



HAL
open science

Guidelines for Designing Interactive Omnidirectional Video Applications

Santeri Saarinen, Ville Mäkelä, Pekka Kallioniemi, Jaakko Hakulinen, Markku Turunen

► **To cite this version:**

Santeri Saarinen, Ville Mäkelä, Pekka Kallioniemi, Jaakko Hakulinen, Markku Turunen. Guidelines for Designing Interactive Omnidirectional Video Applications. 16th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2017, Bombay, India. pp.263-272, 10.1007/978-3-319-68059-0_17. hal-01679774

HAL Id: hal-01679774

<https://inria.hal.science/hal-01679774>

Submitted on 10 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Guidelines for Designing Interactive Omnidirectional Video Applications

Santeri Saarinen, Ville Mäkelä, Pekka Kallioniemi, Jaakko Hakulinen, and Markku Turunen

Tampere Unit for Computer-Human Interaction, University of Tampere, Finland
{santeri.saarinen, ville.mi.makela, pekka.kallioniemi,
jaakko.hakulinen, markku.turunen}@sis.uta.fi

Abstract. Interactive omnidirectional videos (iODV) can offer informative, entertaining, and immersive experiences, especially when combined with novel platforms such as head-mounted displays. However, omnidirectional videos, and interaction with them, present many unique challenges. In the absence of existing guidelines that accommodate for these challenges, we present dos and don'ts for designing and producing interactive omnidirectional videos. We base these guidelines on numerous interactive systems that we have produced in the recent years. Our work offers useful guidance for those working with omnidirectional videos, especially when designing interactivity and navigation within such systems.

Keywords. Omnidirectional videos · Virtual reality · Head-mounted displays

1 Introduction

Omnidirectional videos (ODV) have received a lot of attention lately due to new technologies for recording and producing such content. They are typically recorded with cameras that cover up to 360 degrees of the scene. Active research efforts have been taken in different domains of ODV ecosystem, including capturing, displaying and interaction technologies, and platforms such as YouTube and Facebook have provided their own content distribution channels for sharing ODV content. This study focuses on iODVs, which are ODV applications with additional interaction in addition to looking around the scene [8]. This interaction could be, for example, in the form of activating UI elements for more information on different objects in the scene, or transitioning from one ODV scene to another. We do not limit on which platform the iODVs should be viewed, although our research concentrates more on content displayed with a head-mounted display (HMD).

Albeit being a relatively new technology, omnidirectional videos have already been used in many domains and contexts. Some examples include use in remote operations and telepresence applications [4, 5], consumer products [9], museums [10], and theatre [6]. Many of these applications offer interactive content and have UI elements which

are often crucial features for pleasant user experience [3, 14]. Different interaction techniques in iODV applications have also been researched in many studies. Some examples of these include head-position, dwell-time based interaction [8], gesture-based interaction [2, 13, 16] and second screen interfaces [15].

Despite growing interest in interactive omnidirectional videos, most existing research has focused on applying iODVs in specific contexts, or on evaluating a specific design solution. However, the many unique challenges in the production and design of iODVs have gone largely unreported. Therefore, we believe reporting our extensive experiences with various iODV applications in varying environments are useful to other researchers and practitioners.

In this paper, we present guidelines to help in the design, recording, and production of interactive omnidirectional video applications. In particular, we focus on the design of omnidirectional videos with regards to interaction and navigation. With interaction, we primarily refer to embedded content, such as text and pictures that can be utilized to offer more information on important objects within a scene. With navigation, we refer to the ability to move between several videos, which can be used to e.g., move through an industrial hall, or view an object from several angles. Some of our guidelines are highlighted especially for certain platforms. Argyriou et al. [1] have presented similar guidelines for omnidirectional videos in their research. While their work was comprehensive in regards to immersion and some technical aspects of implementation, we offer new guidelines as well as alternative solutions to some problems.

We base our recommendations on several real-world projects we have developed over the recent years which have utilized iODV content in different ways. In the next chapter, we describe some of these projects, after which we present our guidelines. Finally, we conclude with a summary of our work.

