The VR Motion Tracker: Visualising Movement of Non-Participants in Desktop Virtual Reality Experiences

Adalberto L. Simeone*
School of Creative Technologies
University of Portsmouth

Figure 1: A close-up of the "M314 Motion Tracker", as shown in the game "Alien Isolation" [3]. First shown in the "Aliens" film [13]. It is used by the "US Colonial Marines" to track the movement of alien Xenomorphs and display their positions through blips on its screen, with respect to the tracker's origin.
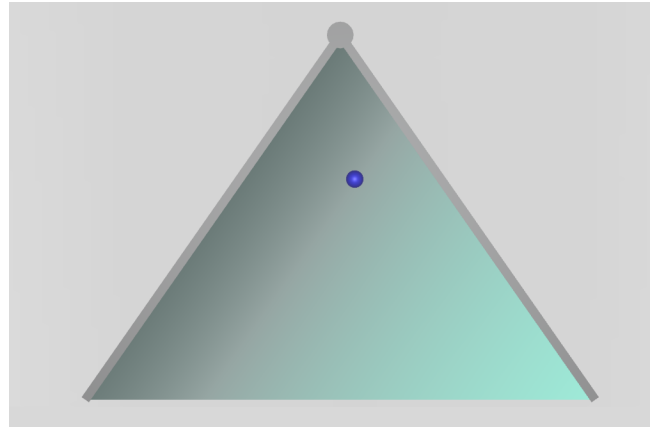

Figure 2: The figure shows a top-down representation of the VR widget. It represents the Kinect's horizontal FOV of $70°$. The blue sphere represents the position of a detected person relative to the Kinect's location (represented as the top vertex of the triangle).

ABSTRACT

In this paper we present the *VR Motion Tracker*: a widget that informs users of VR applications of the movements of non-participants. The design of the widget is inspired by the the motion tracker used in the Alien film franchise. It uses a Kinect to detect other people in the room, besides the user of the VR application. Our system maps this information to a sphere placed within a triangular plane representing the Kinect's field of view. When these non-participants move, the position of the sphere is updated, allowing the user to be aware of nearby movements.

We performed a preliminary study where we presented nine participants with our widget design. We found that they considered the widget to be useful and not distracting. We discuss which features they found interesting, and other information and features they considered useful for a future version of this widget.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Artificial, augmented and virtual realities.

1 INTRODUCTION

Virtual Reality (VR) experiences such as games or applications tend to isolate users from the real world [6]. While wearing a Head-Mounted Display (HMD), the user's vision is focused on the artificial imagery rendered by the VR application. Furthermore, if a user is wearing headphones they might become unaware of any real-world environmental sound.

---

*e-mail: adalberto.simeone@port.ac.uk

In some circumstances, having elements from the real world interfering with the virtual imagery might be undesirable. For example, in a study by Jones et al. [4], a "veil" consisting of a black cloth and plastic cap was used to shield participants from viewing parts of the real world at the periphery of their vision. However, in other circumstances, participants might want to be aware of events happening in the environment they are in [6]. For example, if there are other persons around them, especially if those non-participants want to interact with the VR user in some way.

In this paper we present preliminary work on the design of an in-experience VR widget visualising non-participant positional data detected through a Microsoft Kinect. We were inspired by the design of the "M314 Motion Tracker" as shown in the *Aliens* film [13] and games (Figure 1). We adapted the metaphor of displaying the position of the alien Xenomorphs through blips on the screen to our context. We track the position of other persons who might be walking around in the vicinity of the user. In this way we can display this information within the user's view while they are immersed in a desktop-VR application, to raise their awareness of external events.

We designed a study where participants were asked to play a desktop-VR game while paying attention to the information provided by the VR widget. The game consisted in flying a spaceship through a series of gates. Its flying direction was associated to the head-orientation of the user wearing an Oculus Rift. After completing a round of the game, we asked participants to fill a short questionnaire. Finally, we interviewed them to elicit feedback on the design of the widget.

The preliminary results we found indicate that participants considered such functionality to be useful. From the interviews, we extracted a set of desirable features that participants would like to see in a future iteration of this widget, such as sound detection or

inferring contextual information. We then present a discussion of these potential directions and of the technical considerations associated to the design of an improved version of the widget.

2 RELATED WORK

The commercial availability of HMDs has sparked a renewed interest in research issues focusing on the mixture of reality and VR [1, 9]. In this paper we narrow our focus to those *Mixed Reality* works leaning towards the "Augmented Virtuality" end of Milgram's continuum [8]. That is, works that use information from the real world to augment the expressivity of a Virtual Environment.

