Virtual Architecture in a Real-time, Interactive, Augmented Reality Environment

Project Anywhere and the Potential of Architecture in the Age of the Virtual

Constantinos Miltiadis[1]
[1]*Chair for Computer Aided Architectural Design [CAAD], ETH Zurich*
[1]*http://www.studioany.com*
[1]*c.miltiadis@gmail.com*

Virtual reality opens a new world of a great potential for both research and experimentation by allowing new forms of unbuilt sensible architectural space. The paper starts with a sketch of the current context in Virtual Reality and continues by outlining the development and structure of the research "project Anywhere". The project is an easily deployable, wireless, multi-user, augmented reality app-system that offers full body immersion through body, head and hands tracking. It can host multiple concurrent users, able to move freely in the virtual space, by moving in the real, and also perform actions through a gesture interface to affect their shared environment. Concluding, we describe the inherent properties of such a space, which we propose as a novel spatial medium for architecture, through an example of a potential application.

Keywords: *Augmented Reality, Full-body, Mobility, Immersivity, Multiplayer*

INTRODUCTION AND RELATED APPLICATIONS

The road to virtual reality technology

Since the 1980s, multiple efforts where made in order to develop and bring Virtual Reality (VR) technology to the masses. However, only in the last 2 years, one can truly admit that the technology enabling VR has been advanced to such an extend that renders its implementation both viable and worthwhile. The field that this paper is concerned about is VR through the implementation of Head-Mounted Displays (HMD), in which various companies have recently produced and marketed either tools or devices, that offer to users the experience of stereoscopic real-time animated view. The function of such devices is to render 2 digital images at a time -one for each eye- which, with the display worn in front of the users' eyes, create the illusion of depth and hence stereoscopic view (Sutherland, 1968) in an interactive manner in which any tilting of the head corresponds and affects the virtual point of view in real-time.

Paramount to the application of HMD VR was computer performance (Brooks and Frederick 1999). Because the premise and basis of HMD VR technology is essentially tricking the brain to experience or perceive something virtual (not real), the major obstacles that delayed its development, were maintaining a high refresh rate -greater than 60-75 Hz- and

-imperceptible- low latency; a combination of hardware and software that renders graphics in a very high frequency and that responds very quickly to the user's head rotation. Implementations lacking or underperforming in any of these two parameters result in causing nausea, which is in this case called simulator sickness (Moss and Muth, 2011).

Related technologies and setups

HMDs are mask-like devices, the basic components of which can be broken down to a screen -split in two- for display, an Inertia Measuring Unit (IMU) that calculates the orientation of the user's head, and lenses that on the one hand allow the user to focus on the screen and on the other perform the necessary optical corrections of the displayed graphics. At this moment, HMD VR applications are based on 2 different type of setups.

The first type, is essentially an electronic device that uses the processing and rendering capabilities of a computer to which is connected via HDMI and USB. It is worth mentioning that the most prominent company providing HMD VR technology today, Oculus VR[1], was up to now only offering products for developers (labeled as Developer Kits - "DK1", "DK2") and just recently announced its first consumer product[2] which will reach customers by 2016.

The second type of setup, which is gradually gaining ground, is built for deployment on smartphones and mobile devices in general. Requiring only a passive device mount which makes the device wearable -also including lenses- it depends solely on on-board hardware to process and render, while using the device's screen as the actual display. Offered devices include initiatives like the Google Cardboard[3], Samsung Gear VR[4] (developed with Oculus VR) and Zeiss VR ONE[5].

Both cases involve hardware as well as software, and is safe to say that although functional, they are still in an experimental stage. While the former setup type is clearly the winner in both performance and quality -given that it is a piece of purpose specific hardware that utilises the processing power of a computer- the latter case, which the project used in its implementations, has clearly some advantages to offer. Besides the ease of deployment, given the ubiquity of mobile devices today, it can function without the need of any extra devices besides passive mounts. In addition, in our case, throughout the development of the project, it proved to perform satisfactory, under certain limitations concerning rendering and calculation load.

State of the art in VR implementation

VR has been implemented for years, mostly in the field of the Cognitive Sciences, in the form of CAVE labs (Cruz-Neira et al., 1993), which create a virtual environment through a digital theatre setup. However, applications using and developed for contemporary HMD VR devices are scarce at this moment, as most of the available content regards demo experiences. In the field of research some CAVE labs replaced projection with the implementation of HMDs like the Cyberneum[6] of the Max Planck Institute for Biological Cybernetics. Also in the field of entertainment, a prominent example is the "Birdly"[7] installation, developed at the Interaction Design Department of ZhDK, which immerses users in the mechanics and sensations of a bird flight. VR initiatives, still in the development stage worth mentioning, include the "HTC VIVE"[8] system, developed by Valve and HTC which offers head and hands tracking in a given space using MEMS sensors, and finally the large scale installation "The VOID"[9] which overlaps a real with a virtual indoors built environment, in what they call "Virtual Entertainment Centres", for the purpose of immersive gaming.