2 Interactive Omnidirectional Video Applications

In this section, we present iODV applications based on which we present our guidelines later in this paper. The use cases vary from entertainment to navigation and industrial use, and were carried out in collaboration with several large industrial companies and cultural institutions.

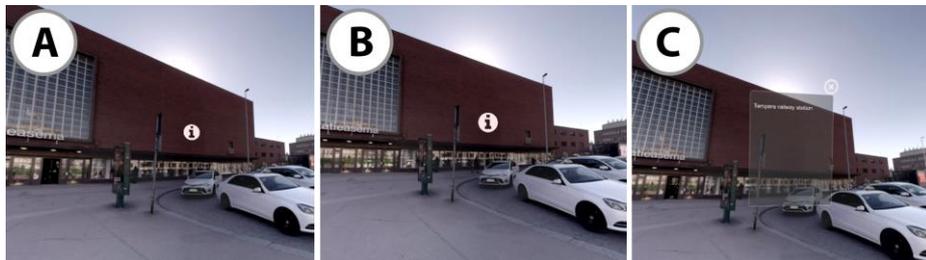


Figure 1. Interaction with hotspots. A: A hotspot is embedded on a building. B: The user focuses the hotspot at the center of the viewport, and the hotspot starts growing to visualize dwell time. C: The hotspot is triggered, and the content is shown.

In most presented cases, embedded content (*hotspots*) appear in the depicted scenes in the form of small icons (Figure 1A). When using a head-mounted display, users generally trigger these hotspots through dwell time, i.e., they center the hotspot in the middle of the viewport and wait for a short period for it to activate. During this time, the hotspot grows larger to visualize that it is being triggered (Figure 1B).

There are two types of hotspots. *Info hotspots* offer additional information on a corresponding object, usually in the form of text (Figure 1C) or pictures. The appearing information dialog is closed by simply moving the viewport away from it. *Navigation hotspots* transfer the user to another ODV, allowing users to move within the depicted location.

2.1 Maintenance Procedures

We developed several systems in cooperation with industrial partners to aid with maintenance of different industrial machines and equipment, for example large-scale fuel engines or vehicle-mounted aerial work platforms (AWPs). These iODV applications provide the user with remote access to the worksite while preparing for maintenance visit or while learning to use different equipment. These industrial sites have the need for both general and in-depth views of the target vehicle/machine as well as the environment it is located in.

With our iODV solution we are able to offer the user the possibility to view the target equipment from different angles, move around it, take a closer look at important parts and to view how different operations are conducted. This is done by combining short, looping iODVs from different angles and distances from the target machine and longer, non-looping iODVs, which present different actions the machine in question can perform. Therefore, a large network of omnidirectional videos is created, in which the user can move freely, and access information about the machine as well as the location.

2.2 Simulator Installations

We created two recreational simulator installations that utilize iODV content. They utilize videos filmed inside or on top of a vehicle while the vehicle in question is moving. Both simulators are currently in active use in an automobile museum.

The first simulator presents a road grader (Figure 2A). It includes a low-tech cabin, including the seat, steering wheel and pedals, and a three-display set surrounding the front side of the cabin. This display is used to present the video material, and the user can interact with it by using the steering wheel to rotate the view and by pressing the gas pedal to simulate speeding up.

The second simulator is a rally car installation that uses the Samsung Gear VR headset and headphones inside an actual, stationary rally car. ODV content is presented in a Gear VR application, which shows a video filmed during a test drive by a professional rally driver (Figure 2B). This allows the user to experience the feeling of sitting inside a rally car, while actually sitting inside a rally car.

The use of iODVs in these simulators is relatively similar. Both utilize long, moving videos filmed inside (or in the case of the road grader, on top of) a vehicle. These films by themselves are not interactive, but the interactivity in the simulators is done in other

ways compared to the industry demos presented earlier. In the road grader simulator, the presentation of iODV content is changed based on user interaction, by rotation, or by adding effects to the video to simulate faster speed. In the rally simulator, the interactivity is limited to starting the rally session, as it is more concentrated on immersion and the experience of rally driving, which most people have no other chance of experiencing. In the next version, we will provide the user with more interactive content such as information on the route driven.

