Comparisons of Mobile TV Standards and Development of 3-D Mobile TV based on DMB

Heung-Nam Kim*[a], Soo-In Lee [b], Jinwoong Kim [b]

[a] Research Laboratory of Electronics, MIT, Cambridge, MA, USA 02139-4307;

[b] Broadcasting System Research Dept., Electronics and Telecommunications Research Institute, Taejeon, Korea 305-700

ABSTRACT

Many countries have developed different mobile TV standards such as DVB-H, DMB, MediaFLO, OneSeg, and ATSC-M/H. In this paper, we compare different mobile TV standards from the viewpoint of video and audio encoding, data encapsulation, transmission network, and advantages and disadvantages of each technology. Also we propose future mobile TV applications 3-D mobile TV to provide backward compatibility with DMB and better bit rates.

Keywords: mobile TV standard; 3-D audio video encoding decoding;

1. INTRODUCTION

Mobile TV is the transmission of TV programs or video for the devices such as cellular phones, PDAs, and wireless multimedia devices. Mobile reception of broadcasting services has recently received much attention worldwide. Digital Media Broadcasting (DMB), Digital Video Broadcasting-Handheld (DVB-H), MediaFLO are such examples. Among them, commercial service of Terrestrial DMB (T-DMB) was launched in Korea for the first time to provide mobile multimedia services in 2005. Multimedia-capable mobile devices are becoming the core of the portfolio of various multimedia information generation and consumption platforms. Many countries have developed different mobile TV standards such as DVB-H, DMB, MediaFLO, OneSeg, and ATSC-M/H. In this paper, we compare different mobile standards from the viewpoint of video and audio encoding, data encapsulation, transmission network, and advantages and disadvantages of each technology.

Various acquisition and display technologies are also being developed to meet ever-increasing demand of users for higher quality multimedia content. Reality is one of the major references in judging high quality. Providing mobility and increased reality in multimedia information services is thus a promising direction for the future. Specifically, 3-D service over DMB system is attractive due to the facts that (1) glassless 3-D viewing with small display is relatively easy to implement and more suitable to single user environments like DMB, (2) DMB is a new media and thus has more flexibility in adding new services on existing ones, (3) 3-D multimedia handling capability of 3-D DMB terminals has lots of potential to generate new types of services if it is added with other components like built-in stereo camera.

In this paper, we propose attractive future mobile TV applications, i.e., 3-D mobile TV to provide backward compatibility with DMB and better bit rates. In section 2, we compare various mobile TV standards. In section 3, 3-D mobile TV technologies based on DMB providing backward compatibility and more efficient bit rates are proposed. Finally, in Section 4 we draw the conclusions.

2. COMPARISONS OF MOBILE TV STANDARDS

The mobile broadcast transmission can be serviced through terrestrial medium like a terrestrial digital TV standards such as DVB-T in Europe, ATSC in U.S.A., and ISDB-T in Japan [1]. However, those standards cannot be applied directly to mobile TV due to the tiny screen, the limited power consumption, and the characteristics of mobility.

*hnkim@mit.edu; phone 1 617 253-8143; fax 1 617 253-7302;

To resolve these problems, many standards have been developed in many countries such as DVB-H in Europe based on DVB-T, DMB in Korea and China based on DAB, OneSeg in Japan based on ISDB-T, and recently ATSC-M/H in U.S.A. based on ATSC. And also Qualcomm developed its proprietary solution, MediaFLO [2][3]. All of these mobile TV broadcasting systems have similar structures which consist of audio/video encoding, multiplex and transport, and transmission systems like below. However, each standard developed its own technologies to provide more efficient bit rates, less power consumptions, and better mobility. We will compare these standards from the view point of A/V encoding, data encapsulation and transport format, channel coding, and modulation technology.
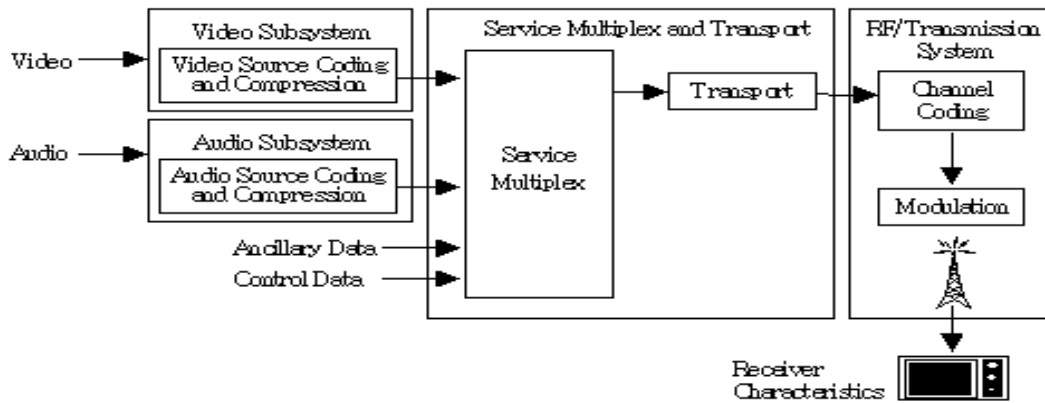

