Managing Mutual Occlusions between Real and Virtual Entities in Virtual Reality

Guillaume Bataille[1,2][a], Valérie Gouranton[2], Jérémy Lacoche[1] and Bruno Arnaldi[2]
[1] *Orange Labs, Cesson Sévigné, France*
[2] *University Rennes, INSA Rennes, Inria, CNRS, IRISA, France*

Keywords: Virtual Reality Display Wall, Mutual Occlusions, Mixed Interactive Systems, Multi-layered Stereoscopy.

Abstract: This paper describes a mixed interactive system managing mutual occlusions between real and virtual objects displayed by virtual reality display wall environments. These displays are physically unable to manage mutual occlusions between real and virtual objects. A real occluder located between the user's eyes and the wall hides virtual objects regardless of their depth. This problem confuses the user's stereopsis of the virtual environment, harming its user experience. For this reason, we present a mixed interactive system combining a stereoscopic optical see-through head-mounted display with a static stereoscopic display in order to manage mutual occlusions and enhance direct user interactions with virtual content. We illustrate our solution with a use case and an experiment proposal.

1 INTRODUCTION

Virtual reality display wall environments use static stereoscopic displays in order to display virtual environments (VE). Virtual reality display wall environments are typically CAVE presented by Cruz et al. (Cruz-Neira et al., 1992), CAVE2[1], Powerwall[2], or immersive rooms. When these displays are physically occluded by real entities from the user viewpoint, these systems cannot properly display mutual occlusion between real and virtual entities. Yet, occlusions are monocular depth cues which are important for the credibility of displayed VEs. Indeed, wrong occlusions are proved to confuse the cognition of users as Sekuler et al. observed it (Sekuler and Palmer, 1992). They are illusion and immersion breakers for virtual reality display wall users. Figure 1(a) emphasizes this situation, where the virtual banana should not be partly occluded by the user's hand.

On the other side, optical see-through head-mounted displays (OST-HMD), like Hololens or Magic Leap, render stereoscopic content which cannot be occluded by real entities. They provide, to a greater or lesser extent, occlusion culling of virtual entities occluded by static or slightly mobile real entities. When available, mutual occlusion management is based on the 3D reconstruction of the user's real environment, as presented by Walton et al. (Walton and Steed, 2017). Figure 1(b) describes the case of an OST-HMD unable to manage mutual occlusions. In that case, the OST-HMD displays a virtual object located behind the user's hand and which should be hidden by it.

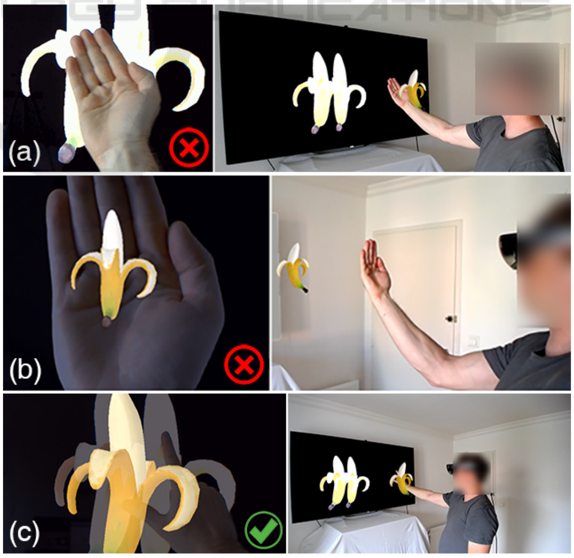

Figure 1: Mutual occlusions between real and virtual. (a) Virtual objects rendered by static stereoscopic displays are likely to be inappropriately occluded by real ones. (b) An OST-HMD requires mutual occlusion management to occlude the virtual by the real. (c) Our mutual occlusion management of a VE rendered by a static stereoscopic display and an OST-HMD.

