Tile-based high resolution VR contents delivery

Mickaël Raulet, Eric Toullec, Lucas Grégory
R&D Research Department
ATEME
Rennes, France

Jean Le Feuvre
Multimedia Lab
TELECOM ParisTech
Paris, France

***Abstract*— This demonstration proposes a distribution solution for High Resolution VR Contents, based on the emphasis of the current user's field of view over the rest of the video. The demonstration shows a complete work flow based on the use of HEVC Tiles and MPEG-DASH distribution.**

Keywords— High Quality VR Contents – MPEG DASH – Tiles - HEVC – 360° Video

I. Introduction

VR content production and consumption have been expanding a lot the last past years, in particular thanks to the growing market of Head Mounted Displays and low budget 360° cameras. Besides the broadly developed entertainment application, VR offers a lot of possibilities in other domains such as Education or Visio conferencing.

Yet one of the biggest challenges surrounding the growth of VR contents consumption is to find a distribution solution fitting the current bandwidth limitations while keeping visual quality a priority, since the purpose of VR is to provide the most immersive experience. Furthermore, even if a 4K contents distribution is conceivable, most of the currently available end devices would not be able to download and decode the resulting streams.

The aim of this demonstration is to provide a solution for High Resolution VR contents delivery, exploiting a live tile-based HEVC encoding of the content, and the adaptive streaming standard MPEG-DASH.

II. Demonstration Overview

When watching VR content (with an HMD or any other device), at any time, only the portion of the original full sphere corresponding to the current FoV (field of view) is displayed. This means that the remaining parts of the sphere are uselessly transmitted and decoded.

Setting out from this observation, it's obvious that transmitting only the useful part of the sphere would allow a significant gain in terms of bandwidth usage. The idea of this demonstration is to spatially segment the original video stream, corresponding to the full sphere, on the HEVC tiling structure basis. Then, only the needed tiles are transmitted depending on the FoV of the user (see Fig. 1). In order to dynamically adapt the displayed video to the user's FoV, every resulting single-tile video stream must be available at any time.

At this point, only a part of the sphere is transmitted and displayed, which might be a problem if the FoV of the user suddenly changes. To avoid this inconvenience, the solution implemented for this demonstration is to permanently download and display a sub sampled version of the full video sphere, in addition to the high-resolution tiles, as illustrated in Fig. 1. This version can be encoded at a very less significant bitrate, since it is not visible most of the time.

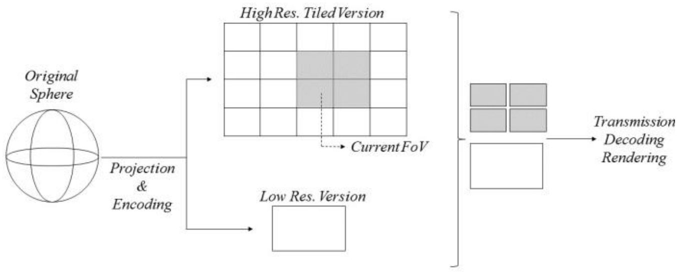

Fig. 1 : Video streams involved for displaying a given field of view.

In order to allow an independent decoding and rendering of the extracted tiles, the High Resolution version must be encoded with the 'motion restriction' option of ATEME's Titan Live encoder, constraining the motion prediction vectors inside the tiles delimitations.

The distribution part is performed using the Adaptive Streaming Standard MPEG-DASH. Besides the temporal segmentation it performs (which is useful to dynamically adapt to the user's viewport), DASH provides a Spatial Relationship Description metadata management, used to inform the DASH Client about the spatial relations (i.e. relative positioning) between the extracted tiles, thereby allowing fast selection of the tiles by the DASH engine for further re-aggregation by the video renderer.

The client side is handled by MP4Client, a DASH client developed by the GPAC team of Telecom ParisTech and using a full hardware based decoding platform in this context. The client has the possibility to do a VR rendering of a 360 video, analyzing the current viewport to select the appropriate tiles in the video (see fig 2).