Fig. 1. Structure of mobile TV broadcasting system

2.1 DVB-H (Digital Video Broadcasting – Handheld)

DVB-H was standardized by the DVB and the ETSI under EN 302 304 in November 2004, and the first commercial service was launched by La3 in Italy, in June 2006. DVB-H uses more efficient bit rate encoding schemes such as H.264 in video and MPEG-4/AAC in audio. H.264 which is MPEG-4 part 10 in ISO/IEC provides about 3 times better efficiency of bit rates than MPEG-2. Also DVB-H uses MPEG-2 TS as transport format.

There are two main technology innovations: MPE-FEC and time slicing. To overcome unpredictable signal reception in mobile environments, DVB-H provides more robust forward error correction in channel coding. It uses Reed-Solomon code RS(255,191) which has 64 parity bytes comparing RS(204,188) in DVB-T. Also DVB-T is not suited for high speed moving environment due to Doppler shift of frequency. So DVB-H uses 4K carriers instead of 8K carriers which is used in DVB-T OFDM modulation.

To save power consumption DVB-H developed a special technology called "time slicing." The receiver switches off power of the tuner for up to 80% of the time while playing video continuously. However, the disadvantage of this technology is that it takes more time to switch channel from one to another. Usually it takes 5 sec compared to 1.5 sec of other standards.

DVB-H is designed to use the existing DVB-T network, so it uses the same spectrum as DVB-T. However, the spectrum for DVB-H might be an issue in many countries because the regulators already allocated the available spectrum for the terrestrial digital TV services. The bandwidth of single DVB-H carrier is at most 8MHz, and it can carry between 20 and 40 video and audio services.

2.2 DMB (Digital Multimedia Broadcasting)

DMB was developed by our institute, ETRI (Electronics and Telecommunications Research Institute) in 2004 based on DAB (Digital Audio Broadcasting) standard which was initiated by Eureka 147 Project. The world's first broadcast mobile TV service was launched in May 2005 by SK Telecom in Korea. DMB standard has been approved by ETSI in June 2005 and by ITU in December 2007.

DMB uses the same video encoding scheme as DVB-H, H.264, however, it uses a different audio encoding technology, MPEG-4 BSAC (Bit Sliced Arithmetic Coding). Also it uses MPEG-2 TS as transport format, and MPEG-4 BIFS (BInary Format for Scenes) for auxiliary data. DMB uses Reed-Solomon code RS(204,188) which has 16 parity bytes, and it uses OFDM modulation. While the DMB does not have the function of time slicing, it deals with a lower bandwidth of 1.54MHz for keeping the tuner power low.

The great advantage of DMB is the availability of the spectrum in Europe and Asia because DMB service is a modification of DAB which uses the same 1.537MHz carriers and spectrum. DMB utilizes band III and L-band where it is easier to get available spectrum than the UHF band which is in use for analog and digital television. Also the cost to initiate DMB is lower than the other standards that require for developing new infrastructure because DMB is based on the DAB and DMB can use the existing DAB infrastructure.