---

[a] https://orcid.org/0000-0002-6751-3914

[1] https://www.evl.uic.edu/cave2

[2] https://www.lcse.umn.edu/research/powerwall/powerwall.html

Bataille, G., Gouranton, V., Lacoche, J. and Arnaldi, B.
Managing Mutual Occlusions between Real and Virtual Entities in Virtual Reality.
DOI: 10.5220/0010213800370044
In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021) - Volume 1: GRAPP*, pages 37-44
ISBN: 978-989-758-488-6

Compared to single display systems, multiple stereoscopic layout systems enhance pros and circumvent cons of their displays. In this manner, they enhance the visual perception of virtual environments. The literature mainly couples static displays and mixed reality devices like tablets or optical see-through head-mounted displays (OST-HMD). A recent approach by Nishimoto et al. combines a virtual reality display wall environment with a Hololens in order to expand its Field of Regard (FoR) (Nishimoto and Johnson, 2019). We propose to extend this system to the management of mutual occlusions between virtual and real objects or body parts. Our system, visible on Figure 1c, is compatible with both stereoscopic TVs and virtual reality display wall environments.

In the next sections, we first describe related work. Second, we outline our solution, which we called Stereoccluder, a mixed interactive system providing a multi-layered stereoscopic rendering dedicated to the mutual occlusion management of virtual reality display wall environments. Third, we present a use case based on the rendering of a cat mummy. Fourth, we describe an experiment that we designed and developed in order to evaluate our system. Finally, we conclude this research and present our future work.

2 RELATED WORK

The problem of occlusions between real and virtual objects has been first addressed by Wloka et al. (Wloka and Anderson, 1995). Their solution uses a stereoscopic camera to estimate the real environment depth. A Video See-Through Head-Mounted Display (VST-HMD) displays virtual objects occluded by real ones depending on the estimated depth. The image-based approach of Walton et al. (Walton and Steed, 2017) to this problem relies on the use of an RGBD camera to capture the current scene. They compose the rendering of virtual content with filtered RGB frames of the real environment. Gimeno et al. (Gimeno, 2018) use a different approach, based on the rendering of a real area occluded by a real object. In that case, they render the virtual twin (Kritzler et al., 2017) of the occluded area. However, these approaches do not apply to the real occlusion of static stereoscopic displays by users or real objects.

Multi-layered 3D-based rendering approaches render an environment from different types of input, called layers, as presented by Kang et al.(Kang and Dinh, 1999). Each virtual layer view synchronizes its viewport with others in order to produce a composite view. The same principle can be extended to multiple display systems composing the user viewpoint.

Combined display systems are studied for their potential in sharing or distributing renderings and interactions. When users simultaneously see multiple displays, they observe a multi-layered rendering of a virtual environment in real-time.

The first sort of association is the combination of two static displays. Projection is used by the Illumiroom system presented by Jones et al. (Jones et al., 2015) in order to enhance a screen FoV. In that case, a projector displays, on the wall around the screen, a larger portion of the screen content. Both displays are colocated but are also static and monoscopic. Also, no occlusion is managed and interactions occur only with one device.

A second association is the combination of two mobile displays, a mixed reality head-mounted display, and a smartphone. Normand et al. (Normand and McGuffin, 2018) augment a smartphone screen with an OST-HMD or a video see-through head-mounted display (VST-HMD). In that case, no screen is static, the smartphone provides static rendering and both see-through head-mounted displays are stereoscopic. Both displays are colocated, occlusions are managed in the VST-HMD case, and interactions are shared between devices.

A third association, the most common one, is the combination of a static display, either a projector or a screen, with a mixed reality display, like a head-mounted display or a tablet. Projector-based spatial augmented reality (SAR) employs projectors in order to interact intuitively with a mixed environment (Raskar and Low, 2001). This mixed environment is obtained by projecting a virtual environment aligned on real surfaces. For example, Roo et al. (Roo and Hachet, 2017) use a projector in order to augment a physical mock-up made out of sand with a volcano image. They also combine a see-through display with a spatially augmented motor. They indicate that the use of projectors has clear occlusion limitations. Sandor et al. (Sandor et al., 2002) presented The SHEEP system in 2002 (MacWilliams et al., 2003). This system projects on a table a video game. Both projector and device screens are monoscopic and colocated by an ART tracking system. An HMD is used but its specifications are not provided. Interactions are provided by devices but occlusions are not managed. Kurz et al. (Kurz et al., 2008) solve mutual occlusions in the case of table-top displays by projecting occluded virtual content on real occluders. But in their case, the projector is static, bounding user locomotion to the table-top space. Alternatively, head-worn projectors could be used instead of OST-HMDs, but they are vulnerable to lighting conditions, materials of projected surfaces, and multiple real occlud-

Table 1: Comparison between mixed interactive systems blending multiple displays.