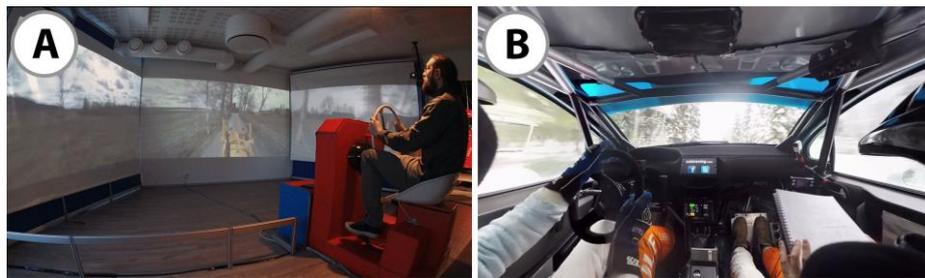


Figure 2. Simulators using ODV content. A: A road grader simulator where the video is projected on the wall using three projectors. B: Rally car simulator used with a VR headset.

2.3 Virtual Tour Applications

We created several virtual tour applications that utilize iODV content, which concentrate on free-form navigation between videos. Tour applications can be utilized in various ways. Our city tour application was used for language learning, where two students needed to collaborate in a foreign language to find the correct location in a city.

Another use case was with a university campus tour project, which was especially intended as a novel way for new students to familiarize themselves with the campus and the buildings and services within. As a third use case, we created virtual industrial complex tours for our industrial partners. These were used, e.g., for promotional purposes in exhibitions.

3 Guidelines for Designing iODV Applications

In this section, we describe guidelines derived from our experiences with iODV applications. In these guidelines, we concentrate on issues related to interactivity and user experience of iODV applications, instead of issues arising from the context of filming ODV content.

3.1 Avoid Objects Very Close to the Camera

Objects very close to the camera can obstruct useful information and can be disturbing to the user, especially so if the video is viewed on a head-mounted display. In ODVs, the camera within the video is a point-like object in the world space. As such, the objects can exist as close to the camera as the developer/producer wishes. However,

when using the iODV application, the user with HMD assumes his body takes up the same amount of space that it normally does. This causes an invading feeling, if some objects are too close. These objects can be anything from walls to tables to other people. The feeling of “in your face” can be off-putting and invasive, and should be accounted for in the production phase. This effect of invading the ‘peripersonal space’ has been studied extensively in both real world [11] and in virtual environments [12]. This was especially noticeable in our virtual city tour application where the videos were shot in the middle of a city. People passing by were interested in the camera, and often lingered around it or looked straight at it, sometimes at very close distance (Figure 3A).

Another example is the rally simulator, as the space inside the car was very cramped, and offered limited options for attaching the camera. Moreover, rules and regulations applied for filming inside rally cars, for instance, the camera was not allowed to reach into the front seat area, and had to remain further back. To get a good view of the windshield and the road ahead, however, we put the camera roughly between the two front seats (Figure 2B). While the primary goal was reached with a good view of the cockpit and the road, the front seats ended up being somewhat disturbing as they seemed to be unnaturally close to the viewport when users were looking around in the car with a HMD (Figure 3B).

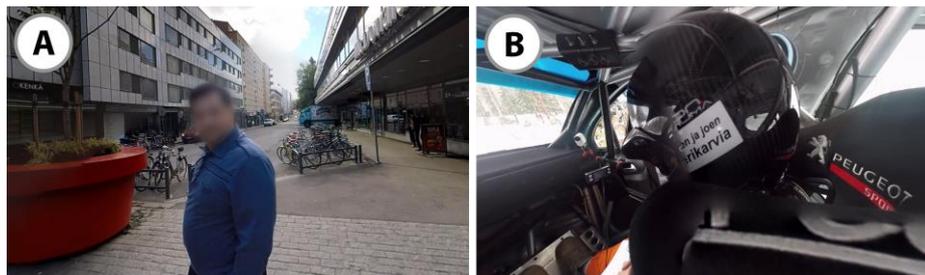


Figure 3. Examples of disrupting objects. A: A man looking straight at the camera as he is passing by. B: The navigator’s seat in the rally simulator.