2.3 MediaFLO (Forward Link Only)

MediaFLO is a proprietary technology of Qualcomm to provide high-quality video and audio using the Qualcomm band of 700MHz. The first service was launched by Verizon Wireless in March 2007, and the second service was launched by AT&T Mobility in May 2008.

MediaFLO provides multiple types of encoding schemes such as H.264, MPEG-4, WMV, and RealVideo. MediaFLO has a multiple-level error correction system and efficient coding which provides efficiencies of 2bps per Hz to allow 12Mbps with 6MHz. It also provides the layered modulation scheme for a high-quality service at 30 fps to be degraded gracefully at 15 fps instead of freezing due to higher distance to receiver or noisy environment.

MediaFLO uses the OFDM to simplify the reception from multiple cells, and the handsets receive the same packet from the multiple cells which can improve the reception. In addition to saving power consumption by only accessing the part of the signal being watched, MediaFLO handsets can provide fast channel switching time which is less than 1.5 sec.

2.4 OneSeg

In Japan, mobile TV based on ISDB-T (Integrated Services Digital Broadcasting – Terrestrial) has been serviced since April 2006. ISDB-T is a Japanese proprietary standard for digital TV, and each channel of ISDB-T is divided into 13 segments. HDTV broadcasting service uses 12 segments, and mobile TV service uses the remaining 13th segment which is the reason why it is called OneSeg.

OneSeg uses H.264 as video coding, MPEG-4 AAC as audio coding, and MPEG-2 TS as its transport format. OneSeg which has the bandwidth of 433KHz can support 312Kbps which consists of 180Kbps video, 48Kbps audio, and 80Kbps data.

The advantage of OneSeg is that every digital TV channel signal can be broadcasted to HDTV devices and mobile TV devices simultaneously using the same transmission infrastructures. However, the disadvantage of OneSeg is that ISDB-T service is required for mobile TV services. If a country decided on a digital TV standard other than ISDB-T, it would be hard for the country to change its digital TV standard to ISDB-T only for the purpose of mobile TV services. This is the reason why only Japan and Brazil adopted OneSeg standard as mobile TV services.

2.5 ATSC-M/H (Advanced Television Systems Committee-Mobile/Handheld)

ATSC-M/H is one of the candidates for the future as mobile TV standard. ATSC-M/H provide backward compatibility with ATSC by using 8VSB (8-level Vestigial SideBand) modulation. This modulation does not provide robustness against multi-path interference and frequency shift by Doppler effect in a mobile environment. To resolve these problems additional technologies are being developed such as additional error correction scheme, synchronization of transmission from multiple towers, and power savings in handheld devices. The field test of mobile TV will be held in 2009 by broadcasting and consumer electronics company.

The advantage of ATSC-M/H is that HDTV and mobile TV can be serviced simultaneously using same infrastructures and contents without additional cost so that ATSC-M/H service can be free while the competing technology, MediaFLO, is subscription-based.

Table 1 shows the comparative summary of four mobile tv standards including DVB-H, DMB, MediaFLO, and OneSeg. We excluded ATSC-M/H from this table because standardization process of ATSC-M/H is going on.

Table 1. Comparisons of Mobile TV Standards

Standards	DVB-H	DMB	MediaFLO	OneSeg
Base Technology	DVB-T	DAB	Qualcomm CDMA	ISDB-T
Standardization body	DVB	WorldDMB	FLO Forum	ARIB in Japan
Video codec	H.264	H.264	Enhanced H.264	H.264
Audio codec	MPEG-4 AAC	MPEG-4 BSAC	MPEG-4 AAC	MPEG-2 AAC
Transmission schema	OFDM	OFDM	OFDM	BST-OFDM
Power saving tech.	Time slicing	Bandwidth reduction	Time slicing	Bandwidth reduction
Channel switching time	5 sec.	1.5 sec.	1.5 sec.	1.5 sec.
Tech. for better mobility	MPE-FEC, RS(255,191)	Forney interleaving RS(204,188)	Turbo code & RS	Viterbi coding RS(204,188)
Frequency bands	UHF	VHF(Band III)/ L-band	VHF/UHF/ L-band/S-band	VHF/UHF
Bandwidth	5/6/8 MHz	1.54 MHz	5/6/7/8 MHz	433 KHz / 1segment
Data rate (Mbps)	15	1.2	11.2	0.312 / channel
Service country	Europe, USA, Asia	Korea, China, Europe	USA	Japan, Brazil
Characteristics	Nokia is leading	Easily migrated from DAB	Qualcomm is leading	HD/SD/Mobile service in 1 channel