Work	static display	MR display	mutual occlusion management
Sandor et al. 2002	monoscopic	monoscopic	no
Kurz et al. 2008	**stereoscopic** (dual)	no	**yes**
Kawakita et al. 2014	monoscopic	monoscopic	no
Jones et al. 2015	monoscopic (dual)	no	no
Benko et al. 2015	monoscopic	**stereoscopic**	no
Baillard et al. 2017	monoscopic	**stereoscopic**	**yes**
Roo et al. 2017	monoscopic	monoscopic	no
Normand et al. 2018	no	**stereoscopic**	**yes** (VST-HMD)
Saeghe et al. 2019	monoscopic	**stereoscopic**	no
Nishimoto et al. 2019	**stereoscopic**	**stereoscopic**	no
Stereoccluder	***stereoscopic***	***stereoscopic***	***yes***

ers. Benko et al. (Benko et al., 2015) present in 2015 the FoveAR system. This system combines an OST-HMD with SAR projections in order to enhance the FoV of the OST-HMD as a hybrid display. Projection is monoscopic while OST-HMD is stereoscopic, both are colocated and no real object occlusion is managed. Rendering is shared by both displays. Monoscopic TVs are also associated with mixed reality devices by several mixed reality systems. Kawakita et al. (Kawakita and Nakagawa, 2014) complete a TV program with a mobile device running augmented reality techniques. Collocation results from the detection of a 2D tag displayed by the TV screen by the mobile device, which produces a pose estimation latency compared to inside-out reconstruction techniques embeded by ARCore or ARKit mobile devices, Hololens or Magic Leap. But real occlusions of the TV tags brake collocation. Still, interactions are distributed between the TV and the mobile device, so any user interaction with one of these devices impacts both. Also, Baillard et al. (Baillard et al., 2017) augment 2D TV programs with an AR device like a tablet or an OST-HMD. Displays are colocated, occlusions are said to be managed, interactions are distributed between displays and the OST-HMD is stereoscopic. Similarly, Saeghe et al. (Saeghe et al., 2019) augment a TV program with an OST-HMD. They explore how user interactions can influence the program storytelling, but do not address occlusion management. Finally, Nishimoto et al. (Nishimoto and Johnson, 2019) augment the FoR of a CAVE2 with an OST-HMD. In that case, the OST-HMD is in charge of displaying virtual content upper and below the CAVE2 stereoscopic screens. This system provides a full stereoscopic rendering. Both OST-HMD and CAVE2 are colocated to provide a consistent multi-layered stereoscopic rendering. Interactions are not shared, are non-direct, and are limited to the use of a PS3 wand. This system does not manage mixed occlusions. Our intuition is that the same system with occlusion management would provide a better user experience and a more significant task performance improvement.

We summarize our comparison of these systems regarding mixed occlusion management in Table 1.

Consequently, we want to combine a stereoscopic display with a stereoscopic OST-HMD in order to manage occlusions between real and virtual entities. We also target to evaluate their impact on user direct gestural interactions. We expect such mixed interactive systems to enhance the user experience by reducing visual breaches due to real occlusions of virtual content.

3 OUR APPROACH

We present in this paper our contribution to solving mutual occlusions occurring between the real and the virtual in virtual reality display wall environments. Our approach is simple and flexible, and relies on known techniques and devices.

Our approach (see Figure 2) extends the approach of Nishimoto et al. (Nishimoto and Johnson, 2019), which consists in combining static stereoscopic displays with OST-HMDs. When a real occluder like a real object or the user's body partly hides this VE (Virtual Environment), the VE located between the occluder and the user's eyes, the green area in Figure 2, should be visible. Our approach consists of :

- detecting and tracking real occluders to estimate which part of the hidden VE should be visible to the user,
- displaying the inaccurately hidden VE with an OST-HMD.