PROJECT ANYWHERE

Research goals

project Anywhere was initiated in May 2014 and was developed by the author at the Chair of CAAD ETH Zurich. Experiencing an immersive environment with the Oculus Rift, motivated us to imagine scenarios of VR applications in architecture, beyond the one of visualisation. Our research goal was to attempt a "viable" virtual space, by enhancing both the immersive experience and the nature of it's space. Consequently, the three main pillars we set were: to include more properties from the human spatial presence in the virtual environment -rather than only vision and head movement- in order to create a greater degree of immersion; to invent intuitive ways of interfacing and performing actions in the Virtual Space (VS); and eventually enable multiple users to co-exist simultaneously in a common, shared VS. In other words, to increase the degree of cognitive self awareness in the virtual world, allow for the user to affect space, and add a social aspect to it. Eventually, our aim was to explore the possibility and potential of experiencing VS under the proposed circumstances and what it could offer to architecture.

While the project is being further developed, what follows is based on its state when it was fist posted online in November 2014[10].

Head Mounted Display platform

The major property that we wanted to include and create a correspondence of, from the real to the virtual space, was physical presence and movement. In order to have the user freely moving in a given space where the project would be installed, we had to minimise cable dependencies and we therefore abandoned the idea of using the Oculus Rift. Instead, we chose to implement a wireless mobile setup using a smart-phone as a HDM platform. At the time, the concept of using smart-phones as VR HMDs was at its beginnings, and the aforementioned commercial products (Google Cardboard, Zeiss VR ONE, Samsung Gear VR) were still not available (Google Cardboard, which was the first to be released, was announced at the Google I/O conference in June 2014[11]). Eventually we made use of the Durovis Dive SDK[12] that provided a framework for adequately utilising a smart-phone as a HMD.

For a smart-phone mount, Durovis offered the "Open Dive"[13] an open CAD drawing for a 3D printable smart-phone head mount which we used while prototyping consecutive versions that solved various problems of the original -inter-pupillary distance adjustment, light-proofing, rigidity, ergonomics etc. The final version labeled "omni"[14] was optimised for an iPhone 5s and can be fabricated on a regular 3D printer at a minimal cost, requiring as extra parts lenses (50mm ø 25mm), padding and an elastic strap (Figure 1).

Figure 1
omni mask
prototype

Real-time tracking

Tracking devices. Using this setup, we already had the head movement of the user captured by the IMU of the phone, and we therefore needed to find solutions for how to do body and gesture tracking.

For the former we decided to implemented Microsoft Kinect cameras, which operate by using infrared light and can track movement in space by calculating 15 points of a person's skeleton in a frequency of about 30Hz. Each camera covers a frustum of 58° by 45° in a depth of roughly 6 meters, and can track up to 6 skeletons without the need of tracker markers for full body interaction (Lange et al. 2011).

Regarding gesture tracking we initially tested 2 commercial devices: the Leap Motion Controller[15], and Thalmic Labs Myo[16]; both of which we rejected for our purpose, the first due to its cables, range and orientation limitations, and the second because of limited gestures and imprecision. Without any other options left we proceeded by prototyping a pair of interactive data gloves. The "inteliglove"[17] system much like other similar devices[18] (Benali-Khoudja, 2014) uses flex sensors for capturing finger movements and a 9 degree of freedom IMU for calculating orientation. It was programmed using Arduino micro-controllers, and the IMU sensor fusion was based on the "Razor 9DOF AHRS" open source algorithm[19]. The circuit also features an XBee radio module and is powered by a Lithium Polymer battery which renders the device wireless. The micro-controller, allowed for a calculation frequency of 50Hz, which was handled adequately by the XBee communication protocol, on the condition that each glove had a dedicated receiver on the other end -the sending rate was very high for one receiver to collect data from more than one transmitters (Figure 2).