3. 3-D MOBILE TV SERVICE

In this paper, we propose a cost-effective method of 3-D mobile TV services based on DMB system. One of the most important requirements of the 3-D mobile TV is the backward compatibility with the existing DMB system. The backward compatibility means that the existing DMB receivers can identify a 3-D DMB signal and recover the 2-D DMB audio-video signal without any problem. To satisfy such a requirement, we define a new object descriptor (OD) of MPEG-4 Systems. Specifically we regard a pair of stereoscopic video signals and multi-channel 3-D audio signals as two objects. Then we consider that the two elementary streams (ESs) of (V_l,V_a) constitute the video object and similarly, the two ESs of (A_s,A_a) constitute the audio object. To verify the proposed scheme of 3-D mobile TV services over DMB, we implemented a prototype of 3-D mobile TV system. Through the experimental results, we show that the proposed system satisfies the required backward compatibility, provides acceptable depth impression even under the limited bandwidth of 3-D video (allocated less than or equal to 0.8 Mbps) and a small-sized auto-stereoscopic 3-D LCD display.

3.1 Proposed method of 3-D mobile TV based on DMB

Figure 2 shows a conceptual block diagram of the 3-D DMB system consisting of a DAB system, MPEG-4 Systems, and MPEG-2 Systems. Figure 2 also shows the added (shaded) blocks in the media processor to deliver 3-D services over DMB system. The proposed 3-D DMB media processor consists of 3-D video encoding, 3-D audio encoding, MPEG-4 Systems encoding, MPEG-4 over MPEG-2 encapsulator, and channel coding.
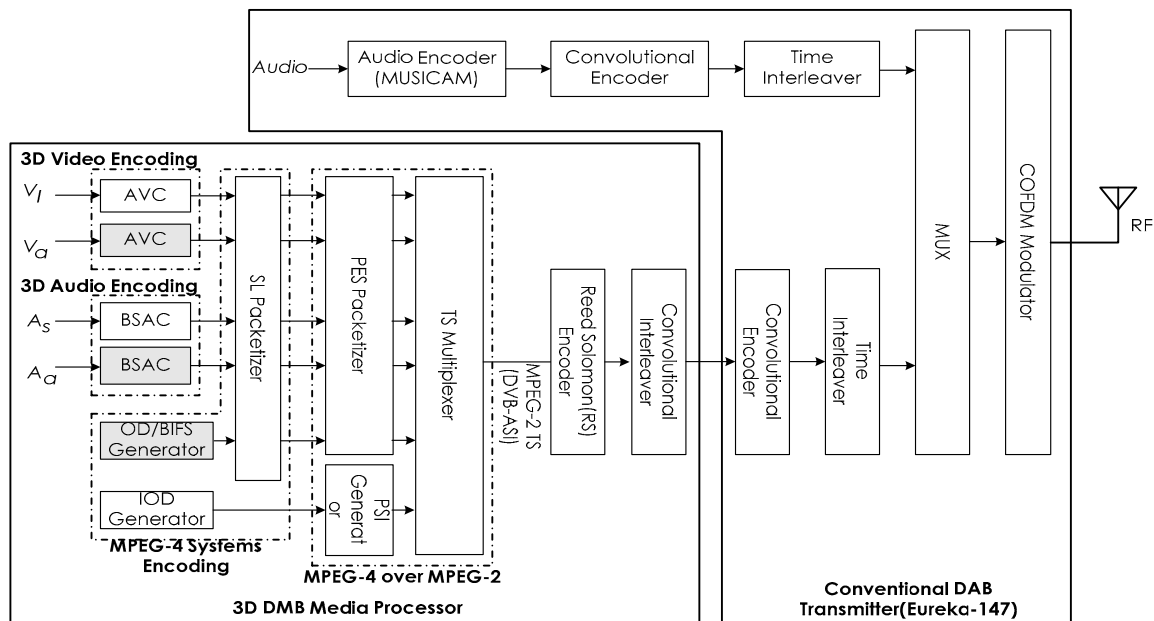
Fig. 2. Block diagram of 3-D mobile TV based on DMB

As input signals, we are considering a pair of stereoscopic video signals and multi-channel 3-D audio signals. In this paper, the left and right (additional) video signals are denoted by Vl and Va respectively. Similarly, the multi-channel 3D audio signals are denoted by As and Aa where As denotes the normal stereo audio signals for the sake of simplicity.

