

A Study of Compressed 3D Video Perception based on Subjective Quality Assessment

Xiangkun Du and Xiuhua Jiang and Caihong Wang

Abstract This paper presents a specific subjective quality assessment to study the perceptual issues about 3D videos processed by the compression system based on H.264/MVC (Multi-view Video Coding). The subjective evaluation is conducted according to the ITU recommendations. The preparation, concrete process and subjective results are presented in detail. By conducting the subjective assessment to both 2D and 3D compressed sequences using Double Stimulus Impairment Scale method (DSIS), this research shows that people have different sensitivity to impairments in 2D and 3D compressed videos. Since ITU recommends three primary assessment dimensions -- picture quality, depth quality and visual comfort -- for subjective assessment of stereoscopic 3DTV systems, our subjective results also showed that 3D video compression based on MVC causes different degree of damage in different assessment dimensions.

Keywords: 3D video, MVC, Video perception, Subjective quality assessment

1 Introduction

With the development of 3DTV techniques, people now can watch 3D programmes and 3D movies at home using a 3D TV set. The introduction of 3DTV services require the development of new image formats, image processing and transmission techniques, whose performance will need to be evaluated through subjective quality assessment.

Studies of 3D video quality assessment based on subjective experiments should be conducted in accord with the international recommendations. Currently, the standardization of subjective quality evaluation for stereoscopic video is led by

Xiangkun Du (✉)

College of Information Engineering, Communication University of China, Beijing, China
e-mail: duxaingkun@cuc.edu.cn

Xiuhua Jiang and Caihong Wang

College of Information Engineering, Communication University of China, Beijing, China

International Telecommunication Union (ITU). Its Study Group 6 (SG6) devotes itself to the research of subjective quality evaluation. These two recommendations published by ITU -- ITU-R.BT.1438 [1] and ITU-R.BT.2021 [2] -- are the most complete standard for subjective quality assessment of 3D videos at present.

Therefore, how to set up the experimental system (including environment, equipment, test material, methodology, etc.) in conformity with the international standards becomes especially important. In addition, the perceptual issues about what kind of subjective feelings people will have when they are watching different compressed 3D videos with different quality and what is the difference between watching 2D and 3D videos are problems worthy of study.

In order to study the 3D video perception, especially the perceptual differences between 2D and 3D videos after compression, we conducted a subjective experiment. The preparation together with the process of the assessment were carried out adopting the ITU recommendations and presented in this paper in detail.

This paper is organized as follows: in Section 2, the preparation of the subjective experiment is introduced; in Section 3, the specific procedure of the assessment is discussed; the subjective results and analysis are given in Section 4; and finally in Section 5, conclusions and the future direction of the work are presented.

2 The Preparation of the Subjective Experiment

In this section, the equipment and test materials used in the experiment will be discussed.

2.1 Equipment and Viewing conditions

The subjective evaluation experiment was carried out at Communication University of China, in a lab especially established to satisfy the recommendations given in ITU-R BT.2021 [2]. The display equipment used for presenting the sequences was a 52" SHARP stereoscopic television, with resolution of 1920x1080 pixels and aspect ratio 16:9. To visualize the 3D sequences, the observers wore the active shutter glasses associated to this TV set. The subjects were positioned at a viewing distance of 3.1 times the height of the TV set (approximately 2m), which is recommended in ITU-R BT.2021. All the sequences were played by a video server which has a Blackmagic DeckLink 3D+ video board in it [3]. We used Media Express as the playback software solution, which is also provided by Blackmagic Design. The video server is connected to the TV set

by HDMI. With all these equipment, we can present full HD 3D videos to the observers.

2.2 Test Material

Two 3D videos from The Institute of Image Information and Television Engineers (ITE) [4] were used in the subjective experiment. The ITE test materials are described in ITU-R.BT.1438 [1], which is informatively referred to by ITU-R.BT.2021 [2]. The screenshots of these 2 sequences “*Festival*” and “*Botanical garden*” are presented in Figure 1, and Table 1 provides a brief summary of the 2 test materials.



