EVALUATION OF 3DTV SERVICE USING ASYMMETRIC VIEW CODING BASED ON MPEG-2

Hari Kalva, Lakis Christodoulou, and Borko Furht

Department of Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL, USA

ABSTRACT

This paper presents the results of the 3DTV quality evaluation on autostereoscopic displays using asymmetric view coding. Asymmetric view coding encodes the stereo views with different quality. It has been shown that the human visual system is able to compensate for this asymmetric view quality and present a good quality 3D video. Asymmetric video coding can be exploited to reduce the bandwidth requirements for 3DTV services. The key factors that affect the asymmetric video coding are the compression algorithms, the human visual system, and the 3D display. We conducted a subjective evaluation of 3D video with asymmetric view quality and encoded using MPEG-2. We also studied the impact of eye dominance on the perceived quality. We show that asymmetric view coding can be used to reduce the bandwidth requirements of 3DTV services based on MPEG-2 view coding.

Index Terms— MPEG-2, H.264, asymmetric view coding, quality evaluation

1. INTRODUCTION

3DTV services are beginning to receive significant attention from researchers and the industry. The advance in video coding together with the advances in 3D displays will allow easier deployment of 3DTV services. One of the keys to 3D TV services is the ease of use and viewing comfort [1]. Autostereoscopic displays are emerging as good candidates for 3DTV. Autostereoscopic displays do not require 3D viewing glasses and the most common ones based on lenticular imaging technologies are relatively inexpensive to manufacture. The quality, cost, and success of 3D video services deployed will depend on the technologies used for compression, communication, and display.

3D video perception requires a pair of views, the left view and the right view, to be presented to the left eye and the right eye of users. The two views can be coded independently requiring twice the bandwidth of the traditional TV or interview coding techniques that allow prediction across view can be employed to improve compression efficiency. The bitrate required for 3D video services can be substantially reduced if the human visual system is properly exploited. Human visual system has the remarkable ability to compensate for the loss of information in one of the views and still present a very good 3D video perception. This is essentially the case where a person with perfect vision in one eye and a slightly blurry vision in the other eye is able to see the world around him normally. This ability of the human visual system can be exploited to reduce the compression of 3D video services by applying asymmetric view coding. In asymmetric view coding the left and the right eye views are encoded with different qualities without degrading the 3D experience. A study on the bounds of asymmetric view coding using H.264 was presented in [2]. A 3DTV service offered using asymmetric video coding allows multiple services with minimal bandwidth requirements. For example, the high quality view can be used for the traditional TV and the additional low quality view can be delivered only to the users with a 3DTV. This approach allows for gradual deployment of 3DTV services.

Given that MPEG-2 hardware is used virtually in all digital TV services, a 3D video based on MPEG-2 video coding should be evaluated. The more recent H.264 video is another candidate for 3DTV services. While MPEG-2 video takes significantly more bandwidth compared to H.264, the existing MPEG-2 infrastructure could be leveraged in 3D TV services. The lower bitrates required by H.264 reduces the bandwidth requirements for 3DTV services but requires new hardware deployments. Since H.264 is not expected to replace MPEG-2 in the short term, an MPEG-2 based service with asymmetric view coding offers service providers a 3DTV service with lower deployment costs. The quality of 3D video depends on the coding algorithm, 3D display used, and the human visual system. In this paper we present a quality evaluation and study the role of eye dominance on the 3D video using asymmetric view coding based on MPEG-2. The goal of this work is to understand the impact of the coding algorithms and human visual system on the perceived quality of the 3D video coded using asymmetric view coding.

The rest of the paper is organized as follows. Section 2 gives a brief overview of 3D perception, experimental methodology is presented in Section 3, the results are discussed in section 4 and conclusions presented in Section 5.

2. 3D PERCEPTION

The human visual system receives two separate projections of a scene; one from each eye. Human eyes are separated by an average horizontal distance of 2.5 inches and the images captured by the eyes are slightly different. The left and right eye views are combined in the brain resulting in a single 3D percept. The combined visual perception of the scene is also known as binocular fusion. Binocular suppression is property where portions of the view in one eye are suppressed by the corresponding view of the other eye. The possibilities of dominance and suppression mechanisms during the binocular fusion exist, but their impact is not yet well understood [3]. Experiments have shown that when the left and right eye views are combined the higher quality view is able to mask coding artifacts in the lower quality view [4, 5].