3.2 Consider an Appropriate Viewpoint

It is important that the viewpoint, i.e., where the camera is placed, supports the context and the use case. For instance, it may be of importance that the camera is placed a) on a platform that is accessible to people, and b) at around the same height as the person’s viewpoint would be if they really were in the location. This guideline also relates to the notion of making the navigation in the virtual environment natural, as mentioned in related work by Argyriou et al. [1]. In some of our industrial cases, iODVs were used as a way for technicians to familiarize themselves with the location before going on-site in person. Therefore, the view in the videos needed to represent what the technicians would actually see when arriving at the site. Moreover, some users reported feeling dizzy when viewing a video that was filmed from the top of a ladder. This was relatively surprising, as the camera was no more than around three meters above the floor. The ladder was used in an industrial setting to provide a full view of an industrial hall, but in this case the solution worsened the user experience.

3.3 Present Details with Embedded Content

Some low-level details and procedures are difficult to present with ODVs. For example, in industrial context, we aimed to record ODVs of maintenance procedures, to be used as reference material and documentation for future maintenance technicians. We found that capturing the fine details of the procedure, such as interacting with a complicated control panel in a correct way, was problematic. Primarily, it was difficult to place the camera close enough to show that much detail properly (Figure 4A). Moreover, the technician conducting the procedure would often obstruct the object from the camera (Figure 4B). In these situations, the role of embedded content, such as pictures, text, and 2D video, is emphasized, to better visualize the details. It is also worthwhile to note that with embedded content (hotspots), one can better guide the user's attention to meaningful objects and events, without the need to rely on the user to always know what to focus on. While Argyriou et al. [1] argued that the UI should be subtle and non-intrusive, in some cases, more disruptive elements could be utilized if it is important to direct the user's attention towards certain elements which she might not otherwise notice. This would be the case in scenarios with a more focused narrative, such as the aforementioned maintenance procedures.



Figure 4. Situations where embedded content is needed. A: A technician operating a large control panel. B: A technician adjusting a small part of a crane, blocking the view with his hands.

3.4 Ensure the Visibility and Clarity of Interactive Objects and Pathways

When interactive hotspots are added to the scene, it should be clear to the user which objects these hotspots are referring to. Whether a hotspot can be overlaid directly on the object depends on a multitude of characteristics, such as the overall clarity of the scene as well as the density of interactive and important objects. In some cases, hotspots can occlude the object they are referring to, or otherwise make the scene difficult to make sense of. In our industrial applications, we noticed that placing hotspots right on top of certain objects made the application more difficult to use, as the users were not aware of which object was under the interactional element. This primarily applied to scenes that contained many fine details, for example, a control panel with a large number of buttons, switches, and screens. Another example is our city tour application: in Figure 5A, the information hotspot on the right seemed confusing to users, as it was

blocking the view forward and it was not clear what it was referring to. In another application, we successfully used transparent hotspots to more clearly highlight important objects without occluding them (Figure 5B).

With regards to navigation, the user should be able to understand where a navigation hotspot will take them. For example, in a scene where a large skylift was shown, and a path to the other side of the skylift was offered, it turned out to be confusing as we placed the hotspot on the skylift itself as if traveling “through” it (Figure 5C). Instead, the hotspot could be visualized to suggest that the user is going around the vehicle. Also, the pathway to the following scene should not be completely covered by the interactive elements. As found by Kallioniemi et al. [7], obscuring the visibility of such routes makes the navigation tasks in 360 degree environments more difficult.