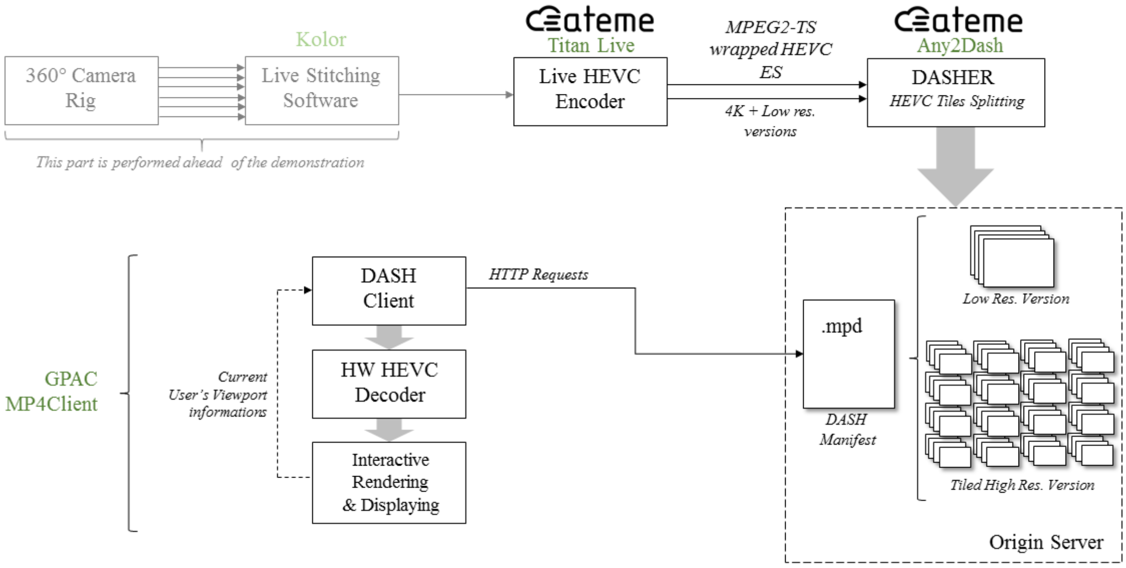

Fig. 2 : Workflow proposed by the demonstration.

Involved equipment :

- **Encoder** : Titan Live - x86 CPU 36 cores– 64 GB RAM ;
- **Dasher** : Any2Dash – x86 CPU ;
- **Server** : WebDAV enabled Apache2 server ;
- **DASH Client** : GPAC MP4Client - x86 CPU, *nVidia 1080 Ti* GPU for hardware HEVC decoding and rendering ;
- **Restitution** : UHD Display.

III. Tests and Results

To measure the gain provided by this demonstration, in terms of reduction of the needed bandwidth, we used two versions of a 360° video, encoded at a 4K resolution and a bit rate of 25 Mbps, uniformly cut in 3x3 and 6x6 tiles respectively. In addition to these streams, a HD version is also encoded at a constant bit rate of 5 Mbps.

Assuming the bit rate is spatially evenly distributed across the picture (see Tab. 1), the resulting gain is measurable by monitoring the liaison between the Origin Server and the DASH Client.

TABLE 1 : Average bit rate per tile.

Sequence	**Tiles structure**	**Bit rate (Mbps)**	
		Total	***Average / tile***
RG2017_3x3	3x3	25	2.78
RG2017_6x6	6x6	25	0.69

In the first case, with the 3x3 tiles version, we observed that the number of tiles needed to cover the FoV of the user fluctuates between 4 and 6, giving a maximum needed bit rate of 16.7 Mbps.

Adding the extra 5 Mbps needed to transmit the sub sampled full sphere, it gives a total required bitrate of 21.7 Mbps. In comparison with the original bit rate of 25 Mbps, it represent a 13.3% gain.

With the second version, 6 to 12 out of 36 tiles are needed to cover the FoV, resulting in a maximum needed bit rate of 8.3 Mbps to transmit the tiles. Once again, adding the extra 5 Mbps corresponding to the low resolution version, it makes a total of 13.3 Mbps. Thus, the solution provides a gain of 46.6% in this case (see fig. 3).

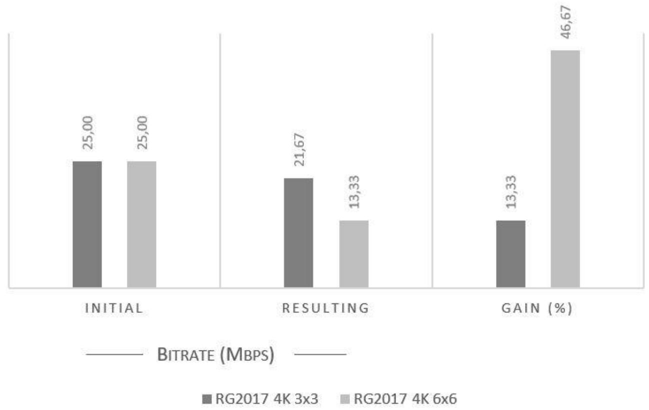

Fig. 3 : Gains provided by the demonstration.

IV. Conclusion

We have shown that it is possible to significantly reduce the required bitrate for a high quality 4K VR experience by up to almost 50%, enabling most devices to be eligible for such services. Our future work will focus on super-high resolution VR 12K and more.

References

[1] K. Misra, A. Segall, M. Horowitz, S. Xu, A. Fuldseth, M. Zhou, "An Overview of Tiles in HEVC", *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 969-977, Dec. 2013.
[2] C. Concolato, J. Le Feuvre, F. Denoual, F. Mazé, E. Nassor, N. Ouedraogo, J. Taquet,, "Adaptive Streaming of HEVC Tiled Videos using MPEG-DASH", *IEEE Transactions on Circuits and Systems for Video Technology*, March 2017.
[3] Y. Wang, T. Stockhammer, '[OMAF] On region-wise packing and its support in the viewport independent HEVC media profile', July 2017.
[4] live360tv.fr, 'Le projet collaboratif' [Online] : http://www.live360tv.fr/projet.php