Poster: Mobile Virtual Reality for Head-mounted Displays With Interactive Streaming Video and Likelihood-based Foveation

Eduardo Cuervo
Microsoft Research

David Chu
Microsoft Research

Immersive virtual reality has long been an aspiration of many. For a truly immersive VR experience, three properties are essential: *quality*, *responsiveness* and *mobility*. By quality, we mean that realistic and life-like visual portrayals in a virtual environment heighten our sense of immersion. By responsiveness, we mean that any motion, especially of the head, must be reflected as quickly as possible in visual feedback because ocular proprioception sensitivity is very high. By mobility, we mean that we ought to be able to move untethered in physical space, free to explore our virtual world. Unfortunately mobile devices like phones are tablets are power-constrained and cannot produce acceptable framerate. While cloud offloading allows them to deliver high-quality and framerate it places high bandwidth requirements and introduces an unacceptable amount of latency.

This demo introduces *Matia*, a stereo HMD system that simultaneously attains quality, responsiveness and mobility. Matia achieves this by offloading rendering work to a high-end GPU across the WAN. The HMD client is mobile and only requires a low-end GPU, yet receives high quality imagery. To overcome WAN latencies, Matia borrows speculative execution techniques inspired by previous work *Outatime* [1]. However, speculative execution alone is insufficient. This is because HMD sensitivities are substantially more stringent along several key dimensions: responsiveness and quality.These factors suggest that we ought to greatly benefit from maximally utilizing all available bandwidth between client and server to deliver the highest resolution imagery possible. At the same time, any latency or bandwidth changes must be handled very responsively.
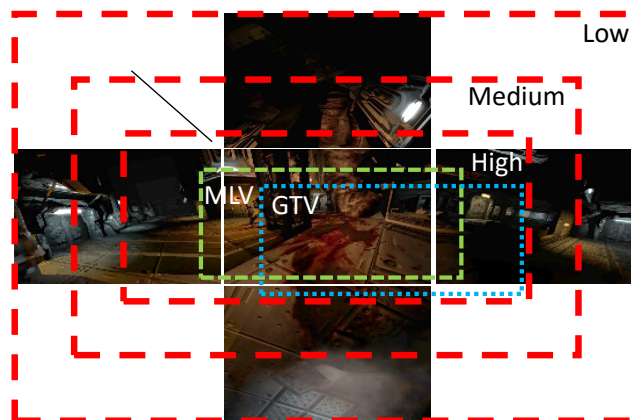
Figure 1: A panorama with foveation regions

To get the highest possible resolution while remaining responsive to network changes, Matia employs *panoramic stereo video* and *likelihood-based foveation*. Specifically, Matia renders a wide field-of-view (FOV) panoramic stereo video. In turn, any possible stereo view (e.g., due to unexpected head movement or network fluctuations) can be generated. Matia then foveates this panorama by reallocating pixels to areas where the user is most likely to look (Maximum Likelyhood View). We divide our panoramic video into three regions described by their image quality: High, Medium and Low. Embedded within is a convex-like optimizer that constantly adapts to real-time data analysis of user head movement and network conditions to figure out the most fruitful foveation configuration, on expectation of these regions.

As a result, we expect to achieve high quality immersiveness in the common case, yet still deliver gracefully degraded experiences in case of network degradation. At the same time, Matia greatly expands the scope of viable HMD content while requiring only modest fixed function computation on the HMD device.

1. REFERENCES

[1] K. Lee, e. a. Outatime: Using speculation to enable low-latency continuous interaction for cloud gaming. MobiSys 2015.