Quality Enhancement for Hybrid 3DTV with Mixed Resolution Using Conditional Replenishment Algorithm

Kyeong-Hoon Jung, Min-Suk Bang, Sung-Hoon Kim, Hyon-Gon Choo, and Dong-Wook Kang

This paper proposes a conditional replenishment algorithm (CRA) to improve the visual quality (where spatial resolutions of the left and right views are mismatched) of a hybrid stereoscopic 3DTV that is based on the ATSC-M/H standard. So as to generate an enhanced view, the CRA is to choose the better substitute among a disparity-compensated view with high quality and a simply interpolated view. The CRA generates a disparity map that includes modes and disparity vectors as additional information. It also employs a quad-tree structure with variable block size by considering the spatial correlation of disparity vectors. In addition, it takes advantage of the disparity map used in a previous frame to keep the amount of additional information as small as possible. The simulation results show that the proposed CRA can successfully improve the peak signal-to-noise ratio of a poor-quality view and consequently have a positive effect on the subjective quality of the resulting 3D view.

Keywords: ATSC-M/H, hybrid 3DTV, quadtree, conditional replenishment algorithm, disparity map, CRA.

I. Introduction

The 3DTV is seen as a major candidate among future broadcasting services, and many approaches have been suggested to help initiate such a service. Two familiar approaches for transmitting the 3D contents are the frame-compatible method [1] and the service-compatible method [2], both of which use a single channel. Also, there have been several hybrid 3DTV approaches that use two channels for the transmission of left and right views. Among these approaches, a new hybrid 3DTV system that can be transmitted via the ATSC-M/H has been proposed [3]. The block diagram of the hybrid 3DTV system via the ATSC-M/H, proposed in [3], is shown in Fig. 1. The left and right views are transmitted through DTV and mobile channels, respectively. If the ATSC-M/H receiver with dual codec is available, then both left and right views are decoded and combined to generate 3D content. The principal merit of this system is to guarantee compatibility with existing broadcasting services. That is, 2D HDTV, mobile 2DTV, and 3DTV can be provided simultaneously.

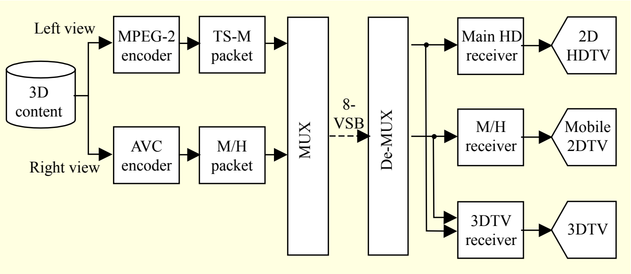


Fig. 1. Block diagram of ATSC-M/H-based hybrid 3DTV system.

Manuscript received Oct. 25, 2013; revised Feb. 24, 2014; accepted May 12, 2014.
This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2014.

Kyeong-Hoon Jung (khjung@kookmin.ac.kr), Min-Suk Bang (bmons85@kookmin.ac.kr), and Dong-Wook Kang (corresponding author, dwkang@kookmin.ac.kr) are with the Department of Electronics, Kookmin University, Seoul, Rep. of Korea.
Sung-Hoon Kim (steve-kim@etri.re.kr) and Hyon-Gon Choo (hyongonchoo@etri.re.kr) are with the Broadcasting & Telecommunication Media Research Laboratory, ETRI, Daejeon, Rep. of Korea.

http://dx.doi.org/10.4218/etrij.14.0113.1092

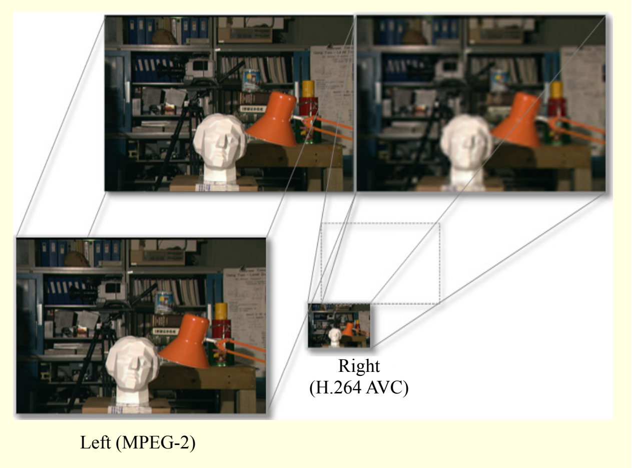


Fig. 2. 3D composition in ATSC-M/H-based hybrid 3DTV system.

There exists a mismatch in quality in the ATSC-M/H-based hybrid 3DTV system because the left and right views have different resolutions, as shown in Fig. 2. However, if one view is blurry while the other is clear, then the combined 3D quality of the scene (produced by both views) is likely to appear almost as clear as the high-resolution view. This fact is supported by binocular suppression theory, which deals with the response of the human visual system to stereoscopic vision with mixed resolution [4]–[5]. As a result, the proposed system could provide quite satisfactory results for many kinds of sequences even though there are considerable mismatches in resolutions [3].

However, 3D visual quality may be degraded when the resolution gap is too high or when there exist severe coding artifacts. Therefore, it is desirable to adopt a method of improving the quality of the view transmitted by a mobile channel. General video codecs, such as the Scalable Video Coding (SVC) [6] or the Multiview Video Coding (MVC) [7], can be considered as a solution to this problem. Naturally, SVC and MVC cannot properly support the service scenario given in Fig. 1 since the legacy DTV in ATSC-M/H should be encoded by MPEG-2. Instead, both codecs are based on H.264/AVC. Even though these codecs can be used, there are some innate drawbacks as they do not fully make use of all the available data; that is, the high-resolution left view is not considered in the SVC and the low-resolution right view is not considered in the MVC.