The process of binocular fusion in the human visual system results in the comparison and combination of the left and right eye views to generate a single 3D percept. The left and right eye views have to be presented to the users using 3D display means to give the sensation of 3D and depth perception. The left and right eye views can be encoded and sent to the receiver and the stereo views can be generated at the receiver. The properties of binocular fusion make possible encoding of left and right eye views at different bitrates. This asymmetric view coding has been exploited to improve compression efficiency [4, 5]. Asymmetric view coding for 3D TV based on H.264 was reported in our prior work [2]. Since the compression artifacts influence 3D perception, the effect of MPEG-2 coding artifacts is expected to the different compared with the H.264 coding artifacts. The past studies using MPEG-2 based 3D coding have not studied the impact of eye dominance on the perceived quality. The effects of the eye dominance and autostereoscopic displays on the 3D video quality cannot be understood from the past MPEG-2 based studies.

The two main approaches for delivering 3D video are 1) stereo coding where the left and right views are encoded and 2) depth image based rendering (DIBR) where a single view and an associated depth map are transmitted to the receiver. DIBR systems synthesize the left and right views at the receiver based on the single view and the depth information. We evaluate 3D TV based on stereo view coding in this work.

3. EXPERIMENTAL METHODOLOGY

The goal of this work is to understand the impact of the the eye dominance and autostereoscopic displays on the quality of the 3D video experiences. We are currently conducting a large user study to evaluate the impact of asymmetrically coded 3D views on the quality of the 3D video rendered on the Sharp autostereoscopic display. The goal of this study is to understand the bounds of asymmetric coding, relationship between the eye-dominance and 3D quality of asymmetrically coded video. The results are reported based on the evaluations from 20 users that have predicated in the study.

The sequences used for these experiments are the Akko & Kayo and the Ballroom sequences created for 3D/mulitview coding work currently underway in the MPEG committee [6]. A pair of views from these sequences was chosen to render stereo video. The video sources are 10 seconds long, 640x480 resolution, 30 FPS, and available in YUV 4:2:0 format. The Akko & Kayo sequence is made specifically for this research and has a number of carefully selected objects that help evaluation of 3D sequences well. The Ballroom sequences capture ballroom dancing and show dancers at multiple levels of depth.

The test sequences were created to test 3D video using asymmetric view coding and coded at different levels of quality. The quality was varied by encoding the left and right eye views at different qualities. Two test cases were created for each video sequence: 1) right eye view at a high quality with left eye view quality varying and 2) left eye view kept constant at a high quality and the right eye view quality varying. The high quality view that was kept constant was encoded with a PSNR of about 42.7 dB for the Akko & Kayo sequence, and the quality of the other view is varied from 42.7 dB to 33 dB. For the Ballroom sequence the high quality view that was kept constant was encoded with a PSNR of about 39.6 dB, and the quality of the other view is varied from 39.6 dB to 30 dB. In both cases, two adjacent views of the multi-view sequences were used as a stereo pair.

The subjective evaluation was done by 20 participants who evaluated the overall quality of video (without looking for specific artifacts) on the standard subjective evaluation scale from 1 (unacceptable) to 5 (excellent). For each user a test was performed to identify the dominant eye and the eyedness (left or right eye dominant) and handedness (left or right handed) information was noted. The study had 12 right eye dominant and 8 left eye dominant participants; 16 were right handed and 4 were left handed. The participants were presented with 10 second test videos separated by a 5 second mid-gray screen. The left and right eye views were encoded with quality varying from high quality to low quality and stereo pairs were created with a high quality view and a lower quality view thereby creating a 3D video with asymmetric view quality. Each video presented was randomly selected from the test set.

