Gigapixel Virtual Reality Employing Live Superzoom Cameras

O. Koskinen
University of Tampere
Faculty of Communication Sciences
33014 Tampere, Finland
Olli.Koskinen@staff.uta.fi

I. Rakkolainen
University of Tampere
Faculty of Communication Sciences
33014 Tampere, Finland
csisra@sis.uta.fi

R. Raisamo
University of Tampere
Faculty of Communication Sciences
33014 Tampere, Finland
csrora@sis.uta.fi

ABSTRACT

We present a live gigapixel virtual reality system employing a 360° camera, a superzoom camera with a pan-tilt robotic head, and a head-mounted display (HMD). The system is capable of showing on-demand gigapixel-level subregions of 360° videos. Similar systems could be used to have live feed for foveated rendering HMDs.

CCS CONCEPTS

• **Human-centered computing** → *Displays and imagers* • **Computing methodologies** → *Mixed / augmented reality* • *Virtual reality* • **Hardware** → *Displays and imagers*

KEYWORDS

Head-mounted display; gigapixel; virtual reality; 360° video; superzoom; foveated rendering

ACM Reference format:

O. Koskinen, I. Rakkolainen, and R. Raisamo. 2018. Gigapixel Virtual Reality Employing Live Superzoom Cameras. In *Proceedings of ACM VRST conference, Tokyo, Japan, November 2018 (VRST'18),* 2 pages.
DOI: 10.1145/3281505.3281586

1 INTRODUCTION

Presumably the first immersive technology was painted panoramas in dedicated cylindrical buildings [1]. Nowadays head-mounted displays (HMD) are often used to show virtual reality (VR), augmented reality (AR) and spherical 360° video. VR and 360° video has recently become affordable for consumers.

Gigapixel images[1] or videos have a resolution of billions of pixels, and they enable zooming far into the details of an image. They are suitable for relatively stationary environments such as cityscapes. They demand vast resources of processing, memory, transmission and storage. Viewing omnidirectional gigapixel images with an HMD poses several additional problems, including latency and user interfaces for zooming and viewing.

Pan-tilt-zoom (PTZ) cameras are used in applications such as teleconferences, surveillance, sports broadcasts, etc. Consumer cameras can embody compact, even 125x superzoom lenses, which can show otherwise imperceptible details of distant objects.

Our contribution is a low-cost, interactive, deeply zoomable 360° video proof-of-concept system. A viewer gets an overview from a 360° camera and can get more details of the environment with a PTZ camera as desired. This saves a lot of computing and networking resources, but enables only one person at a time to control the PTZ camera. As far as we know, no other systems combine live 360° and superzoom PTZ cameras for HMD viewers, thus appearing as live, zoomable gigapixel-resolution VR.

We first describe related PTZ and 360° cameras and gigapixel VR with HMDs, then present our prototype, discuss limitations and potential improvements of it, and finally give conclusions.

2 RELATED WORK

Philco Headsight [4] was the precursor to an HMD and to a robotic telepresence camera. It showed remote places by panning an HMD and a remote camera in synch. Scotti et al. [9] used a dual camera for high-definition surveillance. The robotic PTZ cameras Pixio[2] and Soloshot[3] automatically follow an object wearing an RFID or a GPS tag. IC Realtime IC720 virtual PTZ camera[4] uses a high-resolution 360° camera with digital zoom and HMD.

AR binoculars [7] augment high-precision synthetic objects of live long range telescopic imagery. Clairity Smartbinocular[5] has a 30x optical zoom PTZ camera controlled with hand-held HMD movements. It overlays AR information to airplanes on air traffic controller's vision and thus replaces regular binoculars.

Gigapixel images can be captured with pan-tilt robotic heads[6] with low magnification tele lens [5], and the capture can take a long time, or with synchronized multicamera arrays, which require massive parallel cameras and processing [2]. Partial video clips and looping videos have been added to gigapixel panoramas.

Gigapixel images and videos can be used also for VR. Only a fraction of the 360° VR or gigapixel imagery can be viewed at any one time, and usually viewers are interested only in some regions, e.g., tourist hot spots in a city scene. Much of the view is mostly irrelevant (e.g., blue sky) or static (e.g., distant city scenes). Recent VR streaming systems send only the current viewport and may also predict viewer's head movements.

ACM ISBN 978-1-4503-6086-9/18/11.
https://doi.org/10.1145/3281505.3281586

[1] http://360gigapixels.com/ (All footnotes accessed on Sept. 29th, 2018)
[2] https://shop.movensee.com/en/
[3] https://soloshot.com/
[4] https://www.asmag.com/showpost/19853.aspx
[5] http://360.world/clairity/
[6] http://gigapan.com/

3 APPARATUS

Our goal is to provide a live 360° overview, and to enable an interactive, deep gigapixel zoom for one person at a time. Our approach is to use a 360° camera and a PTZ camera, which sends only the small zoomed subset of the imagery. The user can look around the overall live 360° video view, and also take a closer look anywhere with the superzoom / gigapixel functionality.

The camera hardware consists of a 4K resolution 360° camera (Ricoh Theta V[7]) and a 36x optical superzoom PTZ IP camera (Imporx IMP-4MPH9801Q-MINI) placed on a tripod. Ricoh Theta V can stream the video in real time without noticeable delay. It is above the PTZ camera (see Fig. 1 left). Our PTZ camera can pan 360° and tilt 90° down and 5° up, which limits the PTZ views to the sky, and our construction also slightly limits the PTZ horizontal views. PTZ rotation speed is 180°/s horizontal and 45°/s vertical.