In this paper, a conditional replenishment algorithm (CRA) is proposed to enhance the quality of the right view. The concept of conditional replenishment is not new, and most conventional conditional replenishment methods that have been used for video coding are based on the temporal correlation between consecutive frames [8]–[10]. In contrast with the conventional methods, the CRA is an approach to expand the idea of conditional replenishment to the 3D composition problem, in which the left and right views have different resolutions. Unlike SVC or MVC, CRA can be considered not as a residual coding scheme but rather as a kind of switching process between two candidates; one such candidate being the disparity-compensated view (that is, the high-resolution left view), and the other being the simply interpolated view (that is, the low-resolution right view).

II. CRA

In the ATSC-M/H hybrid 3DTV system, the resolution of the left view is much larger than that of the right view. Herein, we refer to the high-resolution left view as simply the *left* view and the low-resolution right view as simply the *right* view. Thus, the right view needs to be enlarged to the same size as the left view to be composed for 3D display. The most direct way is to resize the right view by using bilinear interpolation or directional filtering. However, there is a strong correlation between the two views; that is, a pixel in one view usually has its corresponding position in the other view. Thus, it is obvious that the quality of the enlarged right view can be improved by using the left view.

The key idea of the CRA is to properly use both the left and right views to generate an enhanced right view. Firstly, the left view needs to be compensated by using disparity vectors, and then the right view needs to be enlarged to the same size as the left view. The left view can contribute to enhancing details or edges, and the right view can cover the occlusions or shadow areas that cannot be represented from the left view. Therefore, if the left and right views are properly selected, then the quality of the enlarged right view is expected to be improved.

The CRA requires two kinds of information; one is the mode information to indicate which substitute is selected, and the other is the disparity vectors when the disparity-compensated view is selected. Since this information needs to be transmitted to the decoder as additional information, it is desirable to keep it to a minimum.

The block diagram of the CRA is shown in Fig. 3. The output of the CRA is a disparity map that contains information about modes and accompanying disparity vectors. Three inputs are entered as candidates into the mode decision block.

First, the disparity-compensated view is generated. To generate it, the disparity vector between the original right (R_{O}) and the decoded left (L_{HD}) views needs to be estimated.

The second input is obtained from R_{MH}, which is the enlarged view of the decoded r_{MH}. The decoded r_{MH} needs to be enlarged so as to be of the same size as the left view. Any kind of interpolation method can be used for this purpose; we have used simple bilinear interpolation.

Meanwhile, there exists much temporal correlation in the disparity map; that is, the disparity map of a current frame

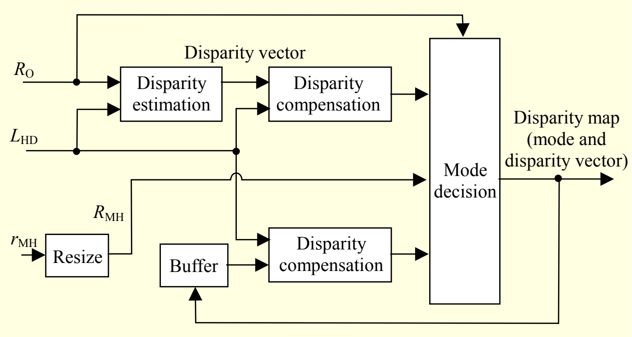


Fig. 3. Block diagram of CRA.

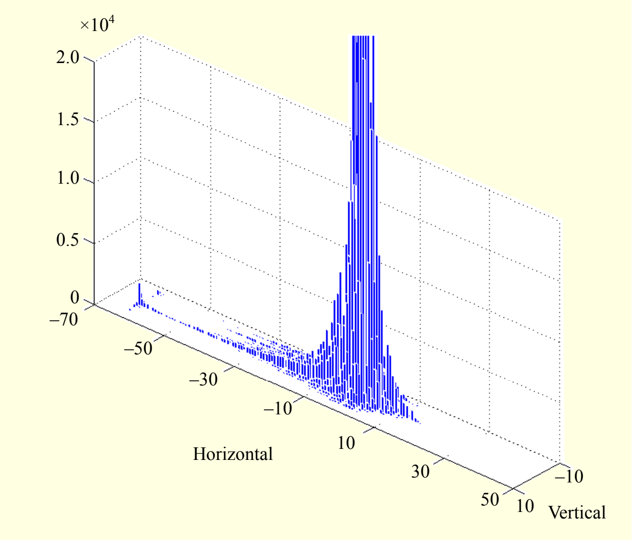


Fig. 4. Distribution of disparity vectors.

shows a similar mode distribution and disparity vector to that of the previous frame. The third input is introduced to reduce such temporal redundancy. This input is compensated by the decoded left view (L_{HD}) and the previous disparity map stored in a buffer. The final step is to determine which input is the most suitable for generating the disparity map. This mode decision step is the core part of the CRA and is described in detail in Section III.

The disparity vectors between the left and right views can be estimated by a block matching algorithm. This algorithm searches for the relative position having minimum mean absolute difference. Theoretically, a disparity vector — the horizontal distance between two stereoscopic views — has only a single component. However, in reality, some vertical misalignment between the different views exists; this is assumed to be mainly due to the mismatch of 3D cameras in many 3D sequences [11]. Figure 4 shows the typical pattern of distribution of disparity vectors. We can see that disparity vectors are distributed in both directions and that the vertical range is smaller than the horizontal range. Therefore, we must estimate the disparity vector vertically as well as horizontally; thus, the shape of the search range becomes more rectangular due to this fact.

