Evaluation of 360VR cinematic stereoscopic DASH service

Julián CABRERA, Marta ORDUNA, Lara MUÑOZ and Narciso GARCÍA.

Abstract— A DASH streaming service of 360VR stereoscopic video has been analyzed to evaluate the effect of the bitrate on the Quality of Experience perceived by the user, and the benefit of the use of advance delivery schemes for bandwidth savings.

Preliminary results show that user (un)stability and dizziness are related low bitrate representations, and that advance delivery schemes can produce significant bitrate savings while maintaining good QoE values.

I. INTRODUCTION

Virtual Reality (VR) videos and their latest versions in the form of 360° content (360VR) are becoming increasingly popular thanks to the emergence in the consumer market of 360VR displays. These displays are in the form of personal devices that connect to processing units such as, PCs or video game consoles, or as complements intended for the latest models of smartphones. In addition, the growing availability of both, cameras and tools to produce this kind of content, has produced a great demand for 360VR content [1].

In this sense, there is an intense research and development activity in both the Industry and the Academia in order to provide solutions to the different challenges that this new type of content brings [2] [3]: encoding and bitrate requirements for a sufficient quality representation, efficient delivery schemes, and display devices. In this paper, a 360VR cinematic stereoscopic streaming system that makes use of current solutions for the encoding, delivery and display stages is considered and analyzed. The analysis focus on two different aspects: (i) the effect of the encoding bitrate in the Quality of Experience (QoE) of the user, and (ii) the benefit in terms of bandwidth savings of the use of advance delivery schemes.

For that purpose, a top-bottom stereoscopic 360VR video has been used to feed the DASH-Server. The 4096x2048 video has been encoded producing H.264 streams contained in ISO-BMFF format at the following bitrates: 1, 2, 3, 5, 6, 8, 10, 15, 20 Mbps. These representations have been split into chunks of 2 seconds and the corresponding *Media Presentation Description* file has been created. This content can be viewed through any standard player which supports stereoscopic 360VR content and DASH.

II. 360VR STEREOSCOPIC DASH SYSTEM EVALUATION

A. Encoding quality assessment

As a preliminary analysis of the minimum requirements in terms of encoding quality, a set of quality assessment experiments have been carried out.

The same 360VR video has been showed several times to the users using only one representation at each time. The display device consisted of a VR glasses on a smartphone. After each viewing, users stated their level of satisfaction with the quality of the representation used.

The following conclusions have been extracted from the results:

- The lower bitrate representations, from 1Mbps to 3Mbps, hinder the user stability while watching the video and can cause dizziness.
- The minimum bitrate for quality acceptance is 5Mbps, although the subjective quality is still low.
- There is no significant difference between 15Mbps and 20Mbps, so there is an upper quality limit due to devices features.

B. DASH extension using Spatial Relationship Description

In the previous scenario, the complete 360° video content is continuously streamed to the client while only a portion of it is actually viewed by the user, depending on his/her point of view with respect to the scene. Therefore, part of the bandwidth used during the streaming session is wasted since it corresponds to the portion of the 360VR video that is not displayed. In order to be more efficient in the use of the network resources, the use of Spatial Relationship Description (SRD) [4] within the DASH specification has been considered.

SRD feature allows to signal spatial relationships among different representations within a DASH MPD. Therefore, the stereoscopic 360VR content can be divided into different spatial areas conforming different spatial representations of the content. In addition, it keeps the DASH paradigm where, in turn, each spatial representation may have different quality representations at different encoding bitrates. The purpose of this media description is to give the client a new degree of freedom (the spatial one) for selecting among several representations of the same spatial part of the full panorama scene. Thus, the client can select representations of different quality at the same time, for each area of the panorama content.

For that purpose and as a preliminary analysis, the whole 360VR sequence has been divided into the areas represented in Figure 1. It is important to note that in the equirectangular projection used, not all the areas of the frames have the same relevance when projecting the 360VR scene in the display. As an example, the upper and lower areas of the image (green and blue parts in Figure 1) suffer a display compression as they belong to the upper and lower part of the projection sphere. On the contrary, the central area of the 360 image (red region in Figure 1) projects on the central part of the projection sphere suffering less distortion.