The 3-D video encoding is performed with two separate AVCs. In order to guarantee backward compatibility, we fix the size of Vl to 320x240(QVGA). On the other hand, two other sizes of Va, 160x240 and 160x120, are tested to find the best compromise between bit-rate overhead and quality of the depth perception. We need to reduce the horizontal resolution to a half, since the interleaved 3-D image will be used for 3-D LCD.

Similarly, the 3D audio encoding is performed by applying two separate BSACs for As and Aa, respectively. Here. the multi-channel audio signals are acquired with a special 3D microphone and then mixed based on the specification described in ITU-R Rec. BS. 775-1 [4].

The MPEG-4 Systems encoding is to generate MPEG-4 initial object descriptor (IOD) as well as OD/BIFS. Next, it performs Sync Layer (SL) packetization for 3D audio-visual elementary streams and OD/BIFS based on MPEG-4 Systems. The MPEG-4 over MPEG-2 encapsulator converts SL packets to MPEG-2 TS packets. Note that the program specific information (PSI) is also utilized in making MPEG-2 TS packets in the same block.

Now we explain the crucial idea of describing AV objects in MPEG-4 Systems in more detail. Let us assume that in the case of 3-D DMB system, there are four objects in total, i.e., Vl, Va, As, and Aa. To meet the backward compatibility, we propose a scheme of using two ODs, each OD consisting of two ESs. Additionally, two ODs are assumed to be independent, but two ESs in each OD are assumed to be dependent. The dependence of ESs can be indicated simply by assigning 'TRUE' boolean value to StreamDependenceFlag and by assigning the same ES_IDs to dependsOn_ES_ID in ES_ Descriptor. Next, according to the definition of the MPEG-4 Systems, we assign 0x21 and 0x40 to object-TypeIndication of Vl and As in Decoder-ConfigDescriptor, respectively. On the other hand in the case of Va and Aa , we assign 0xC0 and 0xC1 to the object-TypeIndication indicating user private in MPEG-4 Systems, respectively. Note that the dependent ESs of Va and Aa are ignored with the current 2-D DMB receivers, but are identified with the proposed 3-D DMB receiver.

We produced 3-D DMB bit-streams according to the new syntax and semantics explained above, and tested them on our prototype system. We have verified that the proposed system satisfies the required backward compatibility with the 2-D

DMB system. As we mentioned previously, the conventional DMB receiver ignored the elementary streams for the additional video (Va) and the additional audio (Aa) because 'OD/BIFS Decoder' in 'MPEG-4 Systems Decoding' block does not identify the objectTypeIndication of ODs.

3.2 Video coding and experimental results

For 3-D video encoding in our 3-D DMB system, we propose a residual-downsampled stereoscopic video coding (RDSSVC) which downsamples the residual data based on H.264/AVC. In order to guarantee monoscopic 2-D video service over conventional DMB system, left-view images are coded with temporal-only prediction structure without downsampling but right-view images are coded with spatio-temporal prediction structure on H.264 shown in Fig.3.

