Visual Discomfort from Camera Panning in 4K-UHD Contents

Sang-IL Kim
School of Media
Seoul Institute of the Arts
Ansan City, Gyeonggi Province,
Department of Holography 3D
Contents Kwangwoon University
Seoul, Korea
silgam@seoularts.ac.kr

Seung Hyun Lee
Graduate School of Information and Contents
Kwangwoon University
Seoul, Korea
shlee@kw.ac.kr

Taeuk Jeong*
Department of Computational Science and Engineering
Yonsei University
Seoul, Korea
iamlogin@yonsei.ac.kr

Abstract—In this paper, we analyze the visual discomfort caused by camera panning in 4K UHD contents. UHD is expected to provide a sense of realness to viewers, with four times more pixels than HD. However, UHD may yield undesirable artifacts, such as the stroboscopic effect from displaying frequency and visual fatigue due to moving objects. To investigate the visual discomfort in the presence of camera panning, subjective evaluation is conducted using video clips that are created considering the distance between the camera, characters and background. Subjective scores illustrate that the effect of camera panning is strongly correlated with the camera panning speed, and weakly related to the distance among three elements. From the subjective test, we conclude that visual discomfort occurs when the object moves at a speed over 2560 pixels per second, or 43 pixels per frame in 4K UHD contents, and 96 cm per second, or 1.6 cm per frame in a 65 inch monitor.

Keywords: visual discomfortness, 4k-UHD, sense of realness, camera panning

I. INTRODUCTION

Ultra High Definition (UHD) imaging technology is now widely spread and used in TV broadcasting and cinema to provide a better visual experience by improving many features of High Definition (HD). UHDTV could be characterized as a realistic TV system producing a sense of realness and of “being there” to viewers, supported by the spatial resolution, which is four times higher than HDTV, and a wide field of view (FOV). However, there are a number of issues in the procedure of creating UHD contents. The technical difficulties include the time and cost limitations of rendering, editing, and real time monitoring in the UHD production workflow due to vast amounts of raw file data. Another difficulty is related to implementation of realness or “being there” in the UHD contents without any artifacts of flicker, motion blur, jerkiness, and color. It may require more consideration on many elements established in HD production to provide a sense of presence to the viewers.

There have been various researches to maximize the quality and effect of immersive UHD videos in spatial and temporal aspects [1-3]. For the spatial resolution, the relation between the sense of realness and viewing distance [4] was reported. Psychological evaluation of the “sensation of reality” and viewing angle [5] was investigated, and guidelines for the optimal viewing distance and the optimal horizontal viewing angle [6, 7] have been proposed. For the temporal artifact, the stroboscopic effect (flicker, motion blur, and jerkiness) for motion reproduction was reported in terms of frame frequency [4]. For the bit depth of the image, more than 10 bit per pixel [8] is recommended to avoid banding (or contouring) effect and to improve the quality and effectiveness of UHD contents compared to HD videos. Beside quality degradation, UHD videos may yield visual discomfort when displayed in a large monitor. For example, viewers may feel motion sickness when the entire region on the screen moves very fast due to camera panning or camera moving.

In this paper, we investigate the visual discomfort from camera panning in 4K UHD contents. In the subjective evaluation, we analyze the effect of camera panning on visual fatigue with various panning speeds and distances between the camera and characters.
II. TEST VIDEOS AND METHOD FOR SUBJECTIVE EVALUATION

A. Preparing Test Videos

Test videos are created by considering the camera panning situation for conventional interview circumstances. As seen in Fig. 1, the camera pans through three characters from left to right, and the distance between two characters is 1 m, which is normal distance in conventional interview stand. The distance between the camera and characters is set to 2m (close interview distance), 4m, and 6m (indoor studio shooting distance), maintaining a bust shot of characters. The camera panning speeds are from 1 second to 5 seconds, and the videos have 1-5 second running time same as the camera working. All videos have the 4K UHD (3840x2160) resolution with 60 frames per second. For the test, 1 second still images are added in the front and back of each video same as the first frame and last frame, respectively.

B. Subjective Evaluation Method

For the subjective evaluation, we employ and modify the single stimulus (SS) method in ITU-R BT.500 [9], which displays the test materials in random order. As seen in Fig. 2(a), the gray still image with 1 second is added in the front for identifying the test video index, and another gray image is followed by the video clip for evaluation with 5 seconds.

