

Effects of Selection of a Reference View on Quality Improvement of Hybrid 3DTV

Min-Suk Bang*, Ji-Won Kim*, Sung-Hoon Kim**, Ki-Doo Kim*, and Kyeong-Hoon Jung*

*Department of Electrical Engineering, Kookmin University, Seoul, Korea

**ETRI, Daejeon, Korea

bmons85@kookmin.ac.kr, kjwon777@kookmin.ac.kr, steve-kim@etri.re.kr, kdk@kookmin.ac.kr

Corresponding author: khjung@kookmin.ac.kr

Abstract— Conditional replenishment algorithm (CRA) was proposed to improve the quality of a hybrid 3DTV system with mixed resolution. In the hybrid 3DTV system, the role of a reference view is important since not only the overall quality of the hybrid 3DTV depends on the quality of the reference view, but also the reference view directly affects the quality improvement of the hybrid 3DTV using CRA. Moreover, since stereo mismatch between left and right cameras can be occurred, carefully choosing the reference view between the left and right views is very important. In this paper, we observe effects of the selection of a reference view on the quality of a hybrid 3DTV which is improved by CRA.

Keywords— Hybrid 3DTV, CRA(conditional replenishment algorithm), Reference view, Mixed resolution, Focus mismatch.

I. INTRODUCTION

For commercialization of a 3DTV service, a hybrid 3DTV system based on DTV main and mobile channels has been proposed [1]. In this system, a left view of stereoscopic video is transmitted by a main channel as a reference view, and a right view is transmitted by a mobile channel as an additional view for the 3D effect. This system is very efficient since an additional bandwidth is not required for the 3DTV service. In particular, a major advantage of this system is that both the 2D HDTV service and the mobile broadcasting service are unaffected. The feasibility of this system was proven through the various field tests [2], and recently, this system has been adopted as a standard of the terrestrial 3DTV broadcasting service in ATSC-M/H under the name of service compatible 3DTV using main and mobile hybrid delivery (SC-MMH) [3].

There is a mismatch issue in the quality of the hybrid 3DTV system, since the left and right views have different resolutions. However, although the additional view is blurry, the combined 3D quality is likely to appear almost as clear as the reference view with high-resolution if the other is clear. 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]. For this reason, the hybrid 3DTV system could provide a quite satisfactory quality for the various kinds of 3D sequences [1].

However, if the quality gap is too high or some coding artifacts in the additional view video are severe, viewers may discern the quality degradation of the poor view, which results in deterioration of the overall quality. To solve this problem, conditional replenishment algorithm (CRA) [5] was proposed, which improves the quality of the additional view by using the reference view.

Meanwhile, a stereo mismatch is a classical problem in computer vision that has been extensively studied [6]. This mismatch may result from a number of causes, such as the cameras focused at different depths, variations in shutter speed, and camera shake causing motion blur. And it usually appears when there is a focus mismatch between left and right cameras. This mismatch between two cameras mostly makes one view have sharp image and another view have blurring image.

The reference view plays an important role since overall quality of 3DTV depends on the quality of the reference view, and particularly in CRA, the reference view is used for the quality improvement of the additional view. The well-matched stereo video can be captured under very carefully controlled conditions. However, capturing well-calibrated high quality video is a challenging problem and not always possible. Thus, the reference view needs to be carefully selected between left and right views.

In this paper, we examine effects of the selection of the reference view on the quality of a hybrid 3DTV which is improved by CRA. Although some methods to automatically select the reference view are needed for using in real broadcasting systems, this paper preferentially deals with effects by the selection of the proper reference view.

II. ROLES OF A REFERENCE VIEW IN CRA

In order to display the hybrid 3DTV, the small-sized additional view needs to be enlarged to the same size as the reference view. The intra view approach, such as a simple bilinear interpolation, can be considered as simple solution, but it does not use any information from the reference view. It is obvious that there is large correlation between two views of 3D contents and a pixel in one view usually has its corresponding position in another view. If the information of

disparity between the reference and additional view can be used, we expect that the quality of the enlarged additional view is greatly enhanced by partially using the reference view. Thus, carefully selecting the reference view between the left and right views is quite important since it affects directly on the quality improvement of the additional view.