The motion and disparity are estimated and compensated for the original resolution. The residual data representing the prediction errors are downsampled before the transform and quantization. In decoding, after the inverse quantization and inverse transform, the reconstructed residual data are upsampled to the original resolution and compensated with the pixels in the correspondent blocks on the reference pictures which have already been decoded. Left-view sequence is encoded and decoded following exactly H.264 specification. By down-sampling the residual data before transmission and then up-sampling in the receiving side, apparent visual quality degrades in return for bit savings. However, final perceptual quality of the stereoscopic 3D video will be about the same as 'full bit rate' version due to 'additive nature' of human stereopsis [5]. Using RDSSVC we could reduce the resulting bit rate while maintaining the perceptual quality of the stereoscopic video.
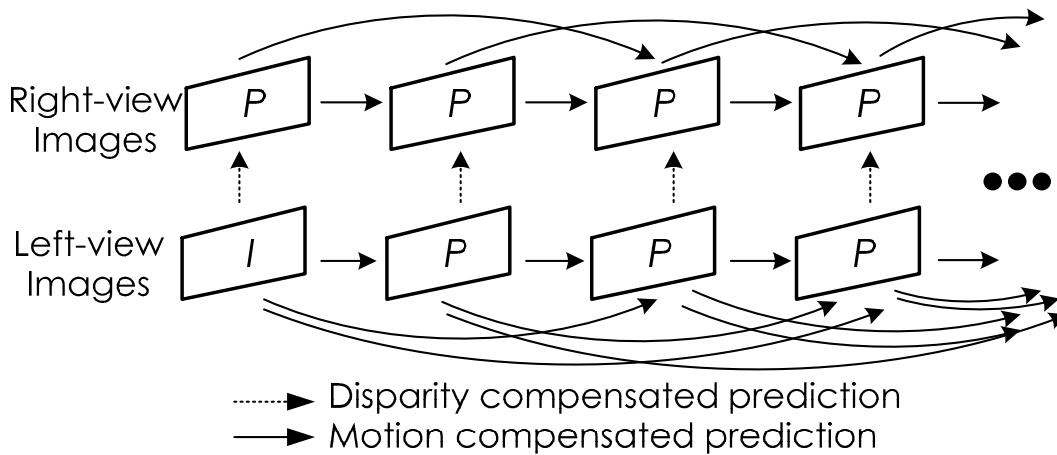

Fig. 3. Structure of reference frames in stereoscopic video coding

Fig. 4. One frame for 'Diving'(left) and 'Soccer'(right) sequences

It is also important to evaluate how much gain is improved by the proposed stereoscopic video coding. Hence, we compared the performance of the proposed coding method against Simulcast and the conventional SSVC (StereoScopic Video Coding). The three coding methods were implemented using the JSVM (joint scalable video model) 4.4 reference software of H.264/AVC. We used two sequences of 'Diving' and 'Soccer' of 320 x 240 (QVGA) with 240 frames as test sequence; they are captured by our stereoscopic camera. Figure 4 shows one frame for 'Diving' and 'Soccer' sequences, respectively.

In Simulcast, the images of left- and right-views are encoded as I-pictures at every second and the rest images are encoded as P-pictures. In the temporal prediction, the three closest preceding frames are referenced as shown in Fig. 3.

For the proposed and conventional stereoscopic video coding, left-view images are encoded as I-pictures at every second, the rest images and all right-view images are encoded as P-pictures. In P-picture type, the structure of reference frames is shown in Fig. 3. The coding method for left- view images is exactly same as that of Simulcast. On the other hand, the coding of right-view images includes temporal prediction from the two closest preceding frames and spatial prediction from one left view frame of the same time in both proposed and conventional stereoscopic video coding.
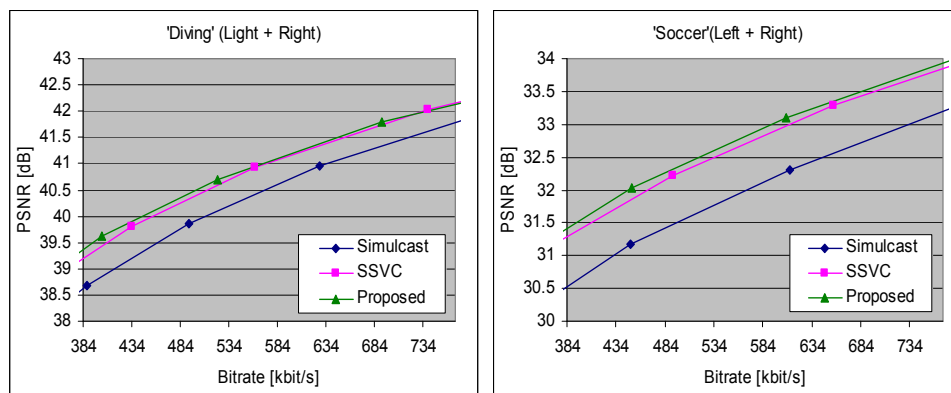