Our system relies on the ability of the OST-HMD to self-locate in its real environment.

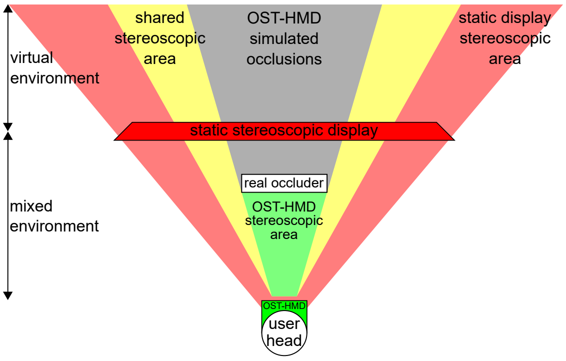

Figure 2: A segmentation of the multi-layered stereoscopic space. The OST-HMD must display the VE part inappropriately hidden by a real occluder in the green area, and simulate real occlusions of the VE in the purple area.

3.1 Setup

Our experimental setup is composed of a Hololens 1 (the OST-HMD), an active stereoscopic TV (the static stereoscopic display), and a computer in charge of the static stereoscopic display rendering (see Figure 3). Stereoscopic TVs are widely available at low cost, but our system can be easily transposed to immersive rooms, CAVEs, or Powerwalls. The stereoscopic TV is placed on a sit/stand workstation to adjust the TV center to the height of the user's gaze.

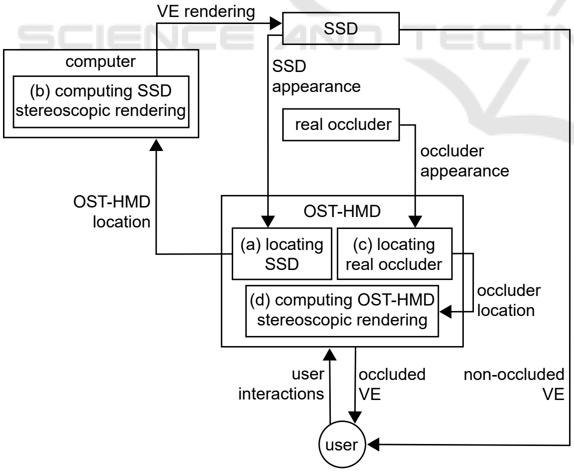

Figure 3: Our Stereoccluder system overview. (a) the OST-HMD locates the pose of the static stereoscopic display. (b) the static stereoscopic display displays the VE from the OST-HMD viewpoint. (c) the OST-HMD locates real occluders. (d) the OST-HMD displays the inappropriately occluded VE.

The user simultaneously wears the OST-HMD and the active stereoscopic goggles paired with the stereoscopic screen, as shown in Figure 4. The active polarizing filter of the goggles interferes with the Hololens internal optics. It hides one RGB color from screen images on a slow cycle. This problem is circumvented by (Nishimoto and Johnson, 2019) since their system is based on passive stereoscopic screens. They position passive filters upon the external glass of the Hololens. In our case, considering the small size of the active stereoscopic goggles, we did not try to use them upon the OST-HMD to prevent color interferences. Displaying only RGB colors composed of at least two primary colors avoids stereoscopic breaches, but still produces minor perceptual concerns due to one fading primary color per cycle.

Figure 4: A user simultaneously wearing active stereoscopic goggles and an OST-HMD. At the top, an overview of the system in use. At the center, a close view on dual display wearing. At the bottom, the right user's eye view.

Our system is implemented with a Hololens 1 and a 65" Samsung 3D TV connected to an Alienware Laptop (RTX 2080, 32Go Ram, I9 9th gen). Both OST-HMD and laptop applications are developed with Unity 2018.4. The Hololens application uses the Mixed Reality Toolkit 2.2.0. Both applications communicate through Wi-Fi, using our own TCP network layer implemented in .NET.