The key idea of CRA is to properly use both the reference and additional views for generating an enhanced additional view. The block diagram of CRA is shown in Figure 1 [5]. Firstly, the disparity vectors between original images of the additional view and decoded images of the encoded reference view are estimated. Next, the view compensated by the disparity vectors is generated and the small-sized additional view is enlarged to have the same size as the reference view.

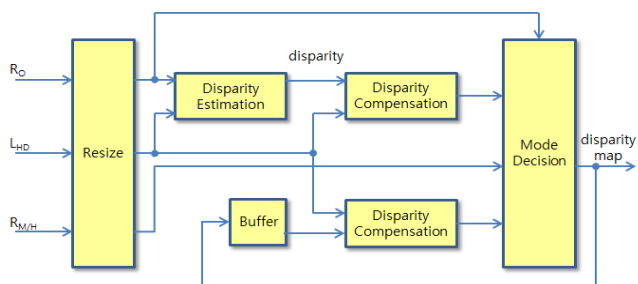


Figure 1. The block diagram of conditional replenishment algorithm.

Meanwhile, CRA generates video enhancement information (VEI) which includes mode information about switching results and disparity vectors for using only the reference view. Thus, the encoding of VEI should be considered; the previous disparity map is used for the generation of the third candidate to reduce the temporal redundancy, and blocks structured by quad-tree is used since the variable-sized block is preferred by observing that there is large spatial correlation in mode information and disparity vectors.

The final step is to determine the mode which conveys information about the best candidate and to produce the disparity map.

III. PERFORMANCE EVALUATION

In general, a stereo inconsistency, as shown in Figure 2 and Figure 3, may happen when taking a left/right stereo image with the camera due to the focus mismatching, differences in shutter speed, and movement of a camera. In these figures, the upper two images represent an enlarged view of the portion of an image inside of red dotted lines. The upper left and right represent the left image and right image, respectively.

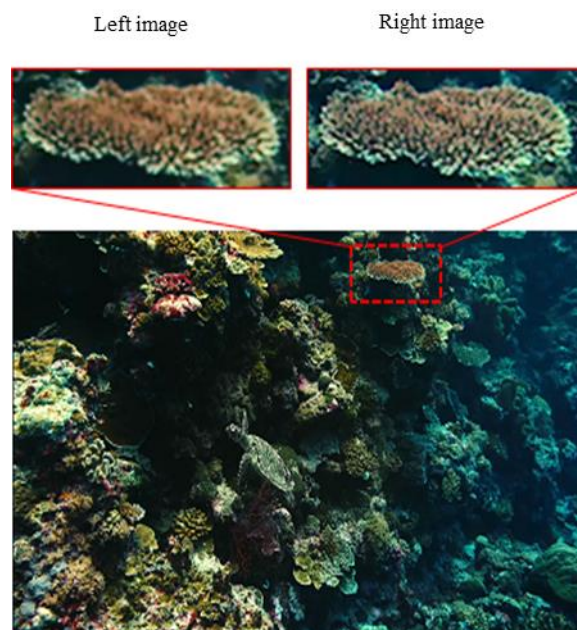


Figure 2. Resolution difference in the left/right camera images of a turtle's movement.

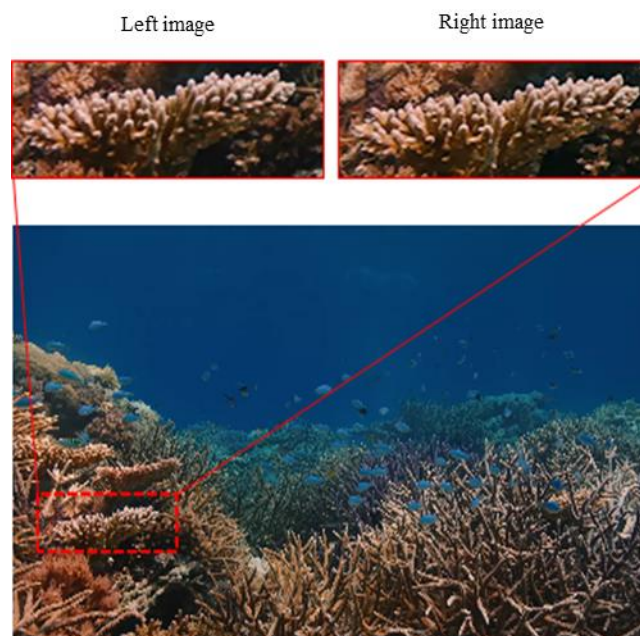


Figure 3. Resolution difference in the left/right camera images of underwater coral reefs.