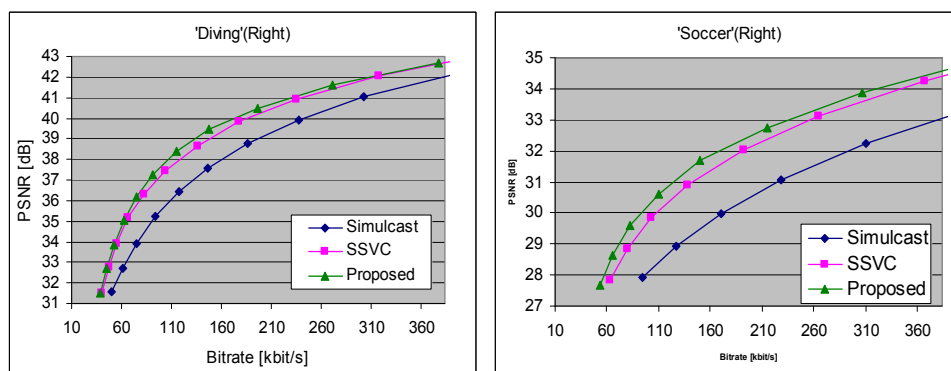Fig. 5. PSNR versus total bitrates of left and right-view sequences

Fig. 6. PSNR versus bitrates of right-view sequence

We show the coding results using RD (rate-distortion) curves of the proposed coding, conventional stereoscopic video coding, and Simulcast. Fig.5 represents the results of total coding efficiency for left- and right-view sequences. The proposed coding method has higher efficiency by up to 0:2dB and by 0.2 ~ 2.2dB than the conventional stereoscopic video coding and Simulcast, respectively, for both test sequences. RD curves in Fig. 6 represent the results of the coding efficiency for only right-view sequence. The proposed coding method has higher efficiency by up to 0.7dB and 0.2 ~ 2.3dB than the conventional stereoscopic video coding and Simulcast, respectively.

It is noted in the RD curves that the gain of the proposed coding method becomes bigger as the coding bit rate decreases. As the bit rate becomes higher, the irreversible loss caused by residual downsampling cannot be restored regardless of how many bits are allocated. In a practical situation, the bitrates of right-view images can be reduced much lower than that of left-view images because human perception is dominated by the high-quality component of a stereo pair. This can be achieved by separate rate control of each view. Hence it is necessary in the future work to study rate-control and test subjective quality for stereoscopic video coding in order to get better coding efficiency.

4. CONCLUSIONS

We compared different mobile TV standards, and introduced a 3-D mobile TV service based on DMB while maintaining the backward compatibility with the existing DMB system. We implemented prototype of transmitting server and 3-D DMB terminals, and have verified that the proposed concept of 3-D mobile TV services over DMB works well. We conclude that mobile TV will be the core platform for interactive multimedia communication in the future, and 3-D mobile TV will be very attractive future application with existing DMB service.

REFERENCES

[1]   Arrnold, J., Frater, M., and Pickering, M., [Digital Television: Technology and Standards], John Wiley & Sons, Inc., Hoboken, USA., (2007).
[2]   Furht, B. and Ahson, S., [Handbook of Mobile Broadcasting-DVB-H, DMB, ISDB-T, and MEDIAFLO], CRC Press, Boca Raton, USA, (2008).
[3]   Kumar, A., [Mobile TV], Focal Press, Oxford, UK, (2007).
[4]   ITU-R Rec. BS. 775-1: "Multichannel Stereophonic Sound System with and without accompanying Picture," (1994).
[5]   Pastoor, S., "3D Television: A survey of recent research results on subjective requirements," Signal Processing: Image Communication, 4(1), 21-32, (1991).
[6]   Cho, S., Hur, N., Kim, J., Yun, K. and Lee, S., "Carriage of 3D Audio-Video Services by T-DMB," Proc. International Conference on Multimedia & Expo (ICME), 2165–2168, (2006).