The Augmented Reality Space Continuum

Thuong N. Hoang[1], Ross T. Smith[2], and Bruce H. Thomas[3]
Wearable Computer Lab – University of South Australia

ABSTRACT

We propose a presentation space classification for augmented reality as a continuum to guide challenges such as object placement, transitioning between presentation regions, and first person perspective effectiveness. The continuum starts from the first person perspective and expands outwards, into four regions: *head, body, distance,* and *remote*. The *head region* is the small area attached to the user's head, while the *body region* contains the spatial area within the user's arm reach. The *distance region* is out of the user's arm reach including everything within the user's vision, while *remote region* covers the rest of the augmented world. The continuum provides a classification that aims to offer useful insights into aspects of AR systems, such as how to transition between presentation spaces effectively. Based on the classification, we developed a 2D continuum for video-based AR to discuss the variations of the technology and how the user's sense of first person perspective is altered along the continuum.

Keywords: AR space continuum, classification of AR space.

Index Terms: H.5.0 [Information Interfaces and Presentation]: General; I.3.6 [Computer Graphics]: Methodology and Techniques.

1 INTRODUCTION

We propose a classification for the presentation space for Augmented Reality (AR) systems, as a continuum of object placement. Unlike the traditional desktop-based computing where the information is restricted by the physical screen, augmented reality systems merge a virtual world directly into the physical surrounding of the user. The augmented world opens up new dimensions for information presentation. We propose a classification of the presentation space for augmented reality system, in a form of a continuum. The classification aims to identify different properties of the presentation regions in the continuum and to offer useful insights into various aspects of AR systems, including classification, display types, interaction, registration, tracking, and collaboration. We also apply the continuum to the physical placements of camera and display for video based AR providing a 2D continuum to explore the different variations of the technology.

Our continuum starts from the first person perspective and expands outwards. The continuum is divided into four regions: *head, body, distance,* and *remote*, see Figure 1:

1. *Head region:* is a small area that is attached to the user's head orientation and located at the near plane of the user's vision. Information displayed in this area is placed closest to the user's eye, but not passing the minimal focusable distance. The main characteristic of this region is the constant visibility of information to the user, because it is always in the line of sight of the user's eyes.
2. *Body region*: is the spatial area surrounding the user that is within the user's arm's reach. There is no fixed boundary for this region as it differs from user to user.
3. *Distance region*: is the spatial area surrounding the user that is out of arm's reach. This region expands to the end of the user's line of vision.
4. *Remote region*: is the spatial area of the augmented world that is not visible from the user's location. This region could be blocked from the user's sight by physical obstruction (walls, buildings, etc.) or by being remotely located.

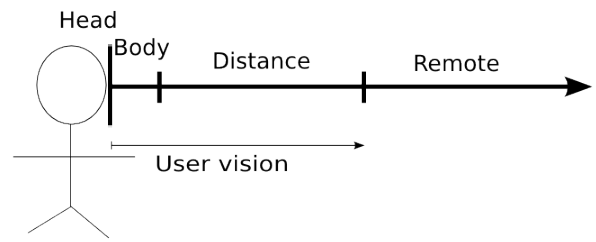

Figure 1: The AR space continuum

The four regions in the continuum cover all the presentation space of an augmented reality world. The next section provides a discussion of various aspects of AR research, with respect to our proposed continuum, which we called the *AR space continuum*.

2 THE AR SPACE CONTINUUM

A similar space segmentation concept is discussed by Cutting and Vishton [1], with regards to the human perception of layout and distance in the surrounding. Their classification divides into *personal space (*approximately within 2m of the user), *action space (*approximately up to 30m), and *vista space (*beyond 30m). Our continuum extends the *personal space* into the *head* and *body regions*. The *distance region* matches with their *action space* and extends into the *vista space*. Our *remote region* extends beyond their *vista space* to include areas that are not visible within the user's vision.

The benefit of an augmented reality system is that it can display information in the *head* and *remote regions* to the user. In everyday life, it is uncommon for a person to have any information attached directly to and constantly visible in their field of vision (*head region)*. Due to visual occlusion, a person cannot see through walls or a scene where they are not currently located (*remote region)*.