The problem of stereo inconsistency often occurring in the stereo camera is mostly caused by the focus mismatching between left and right camera. Thus, most of the left/right image may have some degree of resolution difference and this difference can affect the overall quality of the hybrid 3DTV. By binocular suppression theory, the image quality of the hybrid 3DTV having an asymmetric resolution may follow the resolution of the reference image.

Therefore, in the image quality of the hybrid 3DTV, because the role of the reference image is critical, the reference view needs to be carefully selected between left and right views rather than using only one side view as a reference to enhance the subjective quality of the hybrid 3DTV.

A. Experiment environment

As shown in Figure 4, two sequences, which have focus mismatching, are used for performance evaluation. Both sequences have sharp and clear edges. While sequence 1 has a lot of movements, sequence 2 has relatively slow movements. The color format of each sequence is YUV 4:2:0, and its vertical resolution is 1080p. In order to satisfy the hybrid 3DTV service environment, a reference view video for the HD service is encoded by H.264/AVC with about 10Mbps, and an additional-view video for the mobile service is encoded by H.264/AVC with about 500kbps, where the vertical resolution of the additional view video is 240p. For encoding each video, JM, reference software of H.264/AVC, (ver. 18.5) is used.

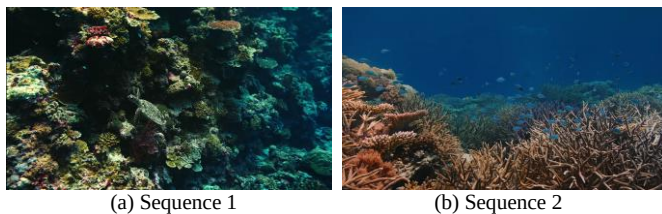


Figure 4. Sequences used for the performance evaluation.

B. Experimental results

Figure 5 shows rate-distortion (R-D) curves of each sequence, and four target bit-rates (120, 240, 360, and 480 kbps) are used for the encoding of VEI. To compare the performance, Bjøntegaard delta (BD)-PSNR [7] can be used, which corresponds to the average PSNR difference for the same bit-rate. For each sequence, the sharpness of the right view is higher than that of the left view. Figure 6 shows that the sharpness of the right view image is higher than that of the left view image. BD-PSNR results claim that when the reference view is the right view, the performances of sequence 1 and 2 are 1.44 dB and 1.63 dB better, respectively.

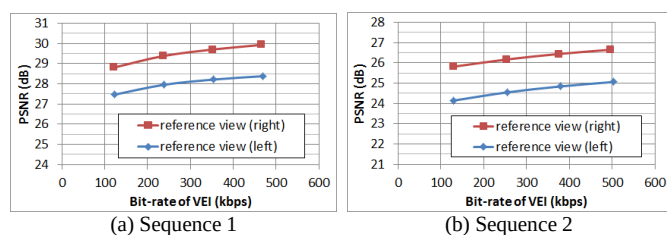


Figure 5. R-D curves of sequence 1 and 2.

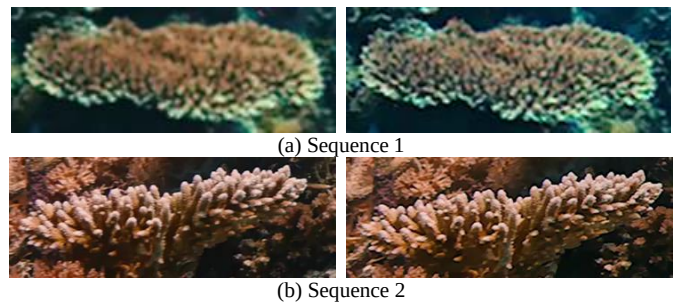


Figure 6. Sharpness differences between left and right views. Images located in left-side are left view and those in right-side are right view.

