with better image quality than other SC schemes while maintaining backward compatibility with a legacy DTV system. In addition, we confirm that it can be easily combined with a legacy DTV platform by updating only the legacy TS-multiplexer.

IV. Conclusion

In this paper, we proposed a novel SC delivery scheme that provides full backward compatibility with a legacy DTV system while retaining the HD 3D visual quality without additional bandwidth or a codec over a legacy broadcasting channel. In terms of maintaining legacy DTV quality, the experiment results show up to a 2.1 dB quality improvement over SCHC and SCNRT. Furthermore, we confirmed that it provides not only a simpler delivery scheme than other SC schemes but also an efficient SC 3DTV service using only a legacy DTV system. We confirm that this SC delivery scheme can be directly applied to a 3DTV service for various rebroadcasting 3D programs that are repeatedly transmitted, such as advertisements, dramas, movies, and so on. Since the proposed SC delivery scheme transmits individual view sequences of a 3D video at different times, in our future work, we will research how to achieve further gain by comparing it to other SC schemes in terms of cost and time consumption.

References

[1] T. Schierl and S. Narasimhan, "Transport and Storage Systems for 3-D Video Using MPEG-2 Systems, RTP, and ISO File Format," *Proc. IEEE*, vol. 99, no. 4, Apr. 2011, pp. 671-683.

[2] ATSC PT1-029r0, "ATSC Planning Team on 3DTV: Interim Report, Part 1 – Visual Sciences," Feb. 28, 2011. http://www.atsc.org/pt1/pt-1-Interim-Report.pdf

[3] N. Hur et al., "An HDTV-Compatible 3DTV Broadcasting System," *ETRI J.*, vol. 26, no. 2, Apr. 2004, pp. 71-82.

[4] Anthony Vetro, "Frame Compatible Formats for 3D Video Distribution," *17th IEEE Int. Conf. Image Process.*, Hong Kong, China, Sept. 26-29, 2010, pp. 2405-2408.

[5] S. Park et al., "A New Method of Terrestrial 3DTV Broadcasting System," *60th Annu. Meeting IEEE Broadcast Symp.*, Oct. 2010.

[6] K. Yun et al., "Efficient Multiplexing Scheme of Stereoscopic Video Sequences for Digital Broadcasting Services," *ETRI J.*, vol. 32, no. 6, Dec. 2010, pp. 961-964.

[7] N. Hur et al., "3DTV Broadcasting and Distribution Systems,"

IEEE Trans. Broadcast., vol. 57, no. 2, June 2011, pp. 395-407.
[8] ITU-T Rec. MPEG-4 AVC | ISO/IEC 14496-10 AVC, *Advanced Video Coding for Generic Audiovisual Services*, Mar. 2003.
[9] G. Ballocca et al., "Tile Format: A Novel Frame Compatible Approach for 3D Video Broadcasting," *IEEE Int. Conf. Multimedia Expo*, Barcelona, Spain, July 11-15, 2011, pp. 1-4.
[10] B. Lee et al., "Adaptive Pre-/Post-Filters for NRT-Based Stereoscopic Video Coding," *ETRI J.*, vol. 34, no. 5, Oct. 2012, pp. 666-673.
[11] A/103, *ATSC Standard: Non-Real-Time Content Delivery*, May 9, 2012.
[12] ISO/IEC 13818-1, *Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: System*, June 2013.
[13] A/53, *ATSC Digital Television Standard, Parts 1 — Digital Television System*, Aug. 7, 2013.

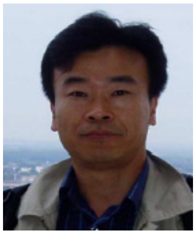
Kugjin Yun received his BS and MS degrees in computer engineering from Chonbuk National University, Rep. of Korea, in 1999 and 2001, respectively. He joined ETRI, Rep. of Korea, in 2001. Currently, he is a senior member of the research staff in the Realistic Broadcasting Research section of the Broadcasting Systems Research Department. He has been a PhD candidate in electronics engineering at Kyung Hee University since 2011. His research interests include MMT/DASH, 3DTV broadcasting, and realistic digital broadcasting.

Won-Sik Cheong received his BS, MS, and PhD degrees in electronics and electrical engineering from Kyungpook National University, Daegu, Rep. of Korea, in 1992, 1994, and 2000, respectively. He joined ETRI, Rep. of Korea, in 2000. Currently, he is a principal member of the research staff in the Realistic Broadcasting Research section of the Broadcasting Systems Research Department. His research interests include 3DTV broadcasting systems, interactive multimedia broadcasting systems, video and image coding, and digital signal processing.

Jinyoung Lee received his BS, MS, and PhD degrees in electrical engineering from Michigan State University, USA, in 1998, 1999, and 2008, respectively. He joined ETRI, Rep. of Korea, in 2004. Currently, he is a senior member of the research staff in the Realistic Broadcasting Research section of the Broadcasting Systems Research Department. His research interests include 3DTV broadcasting systems, MMT/DASH, video source coding, multimedia broadcasting systems, and digital signal processing.

Kyuheon Kim received his BS degree in electronics engineering from Hanyang University, Rep. of Korea, in 1989 and his MPhil degree and PhD in electrical and electronics engineering from Newcastle University, Newcastle upon Tyne, UK, in 1996. From 1996 to 1997, he was with Sheffield University, UK, as a research fellow. From 1997 to 2006, he worked as the head of the Interactive Media Research Team at ETRI, Rep. of Korea, where he standardized and developed the T-DMB specification. He was the head of the Korean delegates for the MPEG standard body from 2001 to 2005. Since 2006, he has conducted research at Kyung Hee University, Rep. of Korea. His research interests include interactive media processing, digital signal processing, and digital broadcasting technologies.

Gwangsoon Lee received his MS and PhD degrees in electronics engineering from Kyungpook National University, Daegu, Rep. of Korea, in 1995 and 2004, respectively. He joined ETRI, Rep. of. Korea, in 2001, and he is currently the director of the Realistic Broadcasting Research section of the Broadcasting Systems Research Department. His current research interests include 3DTV broadcasting, multi-view signal processing, and realistic digital broadcasting.



Namho Hur received his BS, MS, and PhD degrees in electrical and electronics engineering from Pohang University of Science and Technology (POSTECH), Pohang, Rep. of Korea, in 1992, 1994, and 2000, respectively. He is currently with the Broadcasting and Telecommunications Media Research Laboratory, ETRI, Daejeon, Rep. of Korea. He is the manager of the Broadcasting Systems Research Department in ETRI. He was an executive director of the Association of Realistic Media Industry (ARMI), Rep. of Korea. ARMI was established to promote realistic media industry, including the 3DTV and UHDTV broadcasting. Also, he has been an associate professor with the Department of Mobile Communications and Digital Broadcasting, University of Science and Technology (UST), Rep. of Korea, since September 2005. For the collaborative research in the area of multi-view video synthesis and the effect of object motion and disparity on visual comfort, he was with Communications Research Centre Canada (CRC) from 2003 to 2004. His main research interests are digital broadcasting systems, including next-generation DTV, RF/IP converged access technology over HFC, and mobile and 3DTV broadcasting.