Towards Effective Interaction with Omnidirectional Videos Using Immersive Virtual Reality Headsets

Benjamin Petry[1], Jochen Huber[1,2]
[1] Singapore University of Technology and Design, Singapore. bpetry@acm.org
[2] MIT Media Lab, Cambridge, MA USA. jhuber@mit.edu

a) Oculus Rift and mounted Leap Motion **b) User, interacting with the system** **c) The user's point of view**

Figure 1. Overview of the ODV browsing system

ABSTRACT
Omnidirectional videos (ODV), also known as panoramic videos, are an emerging, new kind of media. ODVs are typically recorded with cameras that cover up to 360° of the recorded scene. Due to the limitation of the human vision, ODVs cannot be viewed as-is. There is a larger body of work that focuses on browsing ODVs on ordinary 2D displays, e.g. on an LCD using a desktop computer or on a smartphone. In this demonstration paper, we present a new approach for ODV browsing using an immersive, head-mounted system. The novelty of our implementation lies in decoupling navigation in time from navigation in space: navigation in time is mapped to gesture-based interactions and navigation in space is mapped to head movements. We argue that this enables more natural ways of interacting with ODVs.

Author Keywords
Omnidirectional, panoramic, interactive video, ODV, mobile, mid-air, gesture-based interaction, body-based

AH '15, Mar 09-11, 2015, Singapore, Singapore
ACM 978-1-4503-3349-8/15/03.
http://dx.doi.org/10.1145/2735711.2735785

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Omnidirectional videos have lately received a lot of attention due to emerging technology such as the GoPano lens for the iPhone or standalone recording devices like the 360cam [6]. Being composed of moving images that capture up to 360° of the recorded scene, these videos cannot be viewed as-is due to the limited human field of view. Thus, they pose a unique challenge to the end-user: they offer two degrees of freedom for navigation, users can navigate both in space and time to navigate an ODV.

Users usually focus their attention on a particular region and relocate this very region while watching. However, this oftentimes leads to missing relevant parts of the video that happened in another location at a different time. The central question for our line of research is how to design effective spatial and temporal navigation support?

ODV playback has been implemented using different technology, such as CAVEs [3], head-mounted displays [2], spherical displays [1] and large screens or tablets [5]. CAVEs and head-mounted displays are immersive and offer the advantage of natural spatial navigation through head-movement. Non-immersive displays, however, require other interactions for spatial navigation due to their form factor, such as mouse input. Little research has been done on interacting with ODV content. Ruiz et al. explored ODV interaction in a CAVE-like setup using mid-air gestures [3].

In this paper, we present a novel system for interacting with ODVs as an amalgam of an immersive virtual reality system (with head-tracking capabilities) and a mounted gesture recognition system. This head-worn system provides a high degree of freedom for interaction (see Figure 1b): spatial navigation is mapped to head-based interaction and navigation in time to gesture-based interaction. This enables users to perform both navigation tasks independently and in particular concurrently.

CONCEPT AND IMPLEMENTATION

In the following, we describe the main interaction concept and detail both hardware and software implementation.

Interaction Concept

Previous work exclusively presented systems that allow either temporal or spatial navigation at a time. However, we argue that navigating in both space and time independently is key to avoid missing certain events and therefore to browse an ODV effectively. Conceptually, we map spatial and temporal navigation to *different* input methods that can be easily carried out concurrently. As a concrete example, we chose head-based and mid-air interactions. Spatial navigation, i.e. panning, is mapped to head rotation (see Figure 2a). Temporal navigation is facilitated through mid-air gestures (see Figure 2b). The concrete gesture language is based on [3]: play is performed using a push gesture, pause/stop by halting the arm stretched and fast forward/rewind by moving the hand to the right/left.

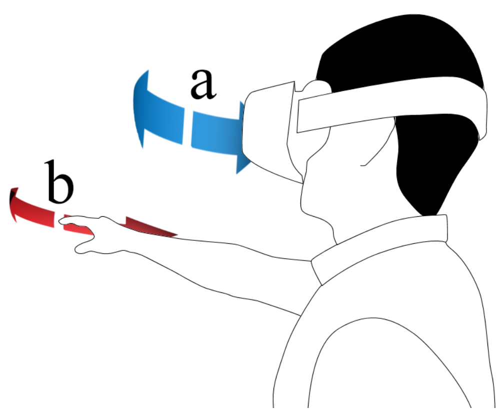

Figure 2. Decoupled spatial (a) and temporal (b) interaction.

Hardware

Our system uses the Oculus Rift DK2 [7] as an immersive virtual reality headset. It also leverages on head-rotation information provided by the Oculus. Mid-air gestures are tracked using a Leap Motion [8] controller, mounted on the Oculus Rift (see Figure 1a). Both are tethered to a laptop.

Software

The system is implemented using openFrameworks [9] and leverages on the Oculus Rift SDK, as well the Leap Motion's virtual reality SDK. Based on the tracked hands and fingers, events for temporal navigation are created. The tracking of the head-rotation for spatial navigation is already included in the Oculus Rift SDK. The system works with annular recordings of ODVs, e.g. as produced by Sony's Bloggie Cam.

SUMMARY, LIMITATIONS AND FUTURE WORK

We presented a system for interaction with ODVs using an amalgam of an immersive virtual reality headset and a gesture recognition system. The main novelty of the proposed system lies in the decoupling of spatial and time-based navigation. Yet, we have not investigated the potential of e.g. bi-manual interactions; time-based interactions are currently mapped to single-handed interactions. Also, using two separate input methods opens up a plethora of possibilities for interacting with different visualizations of ODVs. In the future, we also aim to investigate different visualization approaches such as focus+context and overview+detail inspired interface implementations [4]. Finally, we aim to conduct end-user evaluations in the near future.

ACKNOWLEDGMENTS

This work is partially funded by the SUTD President's Graduate Fellowship.

REFERENCES

1. Benko, H., Wilson, A.D., and Balakrishnan, R. Sphere: Multi-touch Interactions on a Spherical Display. *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, ACM (2008), 77–86.
2. Neumann, U., Pintaric, T., and Rizzo, A. Immersive Panoramic Video. *Proceedings of the Eighth ACM International Conference on Multimedia*, ACM (2000), 493–494.
3. Rovelo Ruiz, G.A., Vanacken, D., Luyten, K., Abad, F., and Camahort, E. Multi-viewer Gesture-based Interaction for Omni-directional Video. *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, ACM (2014), 4077–4086.
4. Ware, C. *Information visualization: perception for design*. Elsevier, 2013.
5. Zoric, G., Barkhuus, L., Engström, A., and Önnevall, E. Panoramic Video: Design Challenges and Implications for Content Interaction. *Proceedings of the 11th European Conference on Interactive TV and Video*, ACM (2013), 153–162.
6. 360cam. http://360.tv/.
7. Oculus VR | Oculus Rift - Virtual Reality Headset for Immersive 3D Gaming. *Oculus VR*. http://www.oculusvr.com.
8. Leap Motion. https://www.leapmotion.com/.
9. openFrameworks. http://www.openframeworks.cc/