IV. CONCLUSIONS

In this paper, we observed that the quality of a reference view affects the quality of the hybrid 3DTV improved by CRA. Several images were used for the performance evaluation. Experimental results showed that just choosing the proper reference view improves the overall quality of the hybrid 3DTV, and the average PSNR gain was about 1.5 dB.

Later, an interesting outcome will be expected if a method of automatic selection of the reference view is added.

ACKNOWLEDGMENT

This work was supported by the ICT R&D program of MSIP/IITP (2015). [15-000-02-002, Development of Service and Transmission Technology for Convergent Realistic Broadcast].

REFERENCES

- [1] B. Y. Kim et al., "A study on feasibility of dual-channel 3DTV service via ATSC-M/H," *ETRI Journal*, vol. 34, no. 1, pp. 17-23, Feb. 2012.
- [2] S. H. Kim, J. Y. Lee, J. S. Choi, J. W. Kim, D. W. Kang and K. H. Jung, "ATSC 8-VSB and M/H hybrid 3DTV system development for terrestrial broadcasting services," in *Proc. Three-Dimensional Imaging, Visualization, and Display 2012*, Maryland, USA, May 2012.
- [3] ATSC A/104, "3D-TV terrestrial broadcasting, part5-service compatible 3D-TV using main and mobile hybrid delivery," *Advanced Television Systems Committee Standards*, Washington, D.C., Aug. 2014.
- [4] L. Stelmach, W. J. Tam, D. Meegan, and A. Vincent, "Stereo image quality: effects of mixed spatio-temporal resolution," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 188-193, Mar. 2000.
- [5] K. H. Jung, M. S. Bang, S. H. Kim, H. G. Choo and D. W. Kang, "Quality enhancement for hybrid 3DTV with mixed resolution using conditional replenishment algorithm," *ETRI Journal*, Oct. 2014.
- [6] C. Doutre and P. Nasiopoulos, "Sharpness matching in stereo images," *Journal of Virtual Reality and Broadcasting*, vol. 9, no. 4, 2012.
- [7] G. Bjøntegaard, "Improvements of the BD-PSNR model," *ITU-T SG16/Q6*, Berlin, Germany, Tech. Rep. VCEG-A11, July 2008.



Min-Suk Bang is a Ph.D's course in the Department of electronical engineering, Kookmin University, Seoul, Korea. He received B.S degree in Department of Electrical Engineering, Kookmin University in 2010. He is currently interested in 3DTV, broadcasting systems, and signal processing.



Ji-Won Kim is a Master's course in the Department of electronical Engineering, Kookmin University, Seoul, Korea. He received B.S degree in Department of Electrical Engineering, Kookmin University in 2015. His research interests include multimedia signal processing, visible light communication and hybrid 3DTV.



Sung-Hoon Kim received BS, MS, and PhD in electronics engineering from Kookmin University, Seoul, Korea, in 1994, 1996, and 2007, respectively. Since 2000, he has been a senior member of staff in the Broadcasting & Telecommunications Convergence Media Research Department, ETRI. His research interests are digital signal processing in the field of video communication, Mobile DTV, 3DTV, and UHDTV broadcasting systems.



Ki-Doo Kim received his BS degree in electronics engineering from Sogang University, Seoul, Korea, in 1980 and his MS and PhD degrees from Pennsylvania State University, University Park, in 1988 and 1990, respectively, both in electrical engineering. In March 1991, he joined the Department of Electronics Engineering, Kookmin University, Seoul, Korea, and is currently a professor. He worked as a visiting scholar in the Department of Electrical and Computer Engineering, University of California, San Diego, from February 1997 to February 1998. His current research interests are mobile communication and signal processing.



Kyeong-Hoon Jung received the BS, MS, and PhD from the Department of Electronics Engineering, Seoul National University, Seoul, Korea, in 1987, 1989, and 1996, respectively. After working at the Korea Broadcasting Institute from 1991 to 1997, he joined Handong University as a faculty member in the School of Computer Science and Electronics Engineering. In 2005, he transferred to Kookmin University, where he is now a professor in the School of Electrical Engineering. His research focuses on broadcasting systems, video coding, and multimedia signal processing.