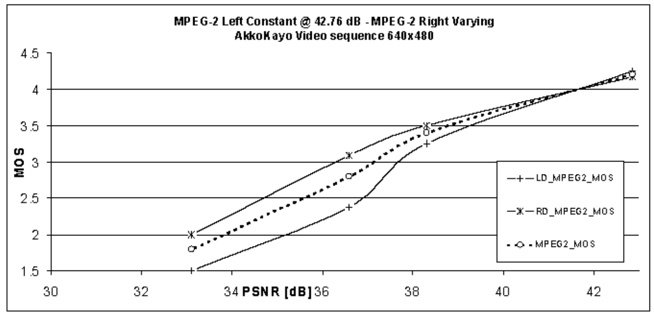
We used the Sharp LL-151-3D autostereoscopic display and a Sharp 3D laptop to evaluate the 3D quality. Both the display is 15-inches, XGA resolution (1024 by 768 pixels). This display which uses lenticular imaging techniques and renders depth very accurately gives a true 3D experience. The perception of depth is achieved by a parallax barrier that diverts different patterns of light to the left and right eye.

4. RESULTS AND DISCUSSION

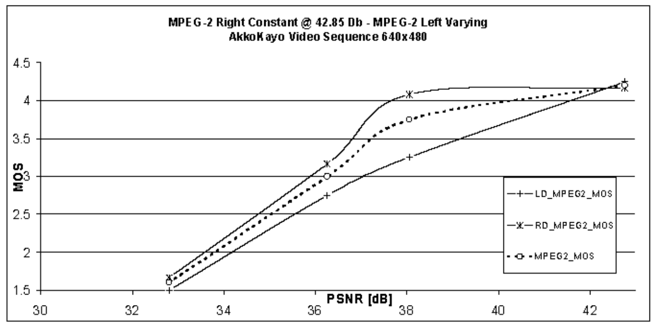
The quality of a single 2D view alone is not an indication of the 3D quality. The PSNR of the individual views can be used as a metric when the quality of both views is high. With asymmetric view coding quality is influenced by the coding artifacts present in the individual views, the human visual system, and the type of 3D display. Developing objective quality metrics for 3D quality is thus very difficult and subjective evaluation is the primary means of evaluating 3D video quality. We evaluated the role of the human visual system by studying the role of eye dominance.

Humans have a preference for one eye over the other, referred to as eye dominance. The significance of eye dominance, however, is not well understood. Humans are mostly right handed (90%) and about 70% are right eyed, 20% left eyed, and 10% exhibit no eye preference [7]. A recent study suggested that the eye dominance just indicates individual sighting preferences and has no function in binocular vision [8]. Another study found that eye dominance improves the performance of visual search tasks by perhaps aiding visual perception in binocular vision [9]. Our results also suggest a role for eye dominance in 3D perception when asymmetric coding is used.

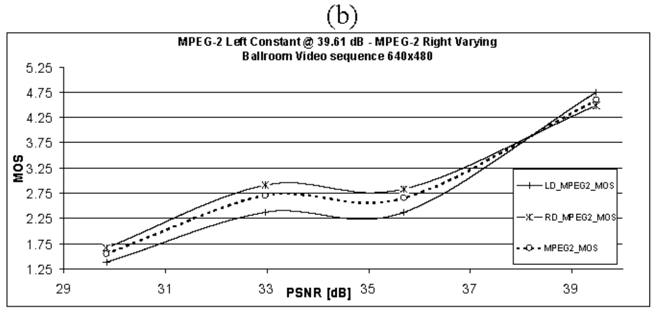
Mean opinion scores were computed for the test sequences based on subjective evaluations. Figure 1 shows the mean opinion scores (MOS) for the Akko & Kayo and Ballroom sequences. Figures 1.a and 1.c show the MOS with left eye view coded at a high quality while the quality of right eye view in each evaluated sequence was varied from high to low. Figures 1.b and 1.d show the MOS with right eye view coded at a high quality while the quality of left eye view was varied. Figure 2 shows the frame 100 of the two views the Ballroom sequence coded with asymmetric quality. As the quality of a view drops, the MOS of the perceived 3D quality also drops. The rate of drop in MOS is different for the Akko & Kayo and Ballroom sequences, The Akko & Kayo sequence is slow motion sequence and artifacts are more easily detected. The Ballroom sequence has lot of motion with dancers moving across the screen and the drop in perceived 3D quality is not as steep. In both cases the perceived quality drops faster as the asymmetry in the coded view quality increases.