III. Generation of Disparity Map

The size of a processing block for disparity estimation and compensation in the CRA needs to be carefully determined. If it is too large, then some unnatural visual artifact may be noticed since only a single disparity vector is given to each processing block. On the contrary, if it is too small, then the additional data rate for the transmission of the disparity map may exceed the reasonable bound. Of course, it is also desirable to take into account spatial characteristics, such as the size or shape of an object, as well as the pattern of depth variation in determining the size of a processing block. Therefore, the CRA includes variable-sized blocks structured by a quadtree to efficiently generate an enhanced version of the right view [12]. Here, the quadtree has hierarchical levels from the bottom upwards. The size of a processing block is at its minimum at the bottom level and is multiplied by a factor of two in each direction for each increase in level.

At each quadtree level, we have to determine the best mode among three candidates for each processing block. The first mode is for generating the disparity-compensated view by using the encoded left view and the current disparity vector. For convenience, we call this the “LD mode”. The second mode, which uses the enlarged version of the right view, is called the “RI mode”. The third mode, which is compensated by using the previous disparity map, is called the “PD mode”. By adopting the PD mode, we can reduce the temporal redundancy in a disparity map; however, this mode cannot be used for instantaneous decoding since it requires information processed at the previous frame.

To determine the optimum mode, a cost function is introduced (see (1)), where D is the distortion, B is the amount of bits needed to represent each mode, and λ is a Lagrange multiplier.

$$J = D + \lambda B\,. \tag{1}$$

For a given level l, we let N be the length of a side of a processing block. We then denote the starting point of a block by (w,h). Then the cost function of the LD mode for a given level l, $J_{\mathrm{LD}}(l;w,h)$, can be obtained through

$$J_{\mathrm{LD}}(l;w,h)=\frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\left\{\begin{matrix} R_{\mathrm{O}}(w+i,h+j) \\ -L_{\mathrm{HD}}(w+i+d_h(w,h), \\ h+j+d_v(w,h)) \end{matrix}\right\}^2+\lambda B_{\mathrm{LD}}\,, \tag{2}$$

where $d_h(w,h)$ is a horizontal disparity vector of a processing block and $d_v(w,h)$ is the corresponding vertical disparity vector.

The amount of bits for the LD mode is represented by B_{LD}.

The cost functions of the RI mode and the PD mode, at a given level l, respectively, are as follows:

$$J_{\mathrm{RI}}(l;w,h)=\frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\begin{Bmatrix} R_{\mathrm{O}}(w+i,h+j) \\ -R_{\mathrm{MH}}(w+i,h+j)\end{Bmatrix}^2 + \lambda B_{\mathrm{RI}}, \quad (3)$$

$$J_{\mathrm{PD}}(l;w,h)=\frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\begin{Bmatrix} R_{\mathrm{O}}(w+i,h+j) \\ -R_{\mathrm{PDC}}(w+i,h+j)\end{Bmatrix}^2 + \lambda B_{\mathrm{PD}}, \quad (4)$$

where B_{RI} and B_{PD} are the amount of bits for the RI mode and the PD mode, respectively. Since the disparity vectors are included in the LD mode, B_{LD} is always larger than B_{RI} or B_{PD}.

In (4), R_{PDC} denotes the disparity-compensated view. This view makes use of the previous disparity map and is obtained by

$$R_{\mathrm{PDC}}(w,h)=\begin{cases} L_{\mathrm{HD}}(w+d_h^{-1}(w,h), \\ \quad h+d_v^{-1}(w,h)) & \exists\, d^{-1}(w,h), \\ R_{\mathrm{MH}}(w,h), & \text{otherwise,}\end{cases} \quad (5)$$

where $(d_h^{-1}(w,h),\ d_v^{-1}(w,h))$ denotes the disparity vector used in the previous frame.

As stated in the previous section, the main purpose of the PD mode is to lower the temporal redundancy in the disparity map. However, this mode requires the use of a previous frame's disparity map; therefore, making it inappropriate for instantaneous decoding. For this reason, we classify the mode decision into two types: one being INTRA type (where the PD mode is not allowed), which is used for random access to every frame; and the other being INTER type (where the LD, RI, and PD modes are all possible candidate modes). Then the cost of a given processing block is replaced by the minimum value as follows:

$$J_{\mathrm{O}}=\begin{cases}\min[J_{\mathrm{RI}},J_{\mathrm{LD}}], & \text{if INTRA type,}\\ \min[J_{\mathrm{RI}},J_{\mathrm{LD}},J_{\mathrm{PD}}], & \text{if INTER type.}\end{cases} \quad (6)$$

As well as deciding upon which mode to use, we must also determine whether to continue to the next quadtree level. This process is also controlled by comparing the rate distortion (RD) costs of a given block and its four sub-blocks. For a given block at level l, its quadtree structure is shown in Fig. 5, and J_{child}, the sum of the four RD costs corresponding to the sub-blocks, is given as

$$J_{\mathrm{child}}(l;w,h)=\begin{Bmatrix} J_{\mathrm{O}}(l-1;w,h) \\ +J_{\mathrm{O}}(l-1;w+2^{l-1},h) \\ +J_{\mathrm{O}}(l-1;w,h+2^{l-1}) \\ +J_{\mathrm{O}}(l-1;w+2^{l-1},h+2^{l-1})\end{Bmatrix}. \quad (7)$$

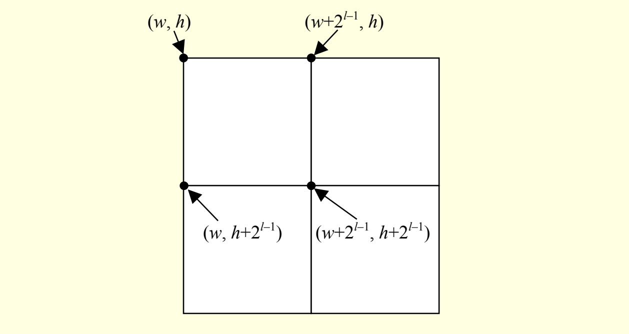


Fig. 5. Quadtree structure at level l.