3.2 Calibration

The calibration phase initially enables the OST-HMD to locate the static stereoscopic display (see Figure 3(a)). While the static stereoscopic display displays a texture, the OST-HMD estimates its pose, thanks to Vuforia[3]. This computation allows the OST-HMD and the computer applications to share a common coordination system in order to render and display a VE from the same viewpoint. In the case of drift, the user can recalibrate the system. An alternative solution would be to fix 2D tags on the static stereoscopic display borders and let the OST-HMD constantly track them, but it entails an unstable tracking and requires additional computing power.

3.3 Multiple Stereoscopic Rendering

Both the OST-HMD and the static stereoscopic display must render the VE from the same viewpoint for a consistent and homogeneous multiple stereoscopic rendering. For this reason, the OST-HMD requires to know the location of the static stereoscopic display (see Figure 3(a)) to render the VE (see Figure 3(d)). Conversely, the computer requires to know the current user's head position to render the VE accordingly (see Figure 3(b)). For this reason, the OST-HMD constantly shares with the computer its own position and orientation. Therefore, the computer is aware of the position and orientation of the virtual cameras simulating the user's eyes in order to render the VE from the same user viewpoint. The computer uses this knowledge in order to calculate the projection matrix of the virtual cameras corresponding to the user's eyes[4].

Nishimoto et al. (Nishimoto and Johnson, 2019) estimate the pose of the OST-HMD with an ARTTRACK system[5] commonly used by virtual reality display wall environments to track the user's head. But this system is exposed to real occlusions, is expensive, and requires to instrument the setup environment initially. A part of our contribution consists in using the inside-out tracking system of the OST-HMD instead. The user's head tracking is slower and less responsive than ARTTRACK systems, but also more flexible since OST-HMD do not require to instrument the static stereoscopic display area with ART trackers. Furthermore, this tracking system is more robust against real occlusions. Indeed, real occluders may hinder ART sensors from tracking the ART markers mounted on the user's head.

3.4 Occlusion Management

Our contribution to the management of mutual occlusions in the case of virtual reality display wall environments is a system combining an OST-HMD and a static stereoscopic display. This system is capable of managing mutual occlusions properly in a simple and portable manner. Our system displays on an OST-HMD the VE inaccurately hidden by a real occluder and located between this occluder and the user's eyes 3(d)). Our mutual occlusion management simulates the presence of real occluders in the real scene by adding the virtual twin (Kritzler et al., 2017) of the real occluder in the VE. This technique is known as the phantom technique (Fischer et al., 2004).

Our system detects and tracks real occluders (see Figure 3(c)) in order to simulate their presence in the VE displayed by the OST-HMD. A first solution tracks the real occluder, which can be a body part of a real object. For example, an OST-HMD (Hololens, Magic Leap, etc), or a Leapmotion (Nasim and Kim, 2016) can track the user's hand at different granularities. A Hololens 1 is able to roughly track a hand's position without neither its orientation nor finger tracking. Conversely, a Leapmotion tracks the position and orientation of fingers' jointures and palm tracking. Tracking can also be obtained by the pose estimation of a 2D tag glued on the object to track. We use a trackable handheld occluding surface presented in Figure 5. This surface is tracked and its virtual twin, the interactive virtual representation of a real entity for Kritzler et al. (Kritzler et al., 2017), is rendered with an occlusion shader. This is an alternative solution circumventing the inaccuracy of hand tracking by Hololens 1. A second solution estimates the depth between the user and the stereoscopic screen to compute the occluded area. Depth sensors can be combined for enhanced accuracy.

In Figure 1(c), the real object is the user's hand and the virtual object is a banana. As part of our contribution, the occluder is simulated on the OST-HMD as a virtual sphere with an occlusion shader assigned to it. From the user viewpoint, the occlusion shader hides virtuals objects behind it and shows virtual objects ahead of it. This sphere is larger than the hand to circumvent any lack of tracking accuracy. On Hololens 1, we use a sphere because this OST-HMD does not provide the hand rotation or the pose

---

[3] https:developer.vuforia.com

[4] https://fr.slideshare.net/N_Baron/view-frustum-in-the-context-of-head-tracking