Fig. 1 The screenshots of “*Festiva*” (left) and “*Botanical garden*” (right)

Table 1 A brief summary of the test materials

Name	Length	Content	Major factors
<i>Festiva</i>	10s	A portable shrine and paper storm	Still and depth resolution
<i>Botanical garden</i>	10s	Waterfall in a botanical garden	Orthostereoscopy

To compress these sequences, we used the encoder *JMVC v8.5* which is based on H.264/MVC (Multi-view Video Coding). By using different quantization parameters (QP) in configuration, we attained several sequences with different bitrates. Then, we conducted a subjective test in a small range to decide which bitrates we can use in the final experiment. We gave consideration to both high quality and low quality sequences. To find out the exact relationship between the bitrate and subjective opinion, we chose 7 bitrates for each sequence at last (Table 2) which were a little more than normal subjective experiments.

Table 2 The 7 bitrates used in the experiment

Bitrate of left view (Mbps)	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7
<i>Festiva</i>	3.26	4.69	5.59	7.72	8.89	10.31	11.90
<i>Botanical garden</i>	1.90	2.57	3.50	4.76	5.48	6.29	7.33

2.3 Observers

In the subjective experiment participated 22 observers, all of them are non-experts in the technical aspects of 3D videos. The ages of the subjects were ranged between 20 and 29 years old, with an average age of 25. Before the experiment, visual test and stereoscopic vision test were taken by all the subjects. We used “*Worth 4 Dot Test*”, “*Circle Test*”, “*Dynamic Random Dot Stereogram Test*” and “*Acuity Test*”, which are recommended in ITU-R.BT.2021 [2] and the results showed that all the observers have normal visual and stereoscopic vision.

3 Evaluation Methodologies

After the preparation of the experiment, the test materials are compressed by the encoder, and the equipment and environment are ready to present these sequences. In this section, the procedure of the evaluation experiment will be described in detail.

3.1 Subjective Methodology

To determine the most suitable subjective methodology for this study, a test had been applied in a small range of subjects. We found out that when using the subjective methodologies described in [5]: *Single Stimulus method (SS)*, *Double Stimulus Continuous Quality Scale method (DSCQS)* and *Absolute Category Rating (ACR)*, subjects need to score a certain sequence directly (without comparison), which is very difficult for observers who are lack of experience in watching 3D videos. Meanwhile, *Double Stimulus Impairment Scale method (DSIS)* described in ITU-R.BT.500 [6] presents both the original sequence and the impaired sequence and the observers need to score based on the injury degree of each impaired sequence. By comparing the quality of the original and compressed video, it is easier for the observers to give a subjective opinion score. Considering these factors, we used DSIS method as our subjective methodology. According to ITU recommendation, a hidden original sequence was used in the experiment, and with the 7 compressed videos, it turned out to be 8 tests in all for each sequence.

3.2 Assessment Dimensions

Three primary assessment dimensions, or perceptual dimensions, are defined in ITU-R.BT.2021 [2]: *Picture Quality*, *Depth Quality* and *Visual Comfort*. *Picture Quality* refers the perceived quality of the picture provided by the system. This is a main determinant of the performance of a video system and is basically identical with the 2D subjective assessment. *Depth Quality* and *Visual Comfort* refer the ability to provide a good sense of depth without discomfort. In our experiment, we finally chose two perceptual dimensions which are *Picture Quality* and *Sense of Depth*.

In order to find out the differences between 2D and 3D video perception, we conducted two subjective assessments for 2D video and 3D video respectively. In the 2D assessment, the monoscopic version (2D videos) had been presented in 3D mode (e.g. the left-view presented to both the left and right eyes using the same 3D hardware settings as for the actual stereoscopic sequence). Figure 2 and Figure 3 show the specific evaluation procedure. The time duration of the whole experiment is within 40 minutes which will not cause visual fatigue according to ITU.