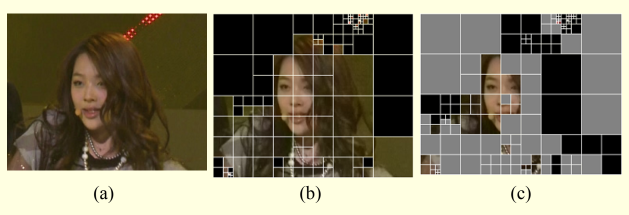

(a) (b) (c)

Fig. 6. Mode distributions in a disparity map: (a) original right view, (b) INTRA-type mode distribution, and (c) INTER-type mode distribution.

Then the criterion for merging at a given level l ($l \geq 2$) is represented as

$$\begin{cases}\text{GOTO the next level,} & \text{if } J_{\mathrm{O}} < J_{\mathrm{child}},\\ \text{STOP at this level,} & \text{otherwise.}\end{cases} \quad (8)$$

The Lagrange multiplier λ plays an important role in the regulation of the bitrate of additional information as well as in the overall performance. By changing λ at each frame, the bitrate of the additional information for the CRA can be controlled. If we use more data for the CRA, then the size of the processing blocks mostly becomes smaller; hence, more gain is expected since details can be well compensated.

Figures 6(b) and 6(c) are typical examples of the different mode distributions of a disparity map. Figure 6(b) illustrates the pattern for an INTRA type, where colored blocks correspond to the LD mode and empty blocks the RI mode. It should be noted that the blocks are successfully partitioned according to the spatial characteristics. The INTER-type mode distribution is given in Fig. 6(c), where gray blocks correspond to the PD mode. By observing that many blocks are determined as PD modes, we can see that the two consecutive disparity maps have similar distribution patterns.

IV. Simulation Result

We used several 3D video sequences with various characteristics for simulation. The color format of each original sequence is YUV 4:2:0, with a vertical resolution of 1,080p. Each sequence consists of 240 frames at a frame rate of 29.97 fps.

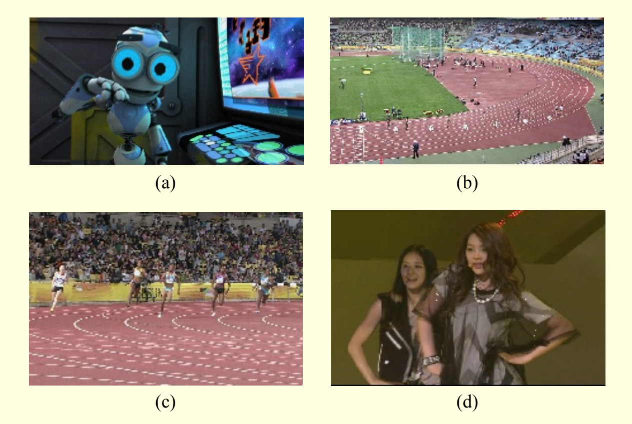


Fig. 7. Test sequences: (a) 1, (b) 2, (c) 3, and (d) 4.

Figure 7 shows four test sequences. Firstly, sequence_1 is an animation that has sharp and clear edges and contains many fast movements. Next, sequence_2 and sequence_3 are obtained from a sports event. Sequence_2 is captured a long distance away from the action; hence, the sizes of objects and subsequently their motion vectors are small. On the contrary, the players in sequence_3, since it is a close-up image sequence, are relatively large and appear to be running fast. The backgrounds in both sequences look complex, but the spatial frequencies are not too high. The crowd in sequence_3 is moving in one direction as the camera is panning. Finally, sequence_4 is obtained from a music show, where two singers are dancing on a stage. The sizes of objects are large and their movements are fast. Also, the illuminations in the background are periodically flickering.

In the proposed hybrid 3DTV system, the left and right views need to be independently encoded to be transmitted via DTV and mobile channels. Table 1 represents the encoding parameters for both views.

The resulting peak signal-to-noise ratios (PSNRs) of the images encoded by the above encoding parameters are summarized in Table 2. As expected, the PSNR range for the left view is from 34.17 dB to 43.63 dB and that of the right view is from 23.61 dB to 32.55 dB. The average difference between the two views is about 10.34 dB.

Table 3 gives the parameters required by the CRA for the generation of additional information. As previously stated, the shape of the search range is a wide rectangle (the horizontal length is from –32 to 32 pixels and the vertical length is from –7 to 7 pixels). To estimate the disparity vectors, we use half pixels (calculated by a bilinear interpolator) as well as integer pixels, and the maximum size of the processing blocks structured by the quadtree is 128 × 128 pixels (level 7). The minimum size reaches a single pixel (level 0) to describe fine details.

Table 1. Encoding parameters for left and right views.

Encoding parameter	Left view	Right view
Format	YUV 4:2:0	YUV 4:2:0
Resolution	1,920 × 1,080	416 × 240
Codec	MPEG-2	H.264/AVC
Profile	Main profile	Main profile
Bitrate	12 Mbps	480 kbps

Table 2. PSNRs for the left and right views.

