360° Vision Using FOV Minification

What's it like to have eyes in the back of your head?

NEIL SORENS, University of Washington

Fig. 1. A Valve Index HMD connected to a PC displayed an FOV-compressed view of the live video provided by the Ricoh Theta V 360° camera. The camera was connected to the PC via the Index's "frunk."

Having "eyes in the back of your head" in real time is now possible thanks to HMDs and video stitched together from multiple cameras. By compressing the resulting 360° video into a field of view small enough that it fits into the center of human vision, we can see all around us without turning our heads or even moving our eyes. The trade-offs are many: dramatically reduced detail, increased latency, poor spatial judgment, and, generally, a large and potentially nauseating departure from normal perception. These trade-offs have been studied using the early wave of VR HMDs. The introduction of vastly improved commercial cameras and HMDs with high FOVs (and thus less potential compression) and higher resolutions (and thus less loss of detail) provides an opportunity for someone with no hardware or software knowledge to experience the impact of these trade-offs for themselves.

1 INTRODUCTION

The augmentation of human abilities, including vision, with the use of electronic devices has long been a staple of science fiction, but it is only in the past decade or so that the technology and computing power have come far enough to start turning fantasy into reality.

Author's address: Neil Sorens, nsorens9@cs.washington.edu, University of Washington.

There have been numerous studies on the subject of enhanced vision - specifically, view expansion - using HMDs, including some that involve FOV minification.

Although most of those studies involved a virtual environment, this project does not break any new ground. Rather, it showcases the ability to experience and tinker with FOV minification without the construction or modification of any hardware and using only the ubiquitous Unity engine.

The expectation was that with improved hardware and easy-to-use development software, the experience would not only be noticeably superior to previous efforts, but it would be accessible to anyone regardless of their level of technical knowledge.

2 RELATED WORK

Although it was not inspired by it, the project is essentially the same concept as FlyVIZ 2.0 [Ardouin et al. 2016], except tethered and with improved camera, HMD, and processor.

The original FlyVIZ [Ardouin et al, 2012] was a cleverly hacked-together assembly consisting of a SONY HMZ-T1 HMD, a catadioptric sensor, and a laptop carried in a backpack. While it was functional, it was limited by the HMZ-T1's now-atrocious 45° FOV. Its 1280x720 per-eye resolution was acceptable, and its 16:9 per-eye aspect ratio are actually quite good in this role. But the custom nature of the build - in particular, the sensor used - meant that few people, if any, built one for themselves.

The 2.0 version used an Oculus DK1, an iPhone 4S, and a GoPano lens attached to the iPhone. This setup more than doubled the display FOV of the previous effort and was easier to replicate, although the lens was an old KickStarter-based product with limited availability. And the per-eye resolution of 640 x 800 was a step down, since much of the vertical resolution was lost to letterboxing (an issue on this project, as well).

Both of these efforts were designed more around portability than visual fidelity and were positioned as a device that could be refined for real-world use. This project takes the opposite tack, using a tethered headset and making no pretenses to being a prototype of anything practical.

3 METHOD

The initial challenge in creating an easy-to-build setup is finding a compatible set of HMD, camera, and engine, with the camera being of particular importance.

The 360° camera has a lens on each side of the camera, and each lens captures video with a 180° FOV. Video stitching combines the video from each lens into a single 360° video in a spherical format. This format is then projected into an equirectangular format, much as a spherical globe of the Earth is projected to a 2D map. While this format distorts the top and bottom (Greenland actually isn't that huge), humans can generally adapt to this kind of distortion fairly well, especially since we are used to watching rectangular-shaped screens. Other types of projection (and there are too many to list) would leave empty spaces where we expect to see things, and we do not adapt well to this.

Even with the camera handling all the nasty work of stitching and projection, we have to handle some work in the engine. In order to perceive detail across the full 360° of video, the video cannot be too close to the edges of the display, as this area is slightly blurry on the Index. Pillarboxing achieves this effect. Furthermore, on a display that is roughly square per eye (1440×1600), it is necessary to letterbox source video with a aspect ratio of 1.67:1, as stretching the video vertically exacerbates the distortion that we already have as a result of equirectangular projection. Of course, the tradeoff of letterboxing and pillarboxing is that all that nice display space on the HMD is going to waste, and some detail is lost as the video is compressed into a narrower space / fewer pixels.

