On the first JND and Break in Presence of 360-degree content: An exploratory study

Roberto G. de A. Azevedo
roberto.azevedo@epfl.ch
EPFL
Lausanne, Switzerland

Neil Birkbeck
birkbeck@google.com
Google, Inc.
Mountain View, CA, USA

Ivan Janatra
janatra@google.com
Google, Inc.
Mountain View, CA, USA

Balu Adsumilli
badsumilli@google.com
Google, Inc.
Mountain View, CA, USA

Pascal Frossard
pascal.frossard@epfl.ch
EPFL
Lausanne, Switzerland

ABSTRACT

Unlike traditional planar 2D visual content, immersive 360-degree images and videos undergo particular processing steps and are intended to be consumed via head-mounted displays (HMDs). To get a deeper understanding on the perception of 360-degree visual distortions when consumed through HMDs, we perform an exploratory task-based subjective study in which we have asked subjects to define the first noticeable difference and break-in-presence points when incrementally adding specific compression artifacts. The results of our study: give insights on the range of allowed visual distortions for 360-degree content; show that the added visual distortions are more tolerable in mono than in stereoscopic 3D; and identify issues with current 360-degree objective quality metrics.

KEYWORDS

360-degree video, visual distortions, visual quality, JND, presence

ACM Reference Format:

Roberto G. de A. Azevedo, Neil Birkbeck, Ivan Janatra, Balu Adsumilli, and Pascal Frossard. 2019. On the first JND and Break in Presence of 360-degree content: An exploratory study. In *11th International Workshop on Immersive Mixed and Virtual Environment Systems (MMVE '19), June 21, 2019, Amherst, MA, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3304113.3326115

1 INTRODUCTION

360-degree images and videos mapped to a planar 2D image [3] and undergoing different processing stages can be affected by visual distortions (or artifacts) that impact the end-user perceived quality of experience (QoE) [1]. Although the perceived QoE of 360-degree content is related to traditional 2D and stereoscopic 3D visual content, the perceptibility of distortions in processed 360-degree content visualized via head-mounted displays (HMDs) brings its own specificities, including new types of distortions [1] such as the magnification of the content, an increased field-of-view, and the fact that the user is completely immersed in the content, which completely changes a users QoE perspective.

In this paper we propose an exploratory study to get a deeper understanding of the perceptibility of compression artifacts in 360-degree content consumed through HMDs. We perform a task-based subjective study in which the user is asked to interactively find the *first just-noticeable distortion* (JND) and the *Break-in-presence* points when incrementally adding distortions to 360-degree visual content. JND is a statistical quantity that accounts for the maximum distortion that stays unnoticeable to a human [11]. We define the "1st JND" as the maximum amount of distortion that can be added to the pristine (non-distorted) content before the user perceives any difference in it. Presence is defined as the sense of "being there," inside a space, even when physically in a different location [5]. We define the Break-in-presence point as the one in which the amount of visual distortion cannot induce presence anymore.

Unlike previous subjective studies on 360-degree content [4, 6, 7, 10], which mainly focus on the overall visual quality (usually mixing different artifacts) we are interested in understanding the perceptibility and the impact of specific compression artifacts on VR quality and the sense of presence. Indeed, although quality assessment of 360-degree video has been an active research area in recent years, none of the previous work focused on how the specific compression artifacts are perceived by end-users wearing an HMD. In particular, we consider the *blur*, *blocking*, *cube map seams*, and *H.264* artifacts. Blur, blocking, and H.264 are also common artifacts in traditional 2D images. In processed 360-degree visual content, however, due to the interaction with the used projection method, rendering process, and the content magnification on HMDs, they are perceived quite differently, e.g., as radial blocking or visible seams [1]. Seams are 360-degree specific and appear due to the lossy compression on discontinuities areas in the planar domain [1].

The rest of the paper is organized into: Section 2 details the methodology, stimulus preparation, and subject profiles. Section 3 presents the results. Section 4 discusses our conclusions.

2 EXPLORATORY STUDY

For our experiment, a Lenovo Mirage Solo HMD (2560x1440@75Hz) was used, and the user sat on a swivel chair so that they were free to explore the content.