Sequences	PSNR for left view (dB)	PSNR for right view (dB)
Sequence_1	39.84	30.21
Sequence_2	34.17	23.61
Sequence_3	36.52	26.43
Sequence_4	43.63	32.55

Table 3. Encoding parameters for CRA.

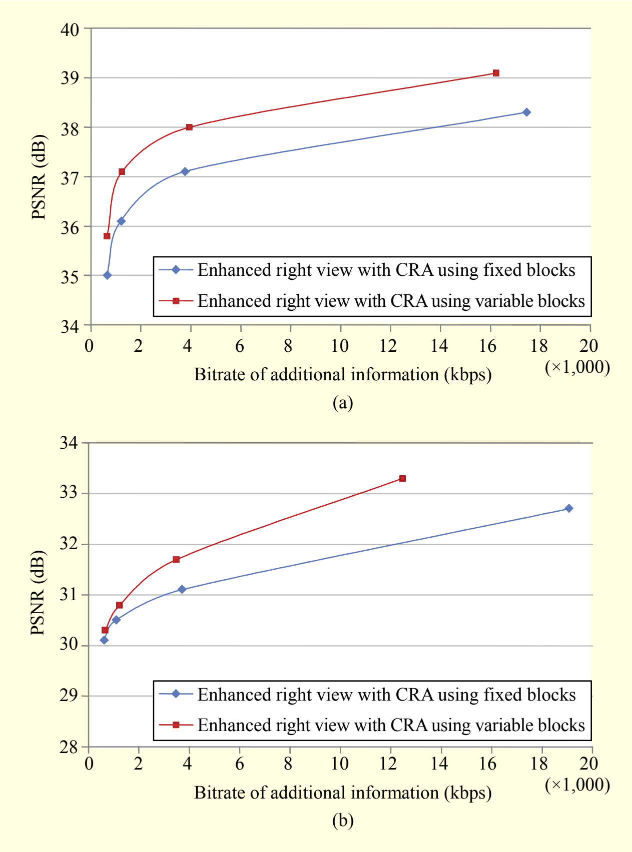
Encoding parameter	Value
Search range of horizontal disparity	–32 to 32
Search range of vertical disparity	–7 to 7
Precision	Half pixel
Maximum block size	128 × 128
Minimum block size	1 × 1
Intra period	30

In addition, the intra period is 30; that is, the image generated by the mode decision with the INTRA type is periodically repeated every 30 frames. To reduce the amount of additional information, the exponential Golomb codes are used after differential pulse-code modulation of neighboring disparity vectors [13].

Our first experiment is to show the validity of the quadtree structure in the CRA. Figure 8 shows the RD curves of the enlarged right views for sequence_1 and sequence_3. We compare the case of variable blocks with that of fixed blocks. The size (in pixels) of the fixed blocks varies from 2 × 2 to 16 × 16. As you can see, there is a considerable gain when the variable blocks are adopted in both sequences. The average of the Bjøntegaard delta (BD)-PSNR [14] increases by 0.68 dB and that of the BD-rate is reduced to 66.9%.

The quality gain when using variable-sized blocks can be clearly seen in Fig. 9. The bitrate of the additional information in each case is the same. We can find blocky and blurred artifacts in the encircled region (Fig. 9(a)) when using blocks of

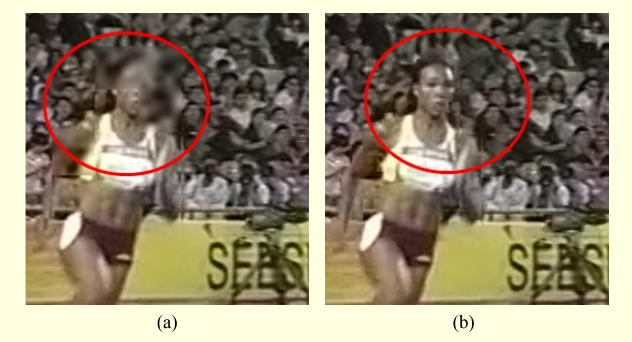
http://dx.doi.org/10.4218/etrij.14.0113.1092



(a)

(b)

Fig. 8. RD curves of enhanced right views using fixed- and variable-sized blocks: (a) sequence_1 and (b) sequence_3.



(a)

(b)

Fig. 9. Enhanced right view using CRA: (a) using fixed-sized blocks and (b) using variable-sized blocks.

a fixed size. In contrast, in Fig. 9(b), it can be seen that when using variable-sized blocks, the details are successfully compensated.

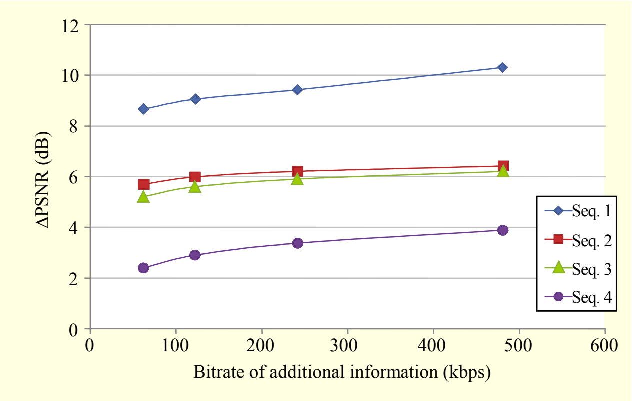
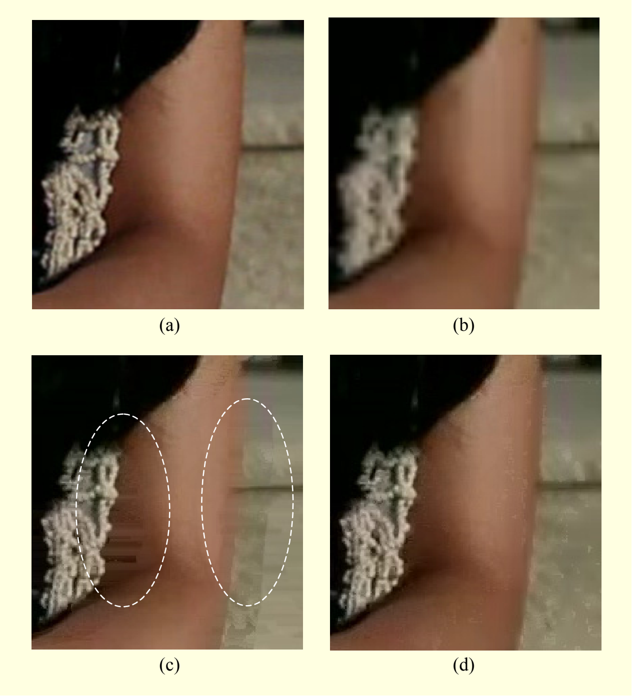