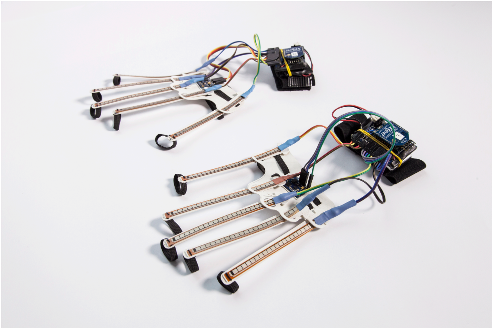

Figure 2
Inteliglove prototype set

Tracking software. Regarding the sensor data, we developed "Omnitracker"; a Java desktop software using the Processing and SimpleOpenNI libraries. Its main functionality is to capture and process data from the Kinect cameras and the inteliglove devices. It allows for a multiple Kinect camera setup, each of which can capture up to 6 skeleton sets at a time. Eventually, a complete tracking dataset for a whole body, includes a 15 vector set for a skeleton's position and for each hand, a pair of 5 finger flex values, and a palm orientation vector, adding up eventually to 73 degrees of freedom.

General Setup and Node Roles

The project is described as an app-system, since though it mainly consists of a mobile app, it also requires, as described before, other hardware and software, controlled by one or multiple computers. Therefore its implementation depends on 2 nodes: the user or Subject Node (SN) on a mobile device and the data server node (DN), on a computer which handles the tracking data. Both are needed to run simultaneously and exchange data in real-time. The way we proceeded to achieve this, was to develop a single application with an embedded networking solution, that can run both on mobile and desktop devices, having though distinct roles for each node type. While both nodes participate in the network by exchanging data, the SN would be an active agent, whereas the DN just a passive element of the Virtual Space.

Application Development

The project's main application was developed with the game engine Unity 3D[20], which was selected because it provides a framework capable of real-time GPU rendering, compatibility with various VR devices and an array of plug-ins ("Assets") that can extend its functionality. Additionally, besides graphics and geometry, it facilitates designing a complete environment including interaction and audio. Last but not least, using the Xamarin Mono[21] framework, Unity 3D provides C Sharp scripting capabilities, along with deployment across all major desktop and mobile platforms (between others: Windows, OS X, Linux Android and iOS).

Networking and Multiplayer functionality

The two networking issues we faced were firstly the effective sending are receiving of sensor data from the DN to the SN, and secondly, allowing the system to host multiple concurrent users in the same Virtual Space. These issues, required a solution for transferring data both on request and on regular intervals. We chose to use the Exit Games Photon PUN[22] cloud service, which was specifically designed for real-time multi-player games and could perform the two requirements by "Remote Procedure Calls" and "Serialisation" update cycles respectively. It can furthermore host up to 100 concurrent users running the application.

The first data cycle regarding sensor datasets starts by the DN capturing and processing the sensor data through the Omnitracker software. The data are then forwarded via the OSC protocol to the Unity 3D application that is running in parallel, and from there are uploaded to the cloud. Instantaneously, the SN related to the specific dataset, downloads and applies it to update its state. This process is asynchronous, because while the Omnitracker can secure a static framerate, its frequency is not necessarily matching the framerate which the application is running (Figure 3).

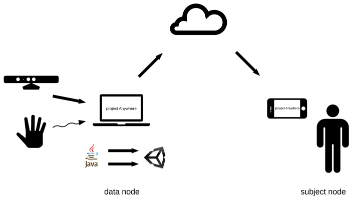

Figure 3
Data stream from Data Node to Subject node

The second data cycle concerns the horizontal updates between concurrent nodes running the application. In this scheme, while each device running the application controls locally its own avatar, it also hosts passive instances of the avatars controlled by remote nodes. Each local instance, besides performing the data cycle described before, is responsible for uploading it's state changes to the cloud, at regular intervals, so as to inform its instances hosted on remote nodes. (Figure 4).

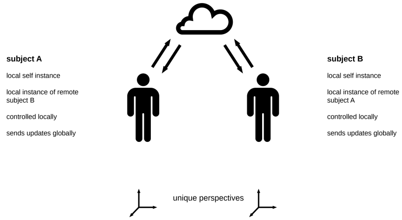

Figure 4
Horizontal data synchronisation of concurrent users

Additionally, actions performed by any SN which affect the shared VS, are asynchronously sent through the cloud to all nodes to apply to their version of the VS.

Digital Avatar and Virtual Kinaesthetics

Of topmost importance for creating the immersive experience we set for, was to achieve virtual kinaesthetics-prioproception. To create a virtual body that would follow the user's real body, and contribute to enhance the sense of presence and self-awareness in the VS (Slater and Usoh, 1994). We therefore used the sensor dataset downloaded from the cloud -responsible for the user's body and hands- along with a rotation vector produced by the smart-phone's IMU, responsible for tracking head orientation, fused together to produce a multidimensional vector describing the user's state in space. To formalise the user's presence, we therefore implemented as a digital avatar, a humanoid 3D model capable of skeletal animation, to which we applied the aforementioned quantisation vector in order to create real-time animation -which proved sufficient for this purpose. Eventually, the users' body, hands and head rotations and movements in real space, were aligned and corollated with the ones of the digital avatar in a 1:1 relationship (Figure 5).