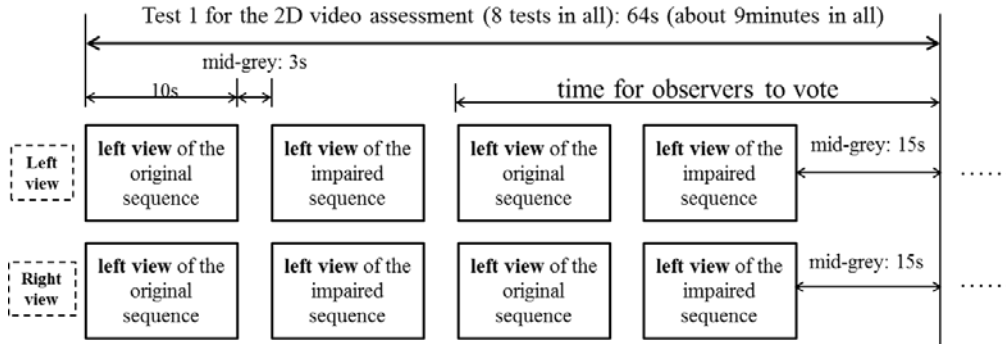


Fig. 2 The assessment procedure of 2D videos

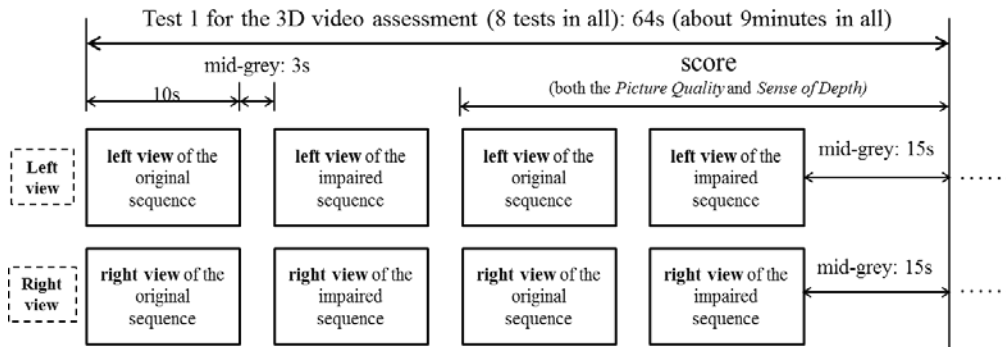


Fig. 3 The assessment procedure of 3D videos

Before the assessment, instructions had been tailored to the dimensions (e.g. *Picture Quality* and *Sense of Depth*) under investigation. The test material “*Red leaves*” used in the instruction is another sequence from ITE.

3.2 Grading

The five-grade impairment scale has been used and Table 3 shows the explanation of each grade. According to the feedback of the observers, these explanations printed on the questionnaire are very helpful when the subjects are voting.

Table 3 The explanation of five-grade impairment scale

Picture Quality	1	poor quality and very annoying
	2	relatively poor quality with obvious impairments
	3	medium quality with distinct differences
	4	can see the differences, but not distinct
	5	unable to detect the differences of video quality
Sense of Depth	1	very obvious differences, cannot get 3D vision
	2	can lead to lapse in judgment of spatial position
	3	indeterminate sense of depth, distinct differences
	4	can see some differences in sense of depth
	5	no differences in spatial position and stereoeffect

4 Results and Analysis

The evaluations collected from the observers that participated in the subjective quality assessment were processed to obtain the Mean Opinion Scores (MOS) and the 95% confidence intervals of the values according to ITU-R.BT.500 [6]. In addition, a screening of the results was done to discard possible random evaluations. The obtained results of the two video sequences in three perceptual dimensions are presented in Figure 4.

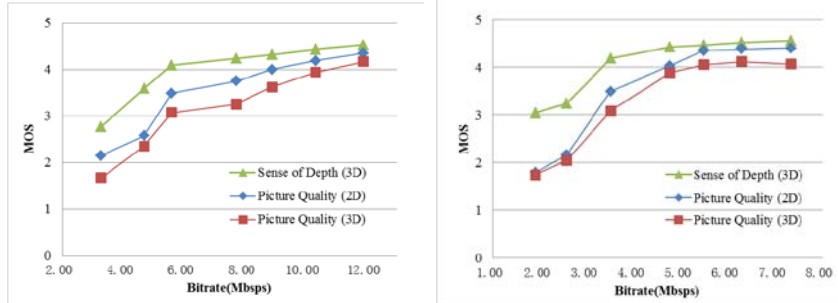


Fig. 4 The subjective MOS of “*Festival*” (left) and “*Botanical garden*” (right)

Firstly, from the two scatter diagrams above, we can see that the MOS of all the three dimensions -- sense of depth of 3D videos (green), picture quality of 2D videos (blue) and picture quality of 3D videos (red) -- increase with the Bitrate in both 2D and 3D sequences, which is to say, a higher bitrate will basically lead to a higher quality.