Fig. 10. Performance of CRA.



(a)

(b)

(c)

(d)

Fig. 11. (a) Original right view, (b) enlarged right view via bilinear interpolation (34.87 dB), (c) disparity-compensated view (29.61 dB), and (d) enhanced right view by CRA (40.60 dB).

The performance of the CRA is given in Fig. 10, where the vertical axis (ΔPSNR) denotes the difference in PSNR of the enhanced right view by CRA and the simply enlarged one. In the case of sequence_1, the PSNRs for the left and right views are 39.84 dB and 30.21 dB, respectively (Table 2). The average ΔPSNR for sequence_1 is 9.37 dB (Fig. 10). This means that the PSNR of the enhanced right view increases to 39.58 dB on average and is consequentially close to the PSNR of the left view.

The PSNR gain depends on the characteristics of a sequence. The average ΔPSNRs of the four RD curves are 9.37 dB, 6.08 dB, 5.73 dB, and 3.14 dB, respectively. Sequence_1, which is from an animation, benefits the most; the result demonstrating that there is a strong correlation between the left and right views in this case.

The performances of sequence_2 and sequence_3, which are from a real sports event, are similar. The PSNR gap in these

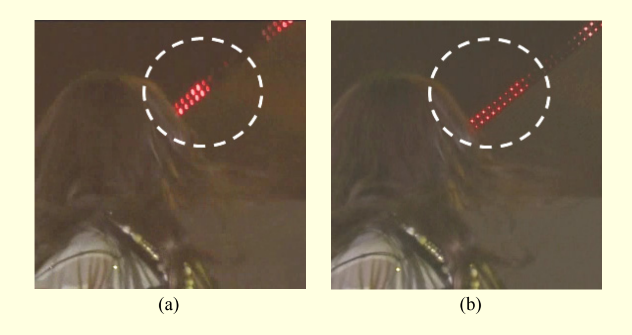


Fig. 12. Mismatch in sequence_4: (a) left view and (b) right view.

sequences is reduced from 10.33 dB to 4.42 dB on average.

Figure 11 shows the necessity of the CRA. Figure 11(a) is the original right view with a resolution of 1,080p. Figure 11(b) is the simply enlarged right view by using a bilinear interpolator from a resolution of 240p; here, some blurring and loss of detail can be easily noticed.

Figure 11(c) is the disparity-compensated view generated by the left view with resolution of 1,080p, where we can see that edges and details are well preserved in most regions; however, there are some artifacts in the regions encircled by a dashed line. Figure 11(d) is the enhanced right view generated by the CRA. It is clear that the visual quality, from both a subjective and objective point of view, can be greatly improved.

In contrast, the PSNR gain for sequence_4 is relatively low. We suppose that this is mainly due to synchronization mismatch or some calibration problems in the left and right views. Figure 12 represents an example of a stereoscopic view of sequence_4. A light in the background can be seen flickering, yet its size and brightness are not consistent in both views. This observation tells us that the two views were not captured at the same time. This temporal discrepancy is one of the critical reasons for the low PSNR gain obtained through using the CRA. In other words, the original stereoscopic pair should be well aligned not only spatially but also temporally in a high-performance CRA.

The next experiment is to compare the performance of the CRA with that of the MVC and SVC codecs. As said before, both codecs are not adequate for the service scenario given in Fig. 1. However, we think that a relative evaluation is still possible by comparing their respective RD curves; this is given that there is no existing conventional target for comparing with our method. For this, we use the reference software JMVM 8.0 [15] for the MVC and the reference software JSVM 9.19.7 for the SVC. Strictly speaking, this environment is slightly unfair and disadvantageous to the CRA because the left view is encoded by MPEG-2 in the CRA but encoded by H.264/AVC in the MVC and SVC. If H.264/AVC is used in the CRA, the quality of the left view is likely to be improved and more PSNR gain is expected.