[5] https://ar-tracking.com/

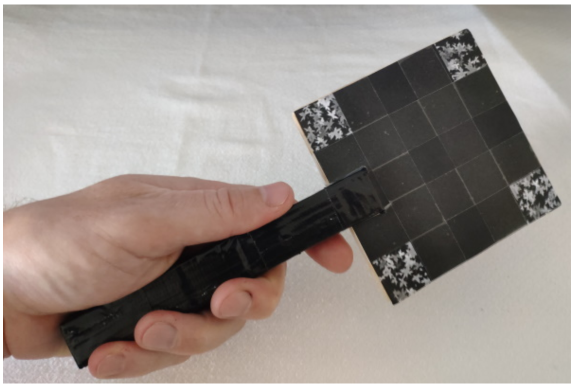

Figure 5: Our handheld occluding surface. The corners of this surface are trackable markers.

of its fingers. The user's hand needs to keep the air-tap ready position to get detected by the Hololens.

4 RESULTS

First, we present in this section a volumetric exploration of a cat mummy in a virtual reality display wall environment using our mutual occlusion management system, Stereoccluder. Second, we describe an experiment aiming at evaluating the benefits of this system for direct interactions between the user's hand and virtual objects displayed by virtual reality display wall environments.

4.1 The Cat Mummy Usecase

In this subsection, we present a multi-layout stereoscopic display of a virtual mummy. This use case allows the volumetric exploration of the internal remains of a cat mummy. The static stereoscopic display displays the mummy's appearance while the OST-HMD displays its internal remains. This interactive system benefits from our occlusion management system. Indeed, our framework allows the user to perceive the occluded virtual remains of the mummy although the static stereoscopic display is partly occluded. The real occluder is a handheld rectangular surface (see Figure 5). With this augmented real object behaving as an x-ray viewer, our system allows the user to see the internal remains of the mummy located between the user's viewpoint and the real occluder. This technique is known as the magic lens technique (Bier et al., 1993). The virtual mummy appearance was reconstructed by photogrammetry. The geometry of its internal parts was obtained by radiography. Figure 6 presents the results obtained with our handheld occlusion surface.

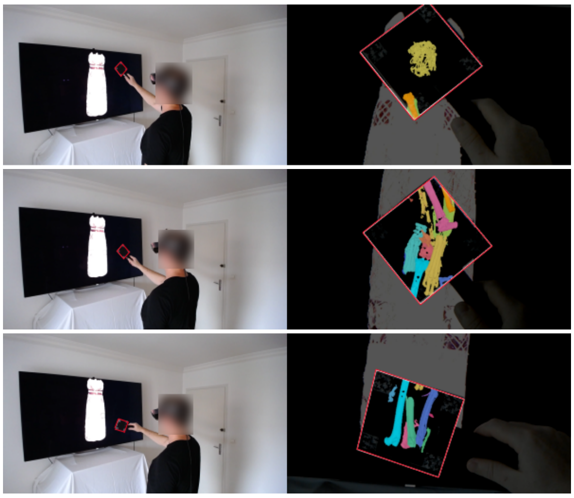

Figure 6: Volumetric exploration of a cat mummy revealing its internal remains. Our handled occluding surface, highlighted by a red border, acts as a virtual x-ray viewer.

Four users experienced Stereoccluder without measurements. The overall acceptance is good and the system is appreciated. The main reasons seem to be the expansion of the field of view and the accuracy and consistency of the multi-layout stereoscopic rendering, despite the tracking latency. They also experienced without measurements the impact of non-managed occlusions in the case of a single stereoscopic layout provided by the static screen. This case is similar to the use of CAVEs with no occlusion management. In that case, non-managed occlusions create a disturbing stereopsis breach. For that reason, we do not consider any experiment of direct user hand gestures without offsetting the virtual hand avatar. Otherwise, non-managed occlusions would impact experiments too heavily to provide relevant results.

4.2 Experiment

This subsection presents an experiment that we have designed and implemented, but not realized at the moment due to the COVID-19 crisis preventing the experiment completion.