2.1 Methodology

Fig. 1 depicts the methodology, which is divided into:

MMVE '19, June 21, 2019, Amherst, MA, USA

ACM ISBN 978-1-4503-6299-3/19/06...$15.00
https://doi.org/10.1145/3304113.3326115

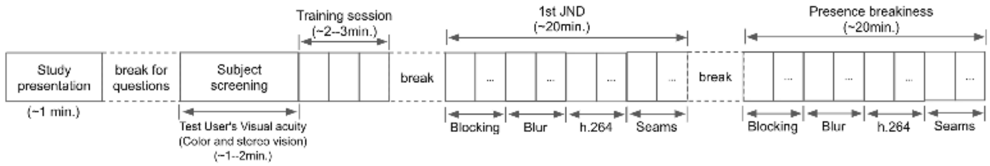

Figure 1: Methodology

Figure 2: User interface

Study presentation The subjects are presented with the goal of the study and informed that it is possible to feel some discomfort and/or sickness (due to the use of the HMD), in which case, they are free to leave the experiment at any time.

Pre-experiment subjects screening The subjects are screened for color blindness and stereoscopic vision issues. The color blindness test is based on the Ishihara Color Test[1]. For the stereoscopic vision test, we show a stereoscopic 3D video pattern on the YouTube VR app using the HMD. The participants are able to see different numbers if they have normal stereoscopic vision.

Training The subject wears the HMD and is presented to the overall process and user interface (UI). An example of 360-degree content, which is not part of the study, is presented so that the user becomes familiar with the overall process and UI. The data gathered in this session is discarded.

1st JND The subject is presented with multiple sequences and asked to choose the point in which they see the 1st JND. For that, the subject uses a slider UI (Fig. 2) that allows to gradually add or reduce the amount of distortion (0%: pristine; 100%: most degraded) until they can notice a difference with regards to the original. After rating the content, the subject advances to the next one. There is no time limit on how long the subject can explore the content.

Break-in-Presence The process and UI are similar to the 1st JND session, but the user is asked to choose the Break-in-presence point. The information provided to the subject on what is presence and the Break-in-presence points are the definitions discussed in Section 1. Given its subjective nature, however, some users were uncertain about how to define the exact Break-in-presence point. In this case —inline with our goal of defining the maximum allowed distortion for subjective studies— we mentioned that the subject could choose the one where the amount of distortion becomes unacceptable. Although our process of directly asking the subjects to define the Break-in-presence point provides us with less information than standard questionnaires for measuring presence [9], it is advantageous because it is easily scalable (there is no need to ask questions for each of the distortions levels). How to adapt presence questionnaires to find the break-in-presence point is interesting, but still an open research question, left as future work.

2.2 Stimuli

Four original pristine 360-degree video source content (Fig. 3) were used in the study, two stereoscopic (*Cartoon* and *Obama*) and two monoscopic (*Alcatraz* and *Aerial*) ones. For the stereoscopic sources, we also include their monoscopic versions, resulting in 6 reference sequences. The source sequences are in the equirectangular (ERP) projection [3] format, with resolutions of 7680x3840 (*Alcatraz* and *Aerial*), 5120x5120 (*Obama*), and 7296x4104 (*Cartoon*). We re-projected them to the equiangular cubemap (EAC) [2] format using a 3840x2160 resolution. Finally, we extracted the first frame to serve as original content for our study, from which, we generated the different distortions levels for each of the different artifacts.

Table 1 shows the ffmpeg (blocking, blur, and H.264) and Matlab (seams) commands used to generate the distortions levels. The blocking distortions are inserted using the H.263 algorithm with different quality parameters (QP). The blur distortion is added through a Gaussian filter on the whole frame. To avoid the appearance of seams in the blur content, the blur command in Table 1 is performed on a padded version of the EAC frame. Then, the correct resolution is cropped from it. The seam distortion is simulated through a blur filter only in the discontinuities (i.e., where the cube faces meet) in the EAC content. (We show only the monoscopic version of the code. The stereoscopic version follows the same principle, but in the vertical discontinuities of the frame.) Finally, the H.264 distortions are created using the H.264 algorithm with different QPs.