AR Classification

The major types of AR technologies have different placements of the display on the AR space continuum. The term 'display' may refer to the physical screen or in the case of projector-based AR, to the actual projection. Head-mounted display (HMD) is an immersive AR branch with hardware located in the *head region,* while mobile phone based AR allows the user to move the display within the *body region*. Spatial AR uses projector to augment the physical world around the user, within the *body* and *distance*

1 thuong.hoang@unisa.edu.au
2 ross.smith@unisa.edu.au
3 bruce.thomas@unisa.edu.au

regions. The *distance region* also accommodates another branch of AR that we refer to as *billboard AR,* where a camera is mounted on top of a display facing the user. The camera feed is then displayed on the screen, and the user can see themselves surrounded by augmented virtual objects, such as in the game Eye Pet by Sony[1]. The display is typically located out of the user's arm reach. The *remote region* is not located within the user's vision; therefore, there is no AR technology to fit into this region. There are also other types of AR display that extends into multiple regions. Head-mounted laser-projectors [2] can be placed into the *body region* or *the distance region,* since the projected lights are seen by the user to be in these two regions. The classification is concerned about the physical location of the augmented images, which can be different from the location of the actual display devices (such as wall mounted or head-mounted projectors).

Display

Various display technologies for AR have the capacity to present information in different regions of the continuum. For this discussion we consider the *perceived* location of the graphics by the user. HMD can display information from all regions, from *head* to *remote*. An example of the display of virtual objects from the *remote region* through the user's HMD is the augmented viewport technique [3], where the feeds from a remotely located camera is shown to the user through a viewport window. With the augmented viewport, the user is aware that the content of the viewport window is from a remote location; therefore, it is considered in the *remote region*. Mobile AR is used for virtual information that is located within the *body, distance,* or *remote regions*. Projector based spatial AR spans from the *body* to *remote regions*. An example work in Spatial AR is the AR X-Ray project by de Miranda et al. [4]. A portable projector is used to project virtual models of internal pipes onto the wall, to create an effect of x-ray vision through the wall. In this project, the projector and the projected light are still located within the *body* and *distance region,* even though the projected imagery shows the pipes that are not visible (i.e., from the *remote region*). Similar observation can be made regarding the work by Bimber and Iwai [5], where a projector is used to augment other media such as photographs, electronic paper display, and x-ray films. As can be seen from this classification, HMD can show information from all regions in AR space continuum, while projectors and mobile devices cover the range from *body* to *remote regions*.

Registration and Tracking

Registration and tracking are two major research areas in the AR community. The problems of registration and tracking vary depending on the region of concern in the AR space continuum.

In the *head region,* registration and tracking of nearby objects or the surroundings is not required. It is a trivial task to display information attached to the user's virtual screen, and there is generally no physical objects near the user's head to be tracked. Eye tracking [6] is an exception in this region, where the tracking of the eyes is generally used as an input mechanism to the system. *Body region,* however, attracts considerable research effort. For example, there are several research works into hand tracking for interaction, such as [7-9]. Tracking and registration for the *distance region* is more challenging. Approaches to the problem include GPS, marker-based [10], SLAM [11], natural features, and sensor-fusion [12]. *Remote region* is generally not considered because it is out of view. Further regions pose more challenges to the problem of registration and tracking due to the growing tracking volume.

[1] http://www.eyepet.com/

Interaction

From the perspective of the AR space continuum, interaction techniques for augmented reality systems are generally represented as *transitions* between the regions. The further away from the user, the more challenging the problems of registration and tracking are. The challenge is due to the larger volumes of tracking. There is a tendency to move virtual objects from outer regions closer to the user to enhance interaction. The image plane technique [13] is an example transition of virtual objects in the *distance region* to the *head region*. Virtual objects are projected onto the user's image plane for interaction. The granularity of the interaction is reduced through the transition process, i.e. small movements in the *head region* translate to larger displacement of the virtual objects in the *distance region*. Ray casting or cursor-based techniques employ a reverse approach, where a cursor or a point in the *head region* is projected outwards to perform selection and transformation in the outer regions.

There is a transition region just into the *distance* region of the continuum, which we refer to as *just-out-of-arms-reach*. Just-out-of-arms reach is the region in which the user has the option of either performing some action at a distance interaction or physically moving to allow some form of direct manipulation. Additionally, finer motor skills interactions may be developed for interactions in the just-out-of-arms-reach region.