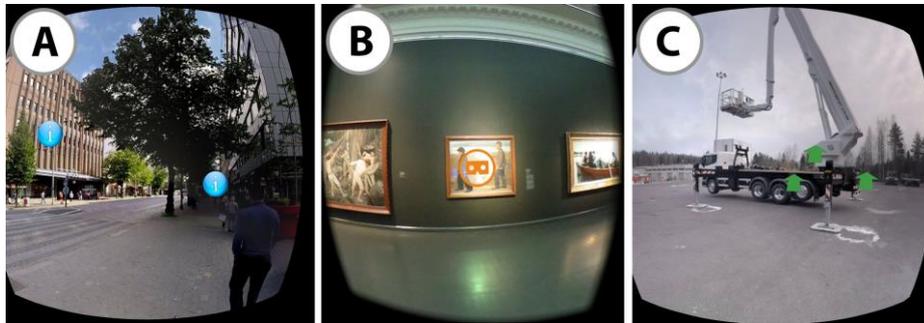


Figure 5. Interactive objects embedded in omni-directional videos. A: The hotspot on the right is blocking the view forward, and it is unclear what the hotspot is referring to. B: An example of a clear, transparent hotspot, containing more information of a painting. C: Examples of unclear navigation hotspots.

3.5 Visualize Transitions When They Are Not Obvious

Moving between several omnidirectional videos may quickly make the user lose sense of where they came from and where their current location is relative to other videos they have moved through. Maintaining navigational awareness is especially important for applications that aim to familiarize users with a remote location before moving on-site.

To support navigational awareness in our campus tour project, we included fast-forwarded non-interactive omnidirectional videos between transitions. These were useful when, e.g., transitioning from one building to another, to visualize which door and path to take to get to the next building. However, if the start point of the transition is visible from the end point, it is not necessary to display a transitional video, but simple fade-in technique is enough.

On a related note, the user should be correctly oriented after each transition. For instance, when transitioning to the next room through a door, the user should be oriented with their back against the door they came through. While this may seem obvious, it is worthwhile to note that in order to achieve this, the relative orientation between

each link must be set. Therefore, if three paths lead to the same video, the orientation for each path is different.

3.6 Plan the Complete Pathways Ahead of Time

The complete pathways and transitions between videos should be planned ahead of time, by e.g., writing a script with the full experience in mind. This becomes especially important with more complex applications. For example, in one of our industrial applications, users could move around an industrial hall both indoors and outdoors, and e.g., travel high above the ground riding a skylift (Figure 5C). Therefore, the application consisted of both looping and non-looping videos. Looping videos were those where users could spend as much time as they want, until they chose to move to a different scene through a hotspot. When users chose to ride up on the skylift, a non-looping crane video would start, at the end of which the users would automatically transition to a new looping video on top of the crane, which would play until the user chooses to take the pathway back down. Therefore, to seamlessly integrate transitions as well as static and moving videos, the full path needs to be carefully planned.

4 Conclusion

Interactive omnidirectional videos (iODVs) can offer a wide variety of useful, exciting, and entertaining experiences. As a relatively new platform, however, iODVs seem to lack basic guidelines to guide practitioners.

Our guidelines for interactive omnidirectional video application, are summarized as follows: 1) Avoid objects very close to the camera, as they obstruct large segments of the surroundings, and may be disturbing to the user. 2) Choose a viewpoint that supports the context and use case of the application. 3) Present details with embedded content, as omnidirectional videos alone cannot always present fine details clearly. 4) Ensure the visibility and clarity of interactive objects and pathways. Using different visual cues, make sure that the link between a hotspot and the related object is clear. In the case of navigational hotspots, communicate clearly where the hotspots will take the user. 5) Visualize transitions when they are not obvious. Navigational awareness is important, especially if the iODV application aims to familiarize users with the depicted location. An especially useful trick is to include fast-forwarded omnidirectional videos between transitions, to e.g. visualize a transition from one building to another. 6) Plan the complete pathways ahead of time. This is especially important when combining free-roaming multi-path videos with static videos – ensuring smooth transitions and avoiding “dead-ends” requires prior planning.

We based our guidelines on numerous projects we have worked on over the recent years. We focused especially on applications that offer interactivity through embedded content, and allow transitioning between multiple videos. With our work, we hope to guide those working with omnidirectional videos, especially when designing interactivity and navigation within such systems.