Table 1: Commands to generate the different artifacts.

Artifact	Command

blocking	`for q in $(seq 1 1 31); do`
	`ffmpeg -i ${frame}.png -vcodec flv -q $q out${q}.flv`
	`done`

blur	`for bb in $(seq 0 0.15 5); do`
	`ffmpeg -i ${frame}.png -vf "gblur=sigma=${bb}" out${bb}.png`
	`done`

seams	`seam_height=32;`
	`for Q=1:30`
	`[h, w, c] = size (frame); F = frame;`
	`for r=(h/2-seam_height/2):(h/2+seam_height/2)`
	`for c=2:w-2`
	`F(r, c) = 1/16*sum(sum([1 2 1; 2 4 2; 1 2 1].*F(r-1:r+1, c-1:c+1)));`
	`end`
	`end`
	`imwrite (F, ['out' Q '.png']);`
	`end`

H.264	`for q in $(seq 0 1 51); do`
	`ffmpeg -i ${frame}.png -c:v libx264 -crf $q out${q}.mp4`
	`ffmpeg -i out${q}.mp4 out${q}.png`
	`done`

2.3 Subjects

We recruited 14 subjects for the experiments, 13 males and 1 female, with ages ranging from 24 to 64. All subjects had normal or corrected-to-normal stereoscopic vision, and two were color blind. One of the color-blind subjects did not complete the experiments; thus, the corresponding data was removed from the experiments. The subjects were randomly divided into two equal-sized groups: one that performed first the 1st JND and then the Break-in-presence sessions; and another that did the opposite.

3 RESULTS AND ANALYSIS

Fig. 4 shows the box plots of all the chosen values for the 1st JNDs and Break-in-presence points for all sequences, grouped by distortion type. The results are useful to define the ranges of visual distortions that are of interest for 360-degree visual quality subjective studies. Also, they help us get interesting insights on the consistency of the chosen levels among the users, between artifacts

[1]https://colour-blindness.com/colour-blindness-tests/ishihara-colour-test-plates/

(a) Cartoon **(b) Obama** **(c) Alcatraz** **(d) Aerial**

Figure 3: Equi-angular cubemap (EAC) versions of source sequences used in the experiment.

across the sequences, and expose trends among the mono/stereo sequences. Overall, the subjects had a better agreement on choosing the first JND while the larger whiskers bars indicate that for some of sequences/distortions they did not have a good agreement on choosing the Break-in-presence point. That is an expected result since presence might have a different meaning for different people. It is also interesting to notice that, despite some outliers, the level of agreement on the Break-in-presence for the seams artifacts is high. Finally, when comparing the mono/stereo versions of *Obama* and *Cartoon*, we can notice a strong trend towards the distortions being first perceived and causing a Break-in-presence earlier in the stereo versions.

Given the different nature of the artifacts, Fig. 4 does not allow us to directly compare the perceptibility across them. To understand how they compare to each other, Fig. 5 shows the average 1st JND vs. the average delta SSIM [8]. Delta SSIM is computed by subtracting the SSIM of the frame chosen as 1st JND from the SSIM of the Break-in-presence one. From Fig. 5, we can notice that: (1) the "blur" Break-in-presence point has lower SSIM, i.e., more distortion is needed before it is perceived; (2) the "blocking" Break-in-presence point has higher SSIM, meaning that it is more perceivable; and (3) being a global metric, SSIM does not discriminate seams well, highlighting the need for new perceptually-oriented objective metrics designed for 360-degree content. Such behavior is also expected for recently proposed metrics specifically developed for 360-degree content, e.g., S-PSNR and WS-PSNR [3], as they are global metrics and are unaware of specific 360-degree localized artifacts.

4 CONCLUSION

The exploratory study provided insights into the range of visual distortions to be used by subjective visual quality studies for 360-degree content. It also shows that the added distortions are less tolerable in stereo than in monoscopic content, and illustrates issues with current 360-degree content metrics, which do not discriminate clearly visible localized distortions, e.g., *seams*. Future work includes the extension to other projections/artifacts through artifacts-based quality experiments taking into account the lessons learned herein.