We have designed an experiment to evaluate the benefits of our multi-layered stereoscopic rendering system for direct interactions between the user's hand and virtual objects. We chose a classical selection task to show that our system performs at least as well as existing ones. The task consists of searching for a secret sphere in a 5x3x4 grid of mystery spheres, all rendered with the same appearance. This task benefits from the large FoV provided by the stereoscopic screen. The spheres' radius measures 5cm. They are separated by 11cm in order to avoid the hand's sphere to simultaneously collide with multiple mystery spheres. In order yo detect the secret sphere, the user must touch mystery spheres with a virtual sphere located between his thumb and his index. This semi-direct interaction avoids occluding the sphere displayed by the screen with the user's hand. When the virtual sphere associated with his hand location collides with the secret sphere, the secret sphere color changes. The tester validates the discovery of the secret sphere with an air-tap gesture. A new sphere grid is then displayed.

Three conditions are considered. The two first conditions employ only one display, either Hololens (condition C1) or a static stereoscopic screen (condition C2). The third condition associates both displays (condition C3). All cases involving Hololens use our occlusion management system. We plan to measure task completion time, the time between spheres collisions, time to perceive the discovered secret spheres, subjective workload with raw NASA-TLX, usability with a Single Ease Question (SEQ), and overall preference. Our hypothesis are:

- H1: the Hololens-only condition (C1) should be slower because of more head movements in order to see the whole grid, for the highest cognitive load,
- H2: the static stereoscopic screen only condition (C2)should be the least pleasant and most disturbing because of the lack of occlusion management, for an average cognitive load,
- H3: the Hololens + static stereoscopic screen condition (C3) should be the most efficient for the least cognitive load, due to its large field of view and its occlusion management.

Figure 7 shows this experiment under C1, C2 and C3 conditions.

5 CONCLUSIONS

In this paper, we have presented how our multi-layered stereoscopic system solves mutual occlusions in the case of virtual reality display wall environments. Stereoscopic layers advantageously associate their strengths and weaknesses in order to provide a consistent rendering of partly occluded virtual environments displayed by static stereoscopic screens. An OST-HMD detects and tracks real occluders, simulates their presence in the virtual environment, and renders the VE part located between the user's head and real occluders. Our contribution grants the use of virtual reality display wall environments as mixed interactive spaces, where the presence of real objects and users is not an illusion breaker. This contribution also enables direct gestural interaction with virtual objects in such environments without vision breaches.

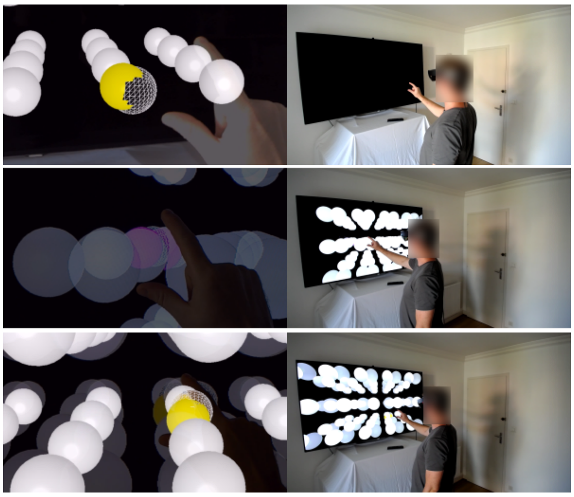

Figure 7: Experiment ”find the secret sphere”. The left column is a capture of the user's view. At the top, the user perceives the VE displayed by the OST-HMD only. At the center, the user perceives the VE displayed by the static stereoscopic display only. At the bottom, the user perceives the VE displayed by both the OST-HMD and the static stereoscopic display.

Future work will run the presented user study. This study aims at evaluating the impact of our system on direct gestural interactions. The intended evaluation task consists of a research task with three conditions, a stereoscopic screen only, an OST-HMD only, and a stereoscopic screen combined with an OST-HMD. Auto-stereoscopic screens and holographic displays are alternative displays against the use of stereoscopic OST-HMDs. Finally, we plan to experiment our solution with an immersive system with multiple screens, since Stereoccluder is compatible with such virtual reality display wall environments.

REFERENCES

Baillard, C., Fradet, M., Alleaume, V., Jouet, P., and Laurent, A. (2017). Multi-device mixed reality TV: a collaborative experience with joint use of a tablet and a headset. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology - VRST '17*, pages 1–2, Gothenburg, Sweden. ACM Press.