To evaluate the visual discomfort from camera panning in a 4K UHD video, the 5-grade impairment scale (Imperceptible, Perceivable but not annoying, Slightly annoying, Annoying, Very annoying) is used after translating the original English words in Fig. 2(b) into Korean words. In adjectival categorical judgements, viewers assign a video clip to one of a set of categories that, typically, are defined in semantic terms.

C. Subjective Test Environments

The subjective test is conducted under the same configuration as the home viewing environment. A consumer-grade 65 inch monitor with 4K resolution and LED type is used for the test. Test video clips are encoded by HEVC codec with MPEG2-TS and displayed in 60 frames per second. Fifteen non-expert viewers participate in the subjective evaluation, who have normal (20/30) visual acuity with or without corrective glasses (per Snellen test or equivalent) and normal color vision (per Ishihara test or equivalent). The viewing distance is set to 1.5H.

III. VISUAL DISCOMFORT FROM CAMERA PANNING

For the subjective evaluation, a total of 60 video clips are created. The 45 test video clips are created considering the distance between the camera and characters (2m, 4m, 6m), and the distance between the characters and background (2m, 4m, 6m) with the camera panning speed (1s, 2s, 3s, 4s, 5s). And the 15 test video clips are additionally included, in which three characters do not exist but the distance between the camera and character position is set to 2m with 5 camera panning speeds. The former clips are used to analyze the effect of camera panning, whether the distance between the characters and background affects visual fatigue, while the latter ones are used to compare the results to see whether the visual discomfort may be influenced with or without the characters when the camera is panning.

The subjective scores from the 15 viewers are averaged to be used as the mean opinion score (MOS) for each clip. From the test sheet, a video clip would be regarded to yield visual discomfort if its MOS is no more than 3 points, and no visual fatigue otherwise. We denote the distance between the camera and characters by d(C, P), and distance between the characters and background by d(P, B).

A. Camera working distance in the test video clip

In each video clip, the camera pans from the left character to the right one with a bust shot during the given panning time. The panned video clip is decomposed with three panoramic images in Fig. 3. The camera working distance for the test video clip is 7680 pixels in 4K UHD contents.

B. The effect of camera panning at varying distances between characters and background

Fig 4 shows the results of subjective scores when d(C, P) is fixed and d(P, B) is varied from 2m to 6m. The MOS increases as the camera pans slowly, and is greater than 3.5 when the camera panning speed is not less than 4 seconds. That is, viewers may feel visual discomfort when the camera panning speed is 1-3 seconds. For the fast panning speed, for example, clips with 1 second panning yield an annoying level of discomfort. The effect of camera panning may be small depending on the distance between characters and background, except for the case of the panning speed 2 when d(C, P)=2, in
which the differences are about 1 point but MOS’s are less than 3 points (slightly annoying level).

(a) d(C, P)=2

(b) d(C, P)=4

Fig. 4. MOS distribution when d(C, P) is fixed.

C. The effect of camera panning for characters

Fig. 5 shows the MOS distribution of test video clips with or without characters along with the panning speed for d(C, P)=2m and d(P, B)=6m. For the same distance between the camera and background, the video clips with or without characters produce small difference in MOS, except for the case of the camera panning speed 2, in which the differences are less than about 1 point but MOS’s indicate less than 3 points (slightly annoying level). The subjective evaluation illustrates that the effect of camera panning may be small, whether the video contents include characters or not, and the visual discomfort is related to the camera panning speed.

D. Critical pixel moving for visual fatigue in camera panning

As mentioned, objects in the test clip move a total of 7680 pixels when the camera is working. The subjective scores show that the critical panning speed is 3 seconds in this test when the viewers feel visual discomfort. By a simple calculation, discomfort comes from the object moving at a speed over 2560 pixels per second, or 43 pixels per frame in 4K UHD contents. Since a 65 inch monitor is used in the test, discomfort occurs when the objects are moving at a speed over 96 cm per second, or 1.6 cm per frame in the monitor.

IV. CONCLUSION

To analyze the visual discomfort in the presence of camera panning, we conduct the subjective evaluation of 60 video clips which contain various panning environments, considering the distance between the camera, characters, and background. According to the subjective test scores, viewers may feel visual discomfort when the object moves at a speed over 2560 pixels per second, or 43 pixels per frame in 4K UHD contents, and 96 cm per second, or 1.6 cm per frame in a 65 inch monitor.

References