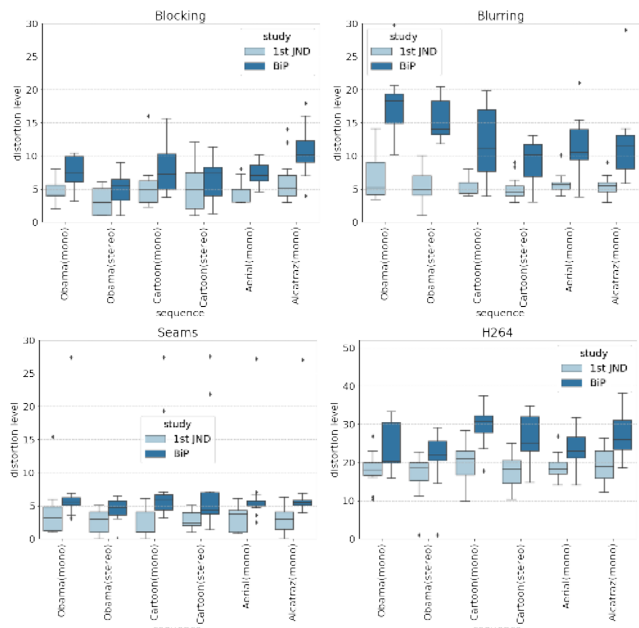

Figure 4: Boxplots of 1st JND and Break-in-presence points.

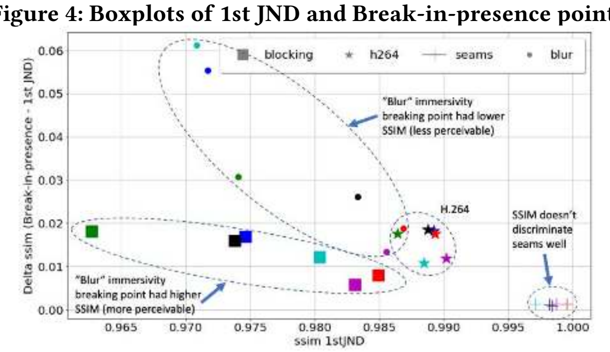

Figure 5: Average 1st JND vs. Average Delta SSIM.

REFERENCES

[1] R. G. A. Azevedo, N. Birkbeck, F. De Simone, I. Janatra, B. Adsumilli, and P. Frossard. 2018. Visual Distortions in 360-degree Videos. *arXiv preprint arXiv:1901.01848* (2018).

[2] C. Brown. 2017. Bringing pixels front and center in VR video. Retrieved March 19, 2019 from https://blog.google/products/google-ar-vr/bringing-pixels-front-and-center-vr-video/

[3] Z. Chen, Y. Li, and Y. Zhang. 2018. Recent advances in omnidirectional video coding for virtual reality: Projection and evaluation. *Signal Processing* 146 (2018).

[4] H. Duan, G. Zhai, X. Min, Y. Zhu, Y. Fang, and X. Yang. 2018. Perceptual Quality Assessment of Omnidirectional Images. In *2018 ISCAS*. 5.

[5] J. Jerald. 2015. *The VR book: Human-centered design for virtual reality.* Morgan & Claypool.

[6] A. Singla, S. Fremerey, W. Robitza, and A. Raake. 2017. Measuring and comparing QoE and simulator sickness of omnidirectional videos in different head mounted displays. In *9th QoMEX*. 1–6.

[7] E. Upenik, M. Řeřábek, and T. Ebrahimi. 2016. Testbed for subjective evaluation of omnidirectional visual content. In *2016 Picture Coding Symposium (PCS)*. 1–5.

[8] Z. Wang, , et al. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 4 (2004), 600–612.

[9] B. G. Witmer and M. J. Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240.

[10] M. Xu, C. Li, Y. Liu, X. Deng, and J. Lu. 2017. A subjective visual quality assessment method of panoramic videos. In *2017 ICME*. 517–522.

[11] X.K. Yang, W.S. Ling, Z.K. Lu, et al. 2005. Just noticeable distortion model and its applications in video coding. *Signal Processing: Image Communication* 20, 7 (2005), 662–680.