Figure 1: Frame segmentation

Therefore, the 360VR scene has been divided spatially into 12 different areas (6 per view of the stereoscopic pair), and each one has been encoded at different bitrates. In order to estimate the bandwidth that can be saved if a client conforms the 360VR scene selecting a high quality representation for the area that corresponds to the current field of view (FOV) of the user, and lower quality representations for the rest of the scene, the following configuration has been analyzed:

- The reference quality has been set to that of the 10 Mbps representation of the complete 360VR sequence.
- Each spatial representation of the segmented sequence has been encoded at 2Mbps, 0.5Mbps, and 0.18Mbps. The 2Mbps representation offers a similar subjective quality to that of the full 360VR sequence encoded at 10Mbps (see Figure 2 and Figure 3).
- The FOV has been set to one of the central rectangles of the frame segmentation (red area in Figure 1)

Figure 2: Comparative between representations at different bitrates. (a) Reference representation: complete VR360 scene at 10Mbps, (b) Spatial representation at 2Mbps, (c) Spatial representation at 0.5Mbps, (d) Spatial representation at 0.1Mbps.

Two different combinations have been evaluated:

1) Scenario 1

The client selects for the FOV the representation at 2Mbps, thus keeping a similar subjective quality to that of the reference configuration. The rest of the central representations are requested at 0.5Mbps. The representations for the floor and ceil are requested at 0.18 Mbps.

Therefore, the total bitrate of the requested representations adds up to 7.72 Mbps, which corresponds to a bitrate saving of 22.8% in bitrate compared to the reference configuration while keeping a similar subjective quality.

1) Scenario 2

The client selects for the FOV the representation at 2Mbps,

thus keeping a similar subjective quality to that of the reference configuration. The rest of the representations for the floor and ceil are requested at 0.18 Mbps.

Therefore, the total bitrate of the requested representations adds up to 5.8 Mbps, which corresponds to a bitrate saving of 42% in bitrate compared to the reference configuration while keeping a similar subjective quality.

Figure 3: Comparative between representations at different bitrates. (a) Lower part detail of original (complete VR360 scene at 10Mbps), (b) Lower part representation at 0.18Mbps, (c) Upper part detail of original (complete VR360 scene at 10Mbps), (d) Lower part representation at 0.18Mbps.

III. CONCLUSIONS

Delivery of 360VR stereo video requires to adopt efficient solutions for encoding and representing the content. The consumption of this type of content imposes new constraints in terms of bitrate for an acceptable QoE mainly due to the nature of the display devices. In this sense, preliminary experiments have shown that effects such as user stability and dizziness are related to using low bitrate representations.

On the other hand, to reduce the bitrate required to provide a reasonable QoE, preliminary experiments have shown that the use of DASH Spatial Relationship Description can achieve a significant bitrate saving in delivering 360VR stereo video. Nevertheless, a deeper study need to be carried out to analyze the effect of this media description when changing dynamically the FOV of the user.

ACKNOWLEDGMENT

This work has been partially supported by the Ministerio de Economía y Competitividad of the Spanish Government under projects TEC2013-48453 (MR-UHDTV) and TEC2016-75981 (IVME) and the Spanish Agency CDTI under project IDI-20150950 (Tranvideoadap).

REFERENCES

- [1] L. Brun and A. Gasparini, "Enabling 360° visual communications: next-level applications and connections", *IEEE Consumer Electronics Magazine*, vol. 5, no. 2, pp 38-43, April 2016.
- [2] L. Turban, F. Urban, and P. Guillotel, "Extrafoveal Video Extension for an Immersive Viewing Experience", *IEEE Journal of TVCG*, Vol. XX, No. YY, Dec. 2015.
- [3] M. Budagavi, J. Furton, G. Jin, A. Saxena, J. Wilkinson, and A. Dickerson, "360 Degrees Video Coding Using Region Adaptive
- [4] M. Oskar van Deventer, E. Thomas, F. Mazé, and F. Denoual, "White paper on Spatial Relationship Description in MPEG DASH part 1 AMD2", "TNO, Canon Research Centre France" France, Oct. 2015.