In previous works [9, 10] we presented the concept of *Substitutional Reality* in VR, where elements of the virtual world are paired to physical objects in the real world, albeit with a degree of mismatch. This use of passive haptics supports the believability of the experience.

McGill et al. [6] presented more explicit ways of bringing elements of the real world into Virtual Environments. In their study, they evaluated various interaction techniques focusing on three main scenarios happening while users are immersed in VR: 1) typing on a keyboard; 2) interacting with objects at arms reach; 3) becoming aware of other persons in the same room. In particular, the last interaction scenario is the one most relevant to the work presented here. Their implementation also uses a Kinect to detect users. However, differently from our widget, they extract a "cut-out" of the detected person and super-impose it within the VR experience. They tested two conditions: one where detected persons are blended within the virtual imagery, and another where these cut-outs are presented fully opaque. In addition, the Kinect was placed in such a way to detect the space around the user. In our work we wanted to focus on domestic scenarios where users are interacting with a desktop PC located near a wall. We thus placed the Kinect beneath the monitor, in order to detect the space behind the user (Figure 3). Their results show that 58% of their participants wanted this information to be communicated in a different way.

Other works have focused on methods to visualise one's hands or legs into the Virtual Environment. For example Steinicke et al. [11] have developed a method for visualising the user's body into the Virtual Environment by means of chroma-keying. Similarly, Tecchia et al. [12] use a depth camera attached to the HMD to re-construct the user's own hands. Lee et al. [5] presented a "cross-dimensional" interface where users interact with a Virtual Environment which contains its own Virtual Environment, displayed through a Virtual PC. Chittaro et al. [2] presented Vu-flow, a system that allows a user to analyse user movement in Virtual Environments. However, this tool is aimed at improving the design of the Virtual Environment and to support its analysis, rather than to provide a real-time in-experience visualisation of movements happening in the real world.

3 APPARATUS

The system consists of a Microsoft Kinect 2.0 connected to a PC. The Kinect sensor is ideally placed beneath the centre of the monitor frame (Figure 3). The Kinect was tilted upwards, in order to maximise the coverage area.

We built a prototype application in Unity 5.3. We used the Microsoft Kinect plugin to interface Unity with the sensors it provides. The Kinect can recognise up to six people and track two, within a horizontal field-of-view (FOV) of $70°$ [7]. From the data provided by the Kinect depth sensor, it is possible to extract a distance in millimetres from the camera plane to the nearest object detected at that particular *x,y* coordinate in the depth frame image.

In our prototype, we used this distance data to implement our "VR Motion Tracker" widget. We designed a 3D visualization of the Kinect's FOV, as shown in Figure 2. It consists of a triangular green plane depicting an arc of $70°$, originating from a cylinder
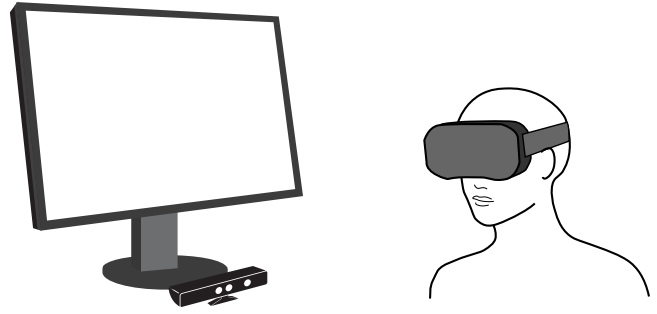

Figure 3: Our "VR Motion Tracker" uses a Kinect 2.0 sensor to detect the movements of persons who might step in its field of view. This information is used to provide a user who is immersed in a desktop-VR experience with a widget tracking these movements.

representing the sensor's origin. The length of the radius represents the maximum detection distance of 4.5 m. Locations within the area of this sector are normalised according to this distance.

As soon as a person enters the Kinect's FOV, it will start tracking its skeleton. We used the root joint ("SpineBase") to extract the *z* distance and derive the *x-y* distances using trigonometry. This information is used to place a blue 3D sphere within the widget's area. The position of the blue sphere is updated at 30 fps, to reflect the movements of the detected person.

Users were wearing an Oculus Rift DK2. Our prototype implementation tracks only one person.

4 PRELIMINARY USER STUDY

We wanted to elicit user feedback on the concept behind this widget: whether they thought it could be useful and how to improve it. To this end, we designed a preliminary user study in which users are immersed in a desktop-VR game where, while playing, the widget provides updated information on the whereabouts of a person who might walk within the Kinect's FOV.