Objects placed in the closer region to the user have higher visibility and cast shadows onto the further region. In other words, closer regions occlude further ones. Therefore, only important information should be placed closer to the user. *Head region* is often the most popular choice to display system UI widgets, which are often required to be constantly accessible to the user. *Body region* can be utilized for proprioception to enhance the UI. Studierstube [14] is a pen-and-pad UI that is placed entirely in the *body region*. Walk-in menu [15] where the user steps on a virtual panel as a menu item is another example of a UI element in this region. Objects placed in the *distance region* are often considered as part of the AR environment, instead of being part of the UI, because this region does not offer the immediate access to the user.

World-in-miniature (WIM) [16] has a further transition that brings both *distance* and *remote regions* to the *body region,* as a miniature model on the user's hand. X-Ray vision is a visualization technique [17] that bridges the border between the *distance* and *remote regions*. Areas that are not visible due to an obstruction of a wall or building are in the *remote region,* and the x-ray vision that allows the user to visualize the occluded area brings it into the *distance region*.

The augmented viewport technique [3] is an interaction technique that brings the *remote region* into the *head* or *body regions*. The technique brings video feeds from remotely located camera to display as a virtual window, either attached to the user's viewing plane or located near the user's location.

In a slightly different context of virtual reality system, the transition of regions does not apply only to the virtual objects, but also to the virtual representation of the user. Arm extending or teleport techniques, such as the Go-go arm [18] bring the user's virtual hand from the *body region* outwards to perform interaction with virtual objects in the *distance* region and vise versa.

Collaboration

In a collaborative work context, the different regions in the AR space continuum have different collaborative values. The *head region* should be almost exclusively used for private information because it is challenging for information displayed in this region to be shared. However, the *head region* of a participant can be used to display identification information for other participants to view.

Body region can be used for both shared and personal space. Information displayed in this region still brings the sense of belonging to the user, but also can be seen by other participants. *Distance region* is the shared space where common information can be displayed. *Remote region* can be transitioned into the *distance region* using techniques such as x-ray vision to support collaborative evaluation tasks. The mapping of the different regions for collaborative work matches with the classification of personal space research in cognitive science [19] and sociology [20] in that personal space is matched to closer regions to the user.

3 The 2D Continuum for video-based AR

The AR space continuum can also be applied to physical objects. In the case of video based AR, positions of the camera and the display may be at different positions on the continuum in different physical incarnations of the technology. In the most common case, the camera is rigidly attached to the display as close to the viewing axis of the display as possible, in other words, they belong to the same region of the AR space continuum. This leads to two independent parameters that determine the quality of the provided first person perspective presentation of the visual information: the position of the camera pose and the physical display. We combine the two versions of the continuum to create a 2D continuum (see Figure 2) to explore the variations of video-based AR as well as to discuss how the user's sense of first person perspective (FPP) is affected.

3.1 Camera pose

As an example, we explore the relation of the viewing direction of a display to an attached camera. This starts with a perfect alignment to the exit pupil in the *head region* to the other extreme of a fixed security camera in the *remote region*. An example of the first position on the continuum is a parallax-free HMD developed by State et al. [21] that aligns the camera with the user's visual axis. This is followed by a very common method of developing video see-through by mounting a camera on the top of the HMD. The slight offset of the camera, still in the *head region,* alters the user's expected view the physical world, thus reducing the feeling of FPP. We now move away from the classical first person perspective to a remote presence mode. The view into this remote location is FPP in that the user's AR view is aligned with camera's view. In the ARVino project, King et al. [22] experimented with placing a camera on a 2m-3m pole that was aligned in the same compass orientation as the display, but tilted slightly down to provide a better view of vegetation 5 to 100 meters away. This allowed the user a more acute viewing angle on GIS data.

We move onto a situation where the position and orientation of the camera are in the *body region* and different from the user's eye, and further reducing the sense of FPP. To the user, the camera's view turns into a third person view of the world, and it does not correspond with what the user can see with their eyes from their current position. Therefore the sense of FPP is gradually reduced. The first of these is a handheld camera connected to an HMD style system. This enables a user to view around a corner, but the augmentation is still first person from the point of view of the camera. The user still has a sense of the location they are occupying, but the direction is not their normal one. An alterative is to place the camera at a fixed location close to the user, mounted on a tripod at eye level with a 2-axis computer controlled platform slaved to user's head. This would emulate a forward observation position. These last two examples depict the independent modification of the two attributes of pose, position and orientation. We place these at the same position of the continuum.

