Culture Assignment: View Direction Adaptive Video for VR Optimization

Asa Hayes – CSCE-221-506

January 24, 2018

While virtual reality is one of the fastest growing parts of the tech industry right now, it is currently trying to overcome a large hurdle: hardware. While the amount of VR applications being made is growing at a high rate due to the possible applications, use and adoption is lagging heavily behind because of the resources required to do it. As far as consumer use, one has to have essentially the top of line for every involved component. Hardware, and specifically graphics hardware is trying to move along to match due to demand, but solutions are still needed for the interim. One of the current paths to optimization is selective compression of the video streaming component, which is the approach taken from the article covered.

The basic premise of the approach is to have the total 360° stream mapped onto a sphere like regular VR video format, but have it divided into a grid of smaller individual videos comprised of cropped parts of the HD source. The optimization comes from the division of these sub-videos, as they can be dynamically down-sampled in quality to reduce the necessary workload for streaming. To determine which of these subsections should stay at source quality and which can be put on the back-burner, the positional sensors are used to detect where on the sphere the user is looking. The algorithm used for detecting where the video is projected and for how high quality takes in the current yaw (horizontal orientation), pitch (vertical orientation) and roll (rotational orientation) of the

VR headset as a composite vector. From this vector, a projection matrix and a rotation are created for the size and position of the priority field of view, namely the area the user is looking at. The projection matrix creates the total area to be shown on the sphere/grid (as a 2:1 rectangle) and the rotation matrix moves and centers that area to where the user is looking. From this new projected area, a simple three-part decision making process is used to determine if a piece of the grid is in the user's field of view and should be streamed at a higher quality: If the area contains a) one vertex of a section b) the center-point of a section or c) the entire section, said section is to be kept in high definition. All other sections are to be streamed in gradually lower quality as their distance from the main area increases, greatly lessening the resources needed for streaming all of the sections. Additional minor factors are the amount of sections, where higher amounts give marginally better bitrates up to a near negligible degree, and the available bandwidth, which would only factor in on the allocated quality when using a currently uncommon wireless VR display.

The societal and economic impact of this implementation would not be immediate if implemented soon as we are currently in a transition period in regards to cost of both graphics and VR hardware, but once the next generations of both come out and the current hardware is more available, this technique will serve to extend the range of people that are able to access and develop for VR systems. Having the technical requirements be lowered from the dynamic down-sampling means that not only will the hardware that can run VR content become more and more attainable, the range of hardware that will support high-range VR applications would increase as well. Despite the current trend of having most VR software as video games right now, if VR is more accessible to others as a utility rather than just a toy for those with disposable income, the demand and utility of non-game VR software would most likely increase along with adoption.

An increase of VR use like this would most likely trigger a big change in tech culture, similar to the boom in mobile devices that started around a decade ago.

References

[1] Kim, S. , Yun, J. , Jo, B. , Kim, J. , Chae, H. and Suh, D. (Jan. 2018) View Direction Adaptive 360 Degree Video Streaming System Based on Projected Area. Journal of Computer and Communications, Vol. 6 No. 1, p. 203-212. doi: 10.4236/jcc.2018.61020.