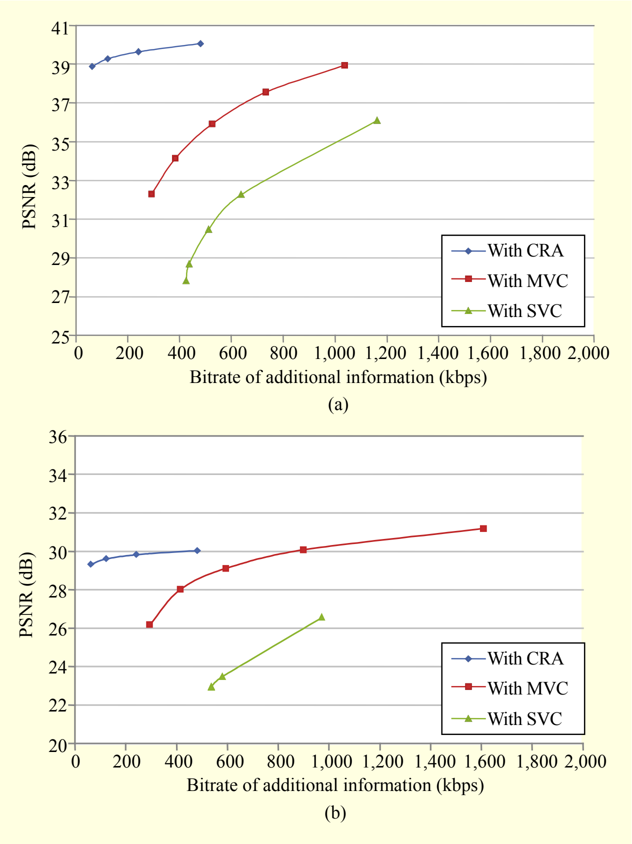
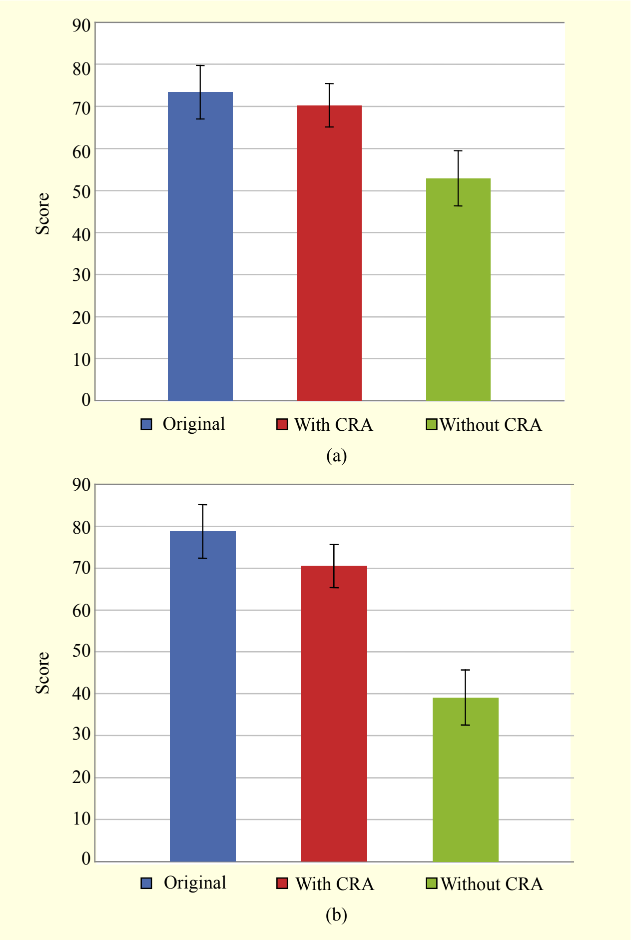


Fig. 13. RD curves of the enhanced right view via SVC, MVC, and CRA: (a) sequence_1 and (b) sequence_3.

The resulting PSNRs of the enhanced right view are given in Fig. 13, where the horizontal axis denotes the bitrate for the enhancement layer in the cases of MVC and SVC. It can be seen that the proposed CRA outperforms both MVC and SVC. In the case of sequence_1, there are 5.95 dB and 10.74 dB gains in terms of BD-PSNR. In the case of sequence_3, we get 2.43 dB and 12.91 dB gains. What is more important than the BD-PSNR gain is the fact that the CRA can give remarkable PSNR gains, especially at a very low bitrate of additional information.

The ultimate goal of the CRA is to improve the perceived 3D visual quality. The double stimulus continuous quality scale (DSCQS) test [16] is used for the assessment of the subjective quality; twenty-two non-experts, aged between 21 and 30, participated in the test. The 3D display was a 47-inch light emitting diode 3DTV of the shutter glasses type. The viewing distance was 3.3 times longer than the height of the 3D display. The resulting scores for sequence_1 and sequence_3 are given in Fig. 14; simpler tests were performed on the remaining sequences, and the results showed that all shared a similar tendency. Here, the colored bars indicate the average score, and the black lines that are marked on the upper part of the colored bars represent ninety-five percent confidence intervals. As



(a)

(b)

Fig. 14. DSCQS scores: (a) sequence_1 and (b) sequence_3.

expected, the perceived visual qualities are greatly improved by using the CRA. In the case of sequence_1, the score is comparable to that of original stereoscopic views. The rise in DSCQS score is over 30 points in the case of sequence_3.

V. Conclusion

In this paper, we proposed an effective algorithm to improve the visual quality of an ATSC-M/H-based hybrid 3DTV system, in which, to display 3D stereoscopic images, the small right view transmitted via a mobile (M/H) channel should be enlarged. The basic idea of the proposed CRA is to determine the better substitute to generate the enlarged right view; that is, to choose from either the disparity-compensated view or the spatially interpolated view. In addition, a disparity map, consisting of mode information and disparity vectors, is to be transmitted to the decoder. Thus, we adopt variable-sized blocks that are structured by a quadtree and use the disparity map used in the previous frame to reduce the extra bitrate.

Computer simulations show the CRA, from both a subjective and objective point of view, successfully improved the visual quality. By applying the CRA, the PSNR gain increased by up to 10 dB with only 400 kbps of additional information. In addition, the DSCQS scores greatly increased by up to about 30 points.

The primary advantage of the CRA is that it can guarantee backward compatibility with current broadcasting systems. Therefore, if the CRA is applied to an ATSC-M/H-based hybrid 3DTV system, then it is expected to play an important role in activating 3D broadcasting services.