4.1 Task

The application is built on top of a VR game sample from the Unity 5 engine where players have to fly a spaceship within a series of gates (Figure 4). They control the direction the ship is flying towards through head movements. We altered this sample in order to place our VR widget in the game and update it accordingly to the information extracted from the Kinect data.

The widget is rendered in such a way to always be in the same position in the screen. Since it is created through 3D solid shapes, whenever the user rotates their head, the widget rotates as well. In this way, the relative position of any detected person will reflect the rotated viewpoint.

4.2 Participants

We recruited nine participants (7 male, 2 female; $M = 27.89$ years, $SD = 9.39$). We described the purpose of the study to each participant, the functionality of the VR widget, and the goal of the game itself. After playing a round of the game, we asked participants to fill in a questionnaire.

4.3 Procedure

We asked the participants to rate the widget visualising movements of non-participants on two factors: its *usefulness* and its perceived *impact* on game performance. Scores were expressed on a scale from 1 (very low) to 7 (very high). Finally we interviewed participants following a semi-structured script.
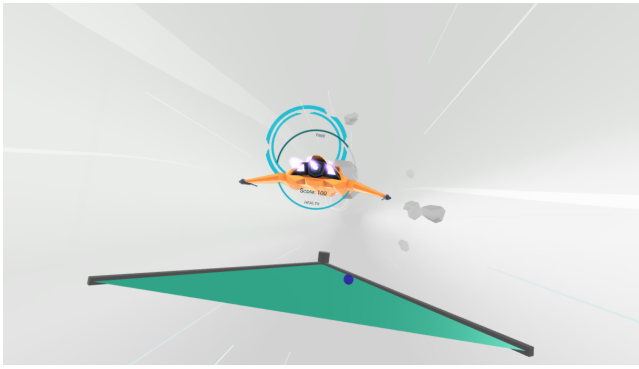
Figure 4: An in-game representation of the widget. It is always rendered in the lower-third of the user's view within the Oculus Rift.
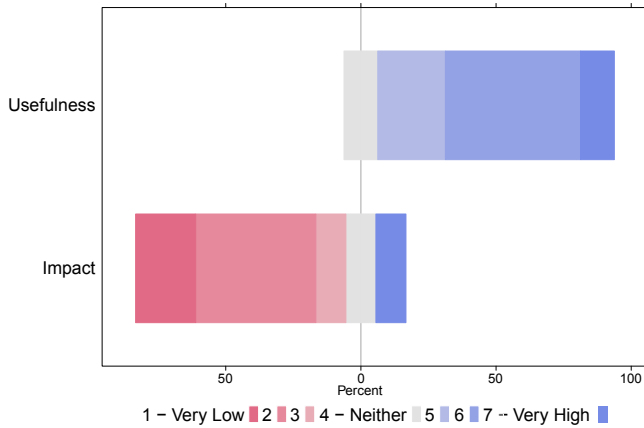

Figure 5: The chart displays the distribution of qualitative scores on the two measures of *usefulness* of the widget and its perceived *impact* on game performance.

We asked three main questions: 1) what kind of information they would like this widget to portray, in addition to the movements of non-participants; 2) whether the widget should work in an always-on manner, on demand, or in any other way; 3) how clearly the design of the widget communicated this information.

5 RESULTS AND DISCUSSION

As shown in Figure 5 participants rated the widget consistently as useful with a mean score of 5.78 ($SD = 0.97$). They also considered it to have a low impact on their overall performance in the game ($M = 2.78, SD = 1.79$).

From the individual interviews, we extracted a set of features that users considered desirable for future iterations of this widget.

5.1 Information Visualization

When asked what other kind of information they would want to see represented, sound was suggested by three users as potentially useful. For example, an ideal system should be able to detect sounds happening in the environment. Estimating their location with respect to the user would be helpful. Further, participants suggested that information about the nature of the event should be conveyed to the user. For example, discerning between a noise and a phone ringing, or detecting when the user's name is being called.

Two other participants suggested that detecting non-verbal communication was also important. For example, a non-participant who

is unfamiliar with head-mounted displays might attempt to wave or gesture at users immersed in the VR experience to get their attention.

Representing visual information about the physical surroundings was deemed useful by two other participant. An overhead mini-map, as commonly shown in some games, could help users avoid collisions with obstacles or walls. They were concerned about these issues in conjunction with VR systems allowing some degree of mobility in the user's immediate surroundings.

5.2 Frequency of Updates

In our prototype the widget was always rendered in the lower part of the screen. Although the data indicates that this did not seem to affect game performance, seven participants agreed that the information should be displayed only when an event is detected. Two others were happy to imagine the widget always rendered if not intrusive.