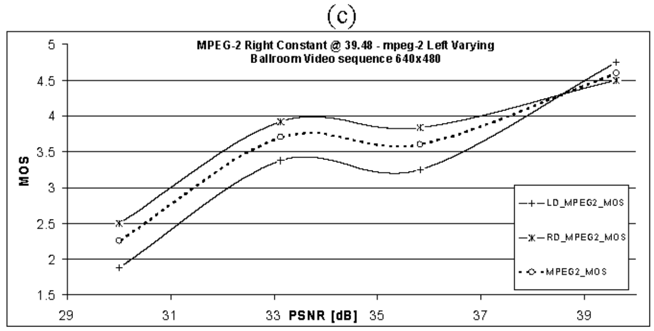
(a)



(b)



(c)



(d)

Figure 1. Mean Opinion Score of subjective evaluation tests for asymmetric view coding

The plots also show the influence of eye dominance. The figures also show the MOS of users grouped by their eye dominance. The perceived 3D video quality is clearly influenced by the eye dominance. The role of eye dominance seems to increase as the asymmetry in the coded view quality increases. The plots show that the right eye dominant users perceive better quality irrespective of the quality of the right eye view. The left eye dominant users are more sensitive to the view quality asymmetry.

Figure 2. Stereo views with asymmetric view quality. Top: right view coded at 6 Mbps (39.6 dB PSNR); Bottom: left view coded at 1 Mbps (30 dB PSNR).

5. CONCLUSION

This paper evaluates the use of asymmetric view coding for 3D video services. With asymmetric view coding, quality is influenced by the coding artifacts present in the individual views, the human visual system, and the type of 3D display. The wide use of MPEG-2 infrastructure in digital TV services means a careful study of 3D TV using MPEG-2 is required. Though MPEG-2 has a higher bandwidth requirements compared to H.264, the use of asymmetric view coding can reduce the overall bandwidth required for 3D video services. Services employing asymmetric view coding have to be carefully deployed as the perceived quality is influenced by the eye dominance of the end users. The right eye dominant uses seem to perceive better 3D quality and less sensitive to view quality asymmetric compared to the left eye dominant users. Since about 20% of the general population is left eye dominant, the stereo views cannot have large asymmetry without affecting the perceived quality of the end users. Asymmetric view coding presents an interesting option for reducing the bandwidth requirements of 3D video services. The asymmetric view coding can be exploited further if content based encoding is employed. The lower quality view can be coded such that the visual cues that contribute to 3D perception are coded with a higher quality compared with the regions without any depth cues.

6. REFERENCES

[1] H. Kalva, L. Christodoulou, L. Mayron, O. Marques, and B. Furht, "Challenges and opportunities in video coding for 3D TV," IEEE International Conference on Multimedia & Expo (ICME) 2006, July 9-12, 2006, Toronto, Canada.

[2] H. Kalva, L. Christodoulou, L. Mayron, O. Marques, and B. Furht, "Design and Evaluation of 3D Video System Based on H.264 View Coding," International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2006), Newport, Rhode Island, May 22–23, 2006, pp. 68-73.

[3] O. Schreer, P. Kauff, and T. Sikora, edts., "3D Video Communications" Wiley 2005.

[4] Lew B. Stelmach, W. James Tam, "Stereoscopic image coding: Effect of disparate image-quality in left- and right-eye views", Signal Processing: Image Communication, Vol. 14, pp.111-117, 1998.

[5] Daniel V. Meegan, Lew B. Stelmach, and W. James Tam, "Unequal Weighting of Monocular Inputs in Binocular Combination: Implications for the Compression of Stereoscopic Imagery", Journal of Experimental Psychology: Applied, Vol. 7(2) 143-153, Jun 2001.

[6] ISO/IEC JTC1/SC29/WG11, "Call for Proposals on Multi-view Video Coding (MVC)," MPEG Document MPEG2005/N7327, July 2005.

[7] D.C. Bourassa, I.C. McManus, and M.P. Bryden, "Handedness and eye-dominance: A meta-analysis of their relationship," Laterality, Vol 1, No. 1, 1996, pp. 5–34.

[8] A.P. Mapp, H. Oho, and R. Barbeito, "What does the dominant eye dominate? A brief and somewhat contentious review," Perception & Psychophysics, Vol. 65, No. 2, 2003, pp. 310– 317.

[9] E. Shneor and S. Hochstein, "Effects of eye dominance in visual perception," International Congress Series, Volume 1282, Vision 2005, September 2005, pp. 719-723.