Benko, H., Ofek, E., Zheng, F., and Wilson, A. D. (2015). FoveAR: Combining an Optically See-Through Near-Eye Display with Projector-Based Spatial Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*, pages 129–135, Daegu, Kyungpook, Republic of Korea. ACM Press.

Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose,

T. D. (1993). Toolglass and Magic Lenses: The See-through Interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pages 73–80, New York, NY, USA. Association for Computing Machinery. event-place: Anaheim, CA.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–72.

Fischer, J., Bartz, D., and Straßer, W. (2004). Occlusion handling for medical augmented reality using a volumetric phantom model. In *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '04*, page 174, Hong Kong. ACM Press.

Gimeno, J. (2018). Addressing the Occlusion Problem in Augmented Reality Environments with Phantom Hollow Objects. page 4.

Jones, B. R., Benko, H., Ofek, E., and Wilson, A. D. (2015). IllumiRoom: immersive experiences beyond the TV screen. *Communications of the ACM*, 58(6):93–100.

Kang, S. B. and Dinh, H. Q. (1999). Multi-layered image-based rendering. In *Graphics Interface*, volume 1, pages 2–13. Citeseer. Issue: 7.

Kawakita, H. and Nakagawa, T. (2014). Augmented TV: An augmented reality system for TV programs beyond the TV screen. In *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, pages 955–960, Marrakech, Morocco. IEEE.

Kritzler, M., Funk, M., Michahelles, F., and Rohde, W. (2017). The Virtual Twin: Controlling Smart Factories Using a Spatially-correct Augmented Reality Representation. In *Proceedings of the Seventh International Conference on the Internet of Things*, IoT '17, pages 38:1–38:2, New York, NY, USA. ACM.

Kurz, D., Kiyokawa, K., and Takemura, H. (2008). Mutual occlusions on table-top displays in mixed reality applications. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology - VRST '08*, page 227, Bordeaux, France. ACM Press.

MacWilliams, A., Sandor, C., Wagner, M., Bauer, M., Klinker, G., and Bruegge, B. (2003). Herding Sheep: Live System Development for Distributed Augmented Reality. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, pages 123–, Washington, DC, USA. IEEE Computer Society.

Nasim, K. and Kim, Y. J. (2016). Physics-based Interactive Virtual Grasping. In *Proceedings of HCI Korea*, HCIK '16, pages 114–120, South Korea. Hanbit Media, Inc.

Nishimoto, A. and Johnson, A. E. (2019). Extending Virtual Reality Display Wall Environments Using Augmented Reality. In *Symposium on Spatial User Interaction*, pages 1–5, New Orleans LA USA. ACM.

Normand, E. and McGuffin, M. J. (2018). Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display). In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 123–133, Munich, Germany.

Raskar, R. and Low, K.-L. (2001). Interacting with Spatially Augmented Reality. In *Proceedings of the 1st International Conference on Computer Graphics, Virtual Reality and Visualisation*, AFRIGRAPH '01, pages 101–108, New York, NY, USA. Association for Computing Machinery. event-place: Camps Bay, Cape Town, South Africa.

Roo, J. S. and Hachet, M. (2017). One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*, pages 787–795, Québec City, QC, Canada. ACM Press.

Saeghe, P., Clinch, S., Weir, B., Glancy, M., Vinayagamoorthy, V., Pattinson, O., Pettifer, S. R., and Stevens, R. (2019). Augmenting Television With Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video - TVX '19*, pages 255–261, Salford (Manchester), United Kingdom. ACM Press.

Sandor, C., Wagner, M., MacWilliams, A., Bauer, M., and Klinker, G. (2002). SHEEP: The Shared Environment Entertainment Pasture. page 1.

Sekuler, A. B. and Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General*, 121(1):95. Publisher: American Psychological Association.

Walton, D. R. and Steed, A. (2017). Accurate real-time occlusion for mixed reality. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, Gothenburg Sweden. ACM.

Wloka, M. M. and Anderson, B. G. (1995). Resolving occlusion in augmented reality. In *Proceedings of the 1995 symposium on Interactive 3D graphics - SI3D '95*, pages 5–12, Monterey, California, United States. ACM Press.