One user suggested that the system should be able to infer the priority of the event. For example, it could prioritise events happening in the close proximity of the user to the point of pausing the VR experience in cases of extreme severity.

5.3 Design of the Widget

We asked participants to comment on the design of the widget. Six of our nine participants appreciated the widget as it was displayed, with two others indicating they would have preferred just a 2D overlay (our widget, although always rendered in the same area, was composed 3D primitives and followed the head's orientation). One participant stated that toast notifications or haptic feedback would have been sufficient. The others did appreciate the positional awareness the 3D widget provided.

Two participants commented that representing a widget using designs reminiscent of mini-maps used in games could lead to confusion in those scenarios where the VR experience makes use of a similar graphical widget to represent information about its own Virtual Environment. For example, a VR flight simulator displaying a radar-like representation of other ships in the vicinity, or a first-person game displaying a mini-map showing other detected characters or points of interest. They felt that in those situations the VR application should use different styles to convey which information belongs to the virtual world and which belongs to the real world. This could be further supported by a tutorial phase where the VR application explains the purpose of the widget.

5.4 Technical Considerations

In our prototype we experimented with a single Kinect placed beneath the user's screen. The Kinect was facing towards the user. This arrangement is particularly suited in those scenarios where a desk is facing the walls of the room, with the user's back facing the rest of the room.

Although this particular configuration tends to incur into some occlusion due to the user blocking part of the view of the Kinect. In our prototype, we placed the Kinect beneath the monitor and tilted it upwards in order to increase the chance of capturing walking persons and minimise the occluded part. It would also be possible to offset the Kinect towards the side most likely other people might be approaching from. For example, if the desk is positioned facing one of the room's corners. The Kinect could also be mounted on top of the monitor frame.

An ideal configuration would conceivably use a $360°$ detection range. This could be achieved either by multiple Kinects or by embedded depth sensors circling the HMD. This last solution would also overcome occlusion issues. Finally, audio hardware necessary for sensing 3D sound location would be necessary to implement the related features previously described.

6 CONCLUSION

In this paper, we presented preliminary work on the design of a *motion tracker* for Virtual Reality applications. It uses a widget to visualise the movements of those who are not participating in the main user's Virtual Reality experience. We presented this widget to nine users and collected their feedback. Users found it to be very useful to raise awareness of events that might happen in their vicinity. We extracted a set of desirable features for future iterations of this widget, such as detecting sound events and inferring contextual information.

REFERENCES

[1] D. Boland and M. McGill. Lost in the rift: engaging with mixed reality. *XRDS: Crossroads, The ACM Magazine for Students*, 22(1):40–45, 2015.

[2] L. Chittaro, R. Ranon, and L. Ieronutti. Vu-flow: A visualization tool for analyzing navigation in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1475–1485, 2006.

[3] Creative Assembly. Alien Isolation. 2014.

[4] J. A. Jones, J. E. Swan, and M. Bolas. Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):701–710, 2013.

[5] P.-W. Lee, H.-Y. Wang, Y.-C. Tung, J.-W. Lin, and A. Valstar. Transection: Hand-based interaction for playing a game within a virtual reality game. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 73–76. ACM, 2015.

[6] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A dose of reality: Overcoming usability challenges in VR head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2143–2152. ACM, 2015.

[7] Microsoft. Kinect 2.0 human interface guidelines. In *http://download.microsoft.com/download/6/7/6/676611B4-1982-47A4-A42E-4CF84E1095A8/KinectHIG.2.0.pdf*, Last accessed on 2016-2-11.

[8] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12):1321–1329, 1994.

[9] A. L. Simeone, E. Velloso, and H. Gellersen. Substitutional Reality: Using the physical environment to design virtual reality experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 3307–3316. ACM, 2015.

[10] A. L. Simeone. Substitutional Reality: Towards a research agenda. In *Proceedings of the 1st Workshop on Everyday Virtual Reality (WEVR)*, pages 19–22. IEEE, 2015.

[11] F. Steinicke, G. Bruder, K. Rothaus, and K. Hinrichs. Poster: A virtual body for augmented virtuality by chroma-keying of egocentric videos. In *Proceedings of the IEEE Symposium on 3D User Interfaces*, 3DUI 2009, pages 125–126. IEEE, 2009.

[12] F. Tecchia, G. Avveduto, R. Brondi, M. Carrozzino, M. Bergamasco, and L. Alem. I'm in VR!: Using your own hands in a fully immersive mr system. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, VRST '14, pages 73–76. ACM, 2014.

[13] Twentieth Century Fox Film Corporation. Aliens. 1986.