We can explore the *body region* of the continuum by removing the condition of the camera being placed at eye level, but relatively close to the user. We now move towards robotic telepresence. In this case the first of these is a camera mounted on a robot slaved to the user's head. This would reflect changes to the user's head orientation, but the camera's location is static. We move on to allowing the user to "drive" the robot to different locations, while still slaving the orientation to the user's head. This transition enables the robot to extend from the *body region* to the *distance region*. When the robot is controlled by another user, this reduces the sense of FPP. Finally we purpose a fixed camera, such as security camera, in the *remote region*. Although the user has a first person perspective on the remote scene, the sense of presence is the lowest. This is outlined in a study of performance in virtual environment, influenced by the level of immersion by Slater et al. [23]. In the study, they compared two levels of immersion: exocentric screen based (*body distance* and outwards) and egocentric HMD based (*head region)*.

In this continuum we have combined the orientation and position of the camera into one parameter: pose. This excludes situations where the orientation is changed but the position is not. A question arises as to which feature affects the user the most in the AR space continuum based on FPP. As this is a continuum, there is no hard isolation of each region. We are attempting to highlight how there are a number of subtle differences in an array of AR systems. These issues do not arise in virtual reality (VR) systems.

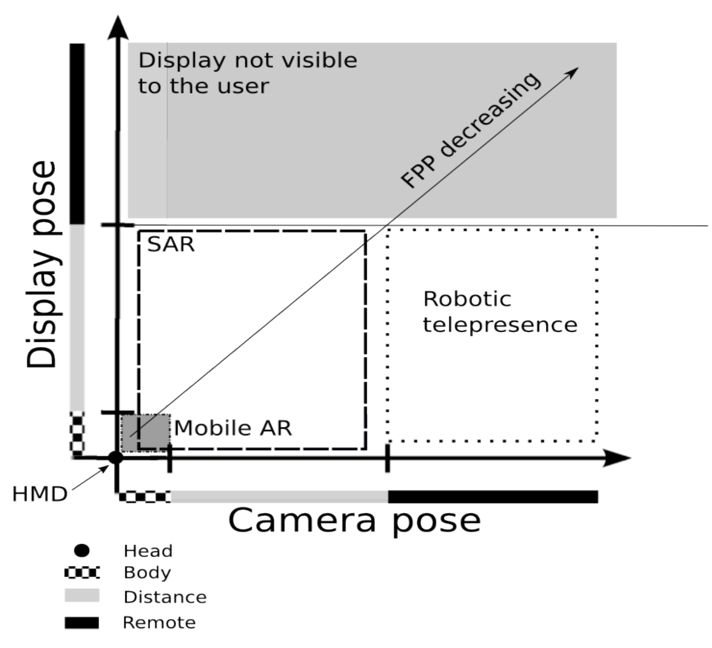

Figure 2: The 2D AR space continuum for video-based AR

3.2 Display pose

The ability to change the display's pose relative to the user's view is very much more limited than with the camera. The most important limitation is that the user must be able to view the display; therefore the *remote region* is omitted in the consideration. What we consider the optimal FPP is 100% alignment of the display with the user's viewing axis that includes the entire field of view for the user. The use of an HMD is a good example of this placement. If you move the display away from the user's eyes and allow other visual information into the user's view such as a handheld display, this reduces the FPP. Rotating the physical display off the viewing axis further reduces this perspective. Proprioception of the display that is worn or held is a strong cue for first person perspective, and this is lessened when the display is free standing. Displays have been mounted onto moveable tripods (Augoscope and ARVino) allowing for larger displays and multiple users to view simultaneously, but reducing the sense of FPP. Fixing the location of the display such as mounting it on a wall further decreases the sense of presence. Multiple projectors in a non-planar configuration such as Spatial AR remove the display's ability to provide FPP.

Figure 2 describes the 2D continuum of display and camera placement for video-based AR. The origin indicates the traditional HMD AR where both the display and camera are located in the *head region*. Mobile AR with the camera attached to the back of the display occupies the area of *body region* for the display and camera pose, because the user is always holding the mobile phone within this region. Projector based spatial AR (SAR), albeit technically not being video based AR, fits in the area that has the camera and display freely located within the *body* and *distance area*. When the camera is remotely located, regardless of the position of the display, it is robotic telepresence AR, as shown on the 2D continuum.