References

[1] K.H. Kim et al., “Efficient Stereoscopic Contents File Format on the Basis of ISO Base Media File Format,” *Proc. SPIE Multimedia Mobile Devices*, vol. 7256, San Jose, CA, USA, Jan. 28, 2009, pp. 72560N-1–72560N-9.

[2] K.J. Yun et al., “Design and Implementation of a Three-Dimensional Broadcasting System Based on DTV,” *J. Broadcast Eng.*, vol. 11, no. 4, Dec. 2006, pp. 483–494.

[3] B.Y. Kim et al., “A Study on Feasibility of Dual-Channel 3DTV Service via ATSC-M/H,” *ETRI J.*, vol. 34, no. 1, Feb. 2012, pp. 17–23.

[4] L. Stelmach et al., “Stereo Image Quality: Effects of Mixed Spatio-Temporal Resolution,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, Mar. 2000, pp. 188–193.

[5] A.K. Jain, A.E. Robinson, and T.Q. Nguyen, “Comparing Perceived Quality and Fatigue for Two Methods of Mixed Resolution Stereoscopic Coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 3, Mar. 2014, pp. 418–429.

[6] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, Sept. 2007, pp. 1103–1120.

[7] 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, Apr. 2011, pp. 626–642.

[8] F.W. Mounts, “A Video Encoding System with Conditional Picture-Element Replenishment,” *Bell Syst. Techn. J.*, vol. 48, no. 7, Sept. 1969, pp. 2545–2554.

[9] D.N. Hein and N. Ahmed, “Video Compression Using Conditional Replenishment and Motion Prediction,” *IEEE Trans. Electromagn. Compat.*, vol. 26, no. 3, Aug. 1984, pp. 134–142.

[10] J.E. Fowler, M. Tagliasacchi, and B. Pesquet-Popescu, “Video Coding with Wavelet-Domain Conditional Replenishment and Unequal Error Protection,” *Proc. Int. Conf. Image Process.*, Atlanta, GA, USA, vol. 2, Oct. 8–11, 2006, pp. 1869–1872.

[11] D.H. Lee et al., “Estimation of Disparity with Vertical Offset to Improve the Quality of 3DTV System with Asymmetrical Resolution,” presented at the Korean Soc. Broadcast Eng.

Summer Conf., Jeju, Rep. of Korea, 2012.
[12] M.-S. Bang et al., “A Variable Block Encoding Method Using the Space-Time Redundancy of Stereoscopic Vectors,” presented at the Korean Soc. Broadcast Eng. Autumn Conf., Seoul, Rep. of Korea, 2011.
[13] M.-S. Bang et al., “A Study on Efficient Encoding Method for Binocular Disparity Map,” presented at the Korean Soc. Broadcast Eng. Summer Conf., Jeju, Rep. of Korea, 2012.
[14] G. Bjontegaard, “Improvements of the BD-PSNR Model,” ITU-T SG16/Q6, Berlin, Germany, Tech. Rep. VCEG-AI11, July 2008.
[15] A. Vetro et al., “Joint Multiview Video Model (JMVM) 8.0,” Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), Geneva, Switzerland, Tech. Rep. JVT-AA207, Apr. 2008.
[16] ITU-R Rec. BT.500-11, *Methodology for the Subjective Assessment of the Quality of Television Pictures,* June 2002.

Kyeong-Hoon Jung received his BS, MS, and PhD degrees in electronic engineering from Seoul National University, Seoul, Rep. of Korea, in 1987, 1989, and 1996, respectively. After working at the Korea Broadcasting Institute, Seoul, Rep. of Korea, from 1991 to 1997, he joined Handong University, Pohang, Rep. of Korea, as a faculty member. Since 2005, he has been with the School of Electronic Engineering, Kookmin University, Seoul, Rep. of Korea, where he is now a professor. His research interests are broadcasting systems, video coding, and multimedia signal processing

Min-Suk Bang received his BS in electronic engineering from the School of Electronic Engineering, Kookmin University, Seoul, Rep. of Korea, in 2010. Since 2010, he has worked toward his MS and PhD joint degree in electronic engineering from the same university. His research interests are 3DTV, broadcasting systems, and signal processing.

Sung-Hoon Kim received his BS, MS, and PhD degrees in electronic engineering from the Department of Electronic Engineering, Kookmin University, Seoul, Rep. of Korea, in 1994, 1996, and 2008, respectively. Since 2000, he has been a principal researcher at the Electronics and Telecommunications Research Institute, Daejeon, Rep. of Korea. His research interests are digital signal processing in the field of Terrestrial & Mobile DTV, 3DTV, and UHDTV broadcasting systems.

Hyon-Gon Choo received his BS and MS degrees in electronic engineering and his PhD degree in electronic communication engineering from Hanyang University, Seoul, Rep. of Korea in 1998, 2000, and 2005, respectively. He has been a senior researcher at the Electronics and Telecommunications Research Institute, Daejeon, Rep. of Korea. His research interests are 3DTV, content-based audio/video analysis, biometrics, multimedia security, and protection

Dong-Wook Kang received his BS, MS, and PhD degrees in electronic engineering from Seoul National University, Seoul, Rep. of Korea, in 1986, 1988, and 1995, respectively. Since 2005, he has been with the School of Electronic Engineering, Kookmin University, Seoul, Rep. of Korea, where he is now a professor. He is a trustee of the Korean Society of Broadcasting Engineers. His research interests include video coding, multimedia signal processing, 3DTV, and digital culture technology.

http://dx.doi.org/10.4218/etrij.14.0113.1092