4 IMPLEMENTATION DETAILS

The system consists of four major pieces: the Valve Index HMD, the Ricoh Theta V 360° camera, a PC to which to connect the hardware, and Unity with the SteamVR plugin.

The Valve Index was selected for its top-of-the-line resolution and FOV and the convenience of its "frunk" USB port for camera connection. Its major downside was that it is tethered to the PC and external power. However, since I was primarily concerned with the visual experience rather than mobility, this was an acceptable disadvantage. Another minor disadvantage was the lack of an obvious way to mount the camera on the headset. A gooseneck clip-on tripod proved to be the most reliable and stable way of attaching the camera, though it adds unwanted distance between the camera lenses and the location of the eyes.

The Ricoh Theta V 360° camera was selected because of its apparent compatibility with Unity, its ability to output real-time high-resolution video at an acceptable frame rate, its auto-stitching feature, and its output in equirectangular format. An under-the-radar driver update does fix its incompatibility with Unity on Windows, and it always worked on Mac.

Unity was selected for its wide adoption in development on VR hardware and its reputation for being simple to use. Once the incompatibility with the camera was resolved, I learned enough about how to use Unity, and I installed the SteamVR plugin, it was a simple matter to stream the video onto a canvas and to implement the desired amount of letterboxing and pillarboxing by tweaking values in the Inspector window to match personal preferences.

I used a high-end desktop PC to eliminate PC hardware as a potential source of extra latency. It has an i8700k CPU, a GeForce GTX 1080 CPU, and 32GB of RAM. It is likely that substantially less powerful hardware will see identical performance, however, as the processors were never heavily taxed.

5 EVALUATION OF RESULTS

The expectation that this setup would provide a big leap over the FlyVIZ 2.0 did not prove to be the case. Latency is a harsh mistress. Even with high-end PC hardware, the delay between events happening and being displayed on the HMD (approx 300ms) was too great for the user to succeed at fast-response tasks, such as catching or dodging a thrown object. Furthermore, the additional per-eye FOV of the HMD added no benefit, as the aforementioned pillarboxing and letterboxing nullified this advantage. Distance judgment, as has been noted in the past [Zhang et al. 2012], is difficult, and this did not improve.

On the other hand, the plug-and-play nature of the setup and easy tweak-ability do allow anyone with compatible hardware to easily try the experience themselves. And the improved pixel density of the HMD displays and camera sensor make it possible to distinguish finer detail.

6 FUTURE WORK

A stereoscopic 360 camera could help with the issue of depth perception created partially by the monoscopic camera.

Reducing latency would vastly improve the experience. Perhaps performing the stitching and projection on high-powered hardware instead of on the camera would reduce latency (although the camera's contribution to latency has not been measured yet). In addition, although Unity is easy to use, it is not known for its speed and efficiency. A custom engine could likely provide some benefit.

And we are not truly a cyborg ninja without other vision enhancements. The ability to manually control the view (for example, zoom) seems like low-hanging fruit. Even better, the view could auto-center on motion, giving our cyborg ninja sniper instant notice when a target appears. Shaders like a Sobel filter could simplify the view, which contains so much information that it is hard to process even with practice.

Tthe addition of more sensors, like an infrared camera, could add important extra information to the video feed.

Finally, the removal of the tether would make this a more practical device.

REFERENCES

ARDOUIN, J., LÉCUYER, A., MARCHAL, M., RIANT, C., MARCHAND, E.: FlyVIZ, a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology* (2012), AMD, pp. 41—44.

ARDOUIN, J., LÉCUYER, A., MARCHAL, M., GUILLERMO, A., NOUVIALE, F., MARCHAND, E.: Enjoy 360° vision with the FlyVIZ. In *SIGGRAPH '16: ACM SIGGRAPH 2016 Emerging Technologies* (July 2016), Article No.: 6.

ZHANG, R., NORDMAN, A., WALKER, J., KUHL, S. A.: Minification affects verbal and action-based distance judgments differently in head-mounted displays. *ACM Transaction on Applied Perception 9*, 3.