4 Conclusion

We propose a continuum for AR presentation space that is based on the first person perspective. There are four regions in the continuum, namely *head, body, distance,* and *remote*. The four regions have different characteristics and considerations for various issues in AR, including classification, display, registration, tracking, interaction, and user interface. We apply the continuum on the physical placements of the camera and display in video-based AR to create a 2D continuum to explore the different incarnations of the technology and the effects on the user's sense of first person perspective.

References

[1] Cutting, J.E. and P.M. Vishton, Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. Perception of space and motion, 1995. **5**: p. 69-117.

[2] Schwerdtfeger, B., et al. Using laser projectors for augmented reality. 2008. ACM.

[3] Hoang, T.N. and B. Thomas. Augmented Viewport: An action at a distance technique for outdoor AR using distant and zoom lens cameras. in International Symposium on Wearable Computers. 2010. Seoul, South Korea.

[4] de Miranda, F.R., et al., Designing and implementing an Spatial Augmented Reality X-Ray. Revista de Informática Teórica e Aplicada, 2009. **15**(3): p. 47-74.

[5] Bimber, O. and D. Iwai, Superimposing dynamic range, in ACM SIGGRAPH Asia 2008 papers2008, ACM: Singapore. p. 1-8.

[6] Beach, G., et al. Eye tracker system for use with head mounted displays. in Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on. 1998.

[7] Foxlin, E. and M. Harrington. WearTrack: a self-referenced head and hand tracker for wearable computers and portable VR. in Wearable Computers, The Fourth International Symposium on. 2000.

[8] Wang, R.Y. and J. Popovic. Real-time hand-tracking with a color glove. in ACM SIGGRAPH 2009. 2009. New Orleans, Louisiana: ACM.

[9] Piekarski, W., et al. Mobile hand tracking using FPGAs for low powered augmented reality. in Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on. 2004.

[10] Kato, H. and M. Billinghurst. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. in Proceedings of 2nd IEEE and ACM International Workshop on Augmented Reality, IWAR99. 1999.

[11] Klein, G. and D. Murray. Parallel Tracking and Mapping for Small AR Workspaces. in Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on. 2007.

[12] Schall, G., et al. Global pose estimation using multi-sensor fusion for outdoor Augmented Reality. in Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on. 2009.

[13] Pierce, J.S., et al., Image plane interaction techniques in 3D immersive environments. Proceedings of the 1997 symposium on Interactive 3D graphics, 1997.

[14] Szalavári, Z., et al., Studierstube: An environment for collaboration in augmented reality. Virtual Reality, 1998. **3**(1): p. 37-48.

[15] Hoang, T.N., et al. Web 2.0 meets Augmented Reality. in 13th International Symposium on Wearable Computers. 2009. Linz, Austria.

[16] Bell, B., T. Hollerer, and S. Feiner, An annotated situation-awareness aid for augmented reality, in Proceedings of the 15th annual ACM symposium on User interface software and technology2002, ACM: Paris, France.

[17] Avery, B., W. Piekarski, and B.H. Thomas. Visualizing Occluded Physical Objects in Unfamiliar Outdoor Augmented Reality Environments. in Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on. 2007.

[18] Poupyrev, I., et al., The go-go interaction technique: non-linear mapping for direct manipulation in VR. Proceedings of the 9th annual ACM symposium on User interface software and technology, 1996: p. 79-80.

[19] Berti, A. and F. Frassinetti, When far becomes near: Remapping of space by tool use. Journal of cognitive neuroscience, 2000. **12**(3): p. 415-420.

[20] Sommer, R., Studies in Personal Space. Sociometry, 1959. **22**(3): p. 247-260.

[21] State, A., K.P. Keller, and H. Fuchs, Simulation-Based Design and Rapid Prototyping of a Parallax-Free, Orthoscopic Video See-Through Head-Mounted Display, in Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality2005, IEEE Computer Society. p. 28-31.

[22] King, G.R., W. Piekarski, and B.H. Thomas. ARVino - outdoor augmented reality visualisation of viticulture GIS data. in Mixed and Augmented Reality, 2005. Proceedings. Fourth IEEE and ACM International Symposium on. 2005.

[23] Slater, M., et al. Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess. 1996