References

1. Argyriou, L., Economou, D., Bouki, V. & Doumanis, I. (2016). *Engaging immersive video consumers: Challenges regarding 360-degree gamified video applications*. In IEEE International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security, IUCC-CSS '16. IEEE 2016, 145-152.
2. Benko, H. and Wilson, A. D. (2010). *Multi-point interactions with immersive omnidirectional visualizations in a dome*. In ACM International Conference on Interactive Tabletops and Surfaces (ITS '10). ACM, New York, NY, USA, 19-28.
3. Berning, M., Yonezawa, T., Riedel, T., Nakazawa, J., Beigl, M. & Tokuda, H. (2013). *360 degree interactive video for augmented reality prototyping*. In Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication, UbiComp '13, ACM (New York, NY, USA, 2013), 1471-1474.
4. Boulton, T. E. (1998) *Remote Reality via omnidirectional imaging*. In ACM SIGGRAPH 98 Conference abstracts and applications (SIGGRAPH '98). ACM, New York, NY, USA.
5. De la Torre, F., Vallespi, C., Rybski, P. E., Veloso, M., & Kanade, T. (2005). *Omnidirectional video capturing, multiple people tracking and identification for meeting monitoring*. Technical report. <http://repository.cmu.edu/robotics/128/>.
6. Decock, J., Van Looy, J., Bleumers, L. and Bekaert, P. (2011). *The Pleasure of Being (There?)*. An Explorative Study into the Effects of Presence and Identification on the Enjoyment of an Interactive Theatrical Performance using Omnidirectional Video. In Proceedings of the International Society for Presence Research Annual Conference (Edinburgh, Scotland, October 26--28, 2011). ISPR 2011. Edinburgh Napier University, Edinburgh, Scotland, UK, 12p.
7. Kallioniemi P., Hakulinen J., Keskinen T., Turunen M., Heimonen T., Pihkala-Posti L., Uusi-Mäkelä M., Hietala P., Okkonen J. & Raisamo R. (2013). *Evaluating landmark attraction model in collaborative wayfinding in virtual learning environments*. In Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM '13). ACM, New York, NY, USA.
8. [Redacted]. (2017) *User Experience and Immersion of Interactive Omnidirectional Videos in CAVE Systems and Head-Mounted Displays*. In Press.
9. Kasahara, S., Nagai, S., Rekimoto, J. (2015). *First Person Omnidirectional Video: System Design and Implications for Immersive Experience*. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video, June 03-05, 2015, Brussels, Belgium.
10. Kwiatek, K. (2012). *How to preserve inspirational environments that once surrounded a poet? Immersive 360° video and the cultural memory of Charles Causley's poetry*. In Virtual Systems and Multimedia (VSMM), 2012 18th International Conference, pp. 243-250.
11. Lourenco S.F., Longo M.R., Pathman T. (2010). Near space and its relation to claustrophobic fear. *Cognition* 119. pp. 448-453.

12. Iachini T., Coello Y., Frassinetti F. & Ruggiero G. (2014). *Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality*. PLoS One, 9(11), e111511.
13. Petry B. & Huber J. (2015). *Towards effective interaction with omnidirectional videos using immersive virtual reality headsets*. In Proceedings of the 6th Augmented Human International Conference (AH '15). ACM, New York, NY, USA.
14. Ramalho, J., and Chambel, T. (2013). *Windy sight surfers: sensing and awareness of 360° immersive videos on the move*. In Proceedings of the 11th european conference on Interactive TV and video, EuroITV '13, ACM (New York, NY, USA, 2013), 107-116.
15. Zoric, G., Barkhuus, L., Engström, A., and Önnvall, E. (2013a). *Panoramic Video: Design Challenges and Implications for Content Interaction*. Proceedings of the 11th European Conference on Interactive TV and Video, ACM, 153–162.
16. Zoric, G., Engström, A., Barkhuus, L., Ruiz-Hidalgo, J., and Kochale, A. (2013b). *Gesture interaction with rich TV content in the social setting*. In *Exploring and Enhancing the User Experience for Television*, Workshop of ACM SIGCHI Conference on Human Factors in Computing Systems, CHI'13 (Paris, France, 04/2013 2013).