Figure 1: Left: a stationary 360° camera and a superzoom PTZ surveillance IP camera on a tripod.

Right: The zoom image can be shown in a window (red, with guiding yellow rectangle to show the overall context as depicted here), or it can also replace the overall view fully.

The 360° video was brought to Unity 2018.2.f1 as a sphere. The PTZ camera streamed video on RTSP format, and Unity could not directly import it. We made custom C# software to bring it all together. We used an MSI GS63VR 7RF-218NE laptop (with Windows 10, 2.8 GHz Intel Core i7-7700HQ, 16 GB RAM, and Nvidia GTX 1060) and an HTC Vive[8] for viewing. The HMD orientation controlled camera pan and tilt over IP and HTTP.

The weights of the Ricoh Theta V and the PTZ camera are 121g and 2230 g, respectively. The PTZ camera is rather heavy, but many similar PTZ web conferencing cameras are much lighter.

4 RESULTS AND DISCUSSION

The user can look around the overall live 360° video view, and also take a closer look anywhere with the live superzoom/gigapixel functionality. The superzoom view can cover a part of the HMD field-of-view (FOV) to retain context (see Fig. 1 right), or it can also replace all of it to better view the details.

Gigapixel VR images raise the issue on how to conveniently control the zooming while using an HMD. The deeper the user zooms into the live image, the smaller the FOV of the lens becomes. Head movements would cause a lot of image shaking and motion sickness. Orlosky et al. [6] presented a gaze-based user interface for a video see-through HMD, which uses various focal length lenses. Chang and Cohen [3] presented improved user interfaces for zooming into gigapixel VR images.

The effect of head's pan-tilt movements for PTZ camera and viewing should usually be dissipated and eventually stopped while zooming deeper into the image. Zooming out would then again increasingly release the freedom of pan-tilt movement. Very deep zooming may result in shaky image also due to vibrations of the PTZ camera or may have atmospheric blur.

The system can provide extreme zooming into the image, but is limited on the number of simultaneous users. It brings relatively little extra on cost, computation, storage and network transfer. The camera setup can be local or remote, and stationary or mobile.

Live gigapixel VR could also be used to enhance foveated rendering [8] for HMDs, which use a standard display for overall view and a moving, extremely high resolution display at the direction of the gaze, (e.g., Varjo HMD[9]). The tiny gaze-directed display has a very small FOV, and as the gaze and head can move quickly, constructing such HMDs is a challenge. Such foveated rendering HMDs can benefit from our dual-camera system by providing high-resolution video content for them.

5 CONCLUSIONS

We presented a feasible way for an HMD to present live overall view with a 360° camera, and use a PTZ superzoom camera to view details on the regions of interest. The proof-of-concept prototype can capture even tera- or petapixel-level images. It is promising for gigapixel VR with HMDs and for foveated rendering HMDs. Such systems have numerous consumer and professional applications. We will continue to develop the hardware and software to test related issues and user interfaces.

ACKNOWLEDGMENTS

We thank our colleagues and reviewers for helpful comments on this paper. This work is partially funded by Tekes (#1444/31/2016).

REFERENCES

[1] R. Barker. 1796. La Nature à Coup d' Oeil. In *Repertory of arts and manufactures.* George and Thomas Wilkie and G.G. and J. Robinson, London, UK.
[2] D. Brady, W. Pang, H. Li, Z. Ma, Y. Tao, and X. Cao. 2018. Parallel cameras. *Optica* 5, 127-137.
[3] H. Chang and M. Cohen. 2017. Panning and Zooming High-Resolution Panoramas in Virtual Reality Devices. In *Proc. of the 30th ACM Symp. on User Interface Software and Technology* (UIST'17). ACM, New York, NY, 279-288.
[4] C. Comeau and J. Bryan. 1961. Headsight television system provides remote surveillance. *Electronics*, 10, 86–90.
[5] J. Kopf, M. Uyttendaele, O. Deussen, and M. Cohen. 2007. Capturing and viewing gigapixel images. *ACM Trans. Graph.* 26, 3, Article 93.
[6] J. Orlosky, T. Toyama, K. Kiyokawa, and D. Sonntag. 2015. ModulAR: Eye-Controlled Vision Augmentations for Head Mounted Displays. In *IEEE Trans. on Visualization and Comp. Graphics*, 21, 11, 1259-1268.
[7] T. Oskiper, M. Sizintsev, V. Branzoi, S. Samarasekera, and R. Kumar. 2013. Augmented Reality binoculars. In *Proc. 2013 IEEE Int. Symp. on Mixed and Augmented Reality* (ISMAR'13). IEEE, New York, NY, USA, 219-228.
[8] A. Patney, M. Salvi, J. Kim, A. Kaplanyan, C. Wyman, N. Benty, D. Luebke, and A. Lefohn. 2016. Towards foveated rendering for gaze-tracked virtual reality. *ACM Trans. Graph.* 35, 6, Article 179.
[9] G. Scotti, L. Marcenaro, C. Coelho, F. Selvaggi, and C. Regazzoni. 2005. Dual camera intelligent sensor for high definition 360 degrees surveillance. In *IEE Proceedings - Vision, Image and Signal Processing*, vol. 152, no. 2, 250-257.

[7] https://theta360.com/en/about/theta/v.html
[8] https://www.vive.com/us/product/vive-virtual-reality-system/
[9] https://varjo.com/