Secondly, considering the differences between three curves -- green, blue and red -- we can see no overlap between them, which means that the impairments caused by the encoder is not the same in 2D and 3D videos as well as in different perceptual dimensions.

The green curve on the top illustrates that when the bitrate is quite low (e.g. about 5Mbps for “*Festival*” and 3.5Mbps for “*Botanical garden*”), the MOS of *Picture Quality* is also very low (about 2.5), but the score of *Sense of Depth* is still greater than 3.5. It means that compression based on H.264/MVC causes less damage to the sense of depth than to the picture quality. As we know, the compression theory adopted by MVC is based on H.264, which is already widely used in the 2D video compression. The damages introduced to the images of left and right view primarily exist in the picture quality of each view, and the sense of depth or the positional relationship of the scenes being shot basically remain the same before and after the encoding. Therefore, we believe that for evaluating the codec systems used for compression and transmission, the subjective experiments should be focused on the assessment of the impairments in *Picture Quality* and *Visual Comfort* caused by the system, and the evaluation of *Depth Quality* may not bring expected results.

Furthermore, in both of the two sequences -- “*Festival*” and “*Botanical garden*” -- the red curve (picture quality of 3D videos) is under the blue one (picture quality of 2D videos) almost in every bitrate. It means that when the observers were presented a 2D impaired video sequence and a 3D one with the same bitrate for the left view, most of the observers gave a higher score to 2D videos and lower to 3D ones, that is to say, the impairments in 3D videos are more visible than in 2D sequences. We know that when watching the stereoscopic videos, the impairments in both of the left view and the right view will be perceived. For the MVC compression method, the two views have different bitrate

and different picture quality, and the perceived impairments always appear in the different areas in every frame in each view. The differences between the impairments in left and right views may result in a lower subjective quality than 2D videos. It can lead to a conclusion that in 3DTV broadcasting, a higher bitrate for each view should be used to attain the same picture quality in HDTV.

5 Conclusions

In this paper, we presented a subjective quality assessment to find out some conclusions about compressed 3D video perception. The subjective assessment for both 2D and 3D videos are set up and conducted based on ITU recommendations. The DSIS method was used as the subjective methodology and the results are presented and analyzed in detail. We found that video compression based on MVC causes less damage to the sense of depth than to the picture quality. It may have guiding significance when choosing the assessment dimensions for subjective quality evaluation against codec systems. We also found out that the impairments in 3D videos are more visible than in 2D ones. It may lead to a conclusion that to acquire the same picture quality in HDTV, a higher bitrate for each view is needed in 3DTV broadcasting. In the future, more subjective assessment should be done using a larger range of test materials as well as more observers to obtain more universal rules about the relationship between bitrate and 3D video quality.

Acknowledgments This work is supported by the National Sci-tech Support Plan of China: The Research of Stereoscopic Content Transmission Technology and Systems. (No. 2012BAH39F02).

References

1. International Telecommunication Union, 2000. ITU-R.BT.1438: Subjective assessment of stereoscopic television pictures. Geneva: ITU.
2. International Telecommunication Union, 2012. ITU-R.BT.2021: Subjective methods for the assessment of stereoscopic 3DTV systems. Geneva: ITU.
3. Blackmagic Design, 2013. Blackmagic Design: DecLink. [online] Available at: <<http://www.blackmagicdesign.com/products/decklink>> [Accessed 11 June 2013].
4. ITE, 2013. Stereoscopic Imaging Standard Charts_Ver.2. [online] Available at: <http://ite.or.jp/en/p_t/test_chart/?mode=disp&key=29&lid=&sort=&word=&page=1> [Accessed 11 June 2013].
5. Kawano T., 2012. Performance comparison of subjective assessment methods for 3D video quality. Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on. IEEE, 2012: 218-223.
6. International Telecommunication Union, 2012. ITU-R BT.500-13: Methodology for the subjective assessment of the quality of television pictures. Geneva: ITU.