It is important to note, that since the virtual body of the user is entangled to the real, and therefore the virtual view point moves with the user's body, viction, which describes the illusory perception of self-motion, and is one of the main causes of disorientation in VR applications, in our case was minimal (Hettinger et al. 1990).

Additionally, the digital avatar featured physics behaviour, meaning they could physically interact -collide- when in contact with other virtual objects or avatars of the VS.

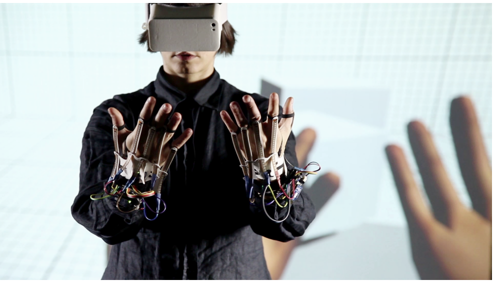

Figure 5
By placing their hands in front of their eyes, the users can see their virtual hands animating in real-time (The background is projection of what the user currently sees -in mono)

Active Presence
Having solved the previous, we wanted to add some actions, in order to display and test the interactive nature of the project, and therefore prove the active presence of the user in the VS. As a primitive interface scenario, we programmed hand gestures to perform manipulations on geometrical objects. These were object creation, scaling, colouring, rotation and movement, each of which was assigned to a unique gesture. Using this tactile interface, users could intuitively perform specific actions -geometric in this case- which offered a hint that the shared VS could be affected by each of its users -given that the created objects also take part in the network (Figure 6).

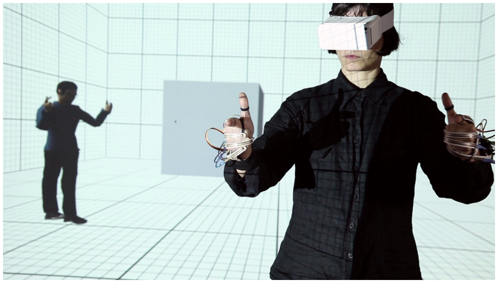

Figure 6
A user creating and scaling a box using gestures (The background is a projection of the current view of the virtual environment)

Virtual Environment
As a virtual environment, the appearance of the virtual space overlapping the real, the game engine we used to develop the application, allowed the use 3D digital geometry. We were therefore able to import 3D CAD models with textures which successfully dressed the VS, and were rendered in real-time by the application. With the only constrain of using the available space of the installation which was effectively tracked by the motion capture sensors, we could overlay any context, similar or dissimilar to the real. Not needing to play any functional role, since it is immaterial, this environment simulating a built object or even landscape, does not additionally need to either follow the way we design to build, what our real environment -urban or natural- looks like or its physical laws (Sutherland, 1965). Furthermore, surpassing the limitations of what we conceive as built environment, the virtual environment can also be programmed to animate and alter in time.

The Characteristics of a New Spatial Medium
Putting everything together, the project creates a first of its kind experience, in a novel spatial form. To justify our claim, we intend to offer a list of characteristics or properties that define the singularity of this type of virtual space. To render however, these properties easier to grasp, we will provide an example of a potential application of the virtual installation described through the project, in the form of an exhibition space, drawing from the author's proposal for the "Museum of Science Fiction" architectural competition[23].

The general concept for this hypothetical application, supposes the hosting of a virtual exhibition in a physical building. The setup requires for the space to be covered by a layer of motion capture sensors, which are responsible for quantising the visitors body position and movements. Furthermore, each visitor, is handed a mobile HMD -smart-phone with head mount- connected to the internet and running the application. A data centre controlling the motion capture sensors, scans the space and sends each skeletal dataset to the device of the visitor it corresponds, for it to update their position and avatar state, in real-time. The virtual space where the exhibits are hosted is overlapping the real. The visitors, consequently, wandering in a seemingly empty space, just by wearing their HMD, are exposed to virtual space overlapping and enhancing the real, by creating a mixed spatial experience full of exhibits. Eventually the museum emerges from the augmentation of the real space where the exhibition is hosted and the exhibition content which is found in a digital form in the smart-phone application. The characteristics of this augmentation are as follows:

- The exhibition space itself, does not necessarily need to follow spatial requirements typical for museums. Since only the floor plan is of relevance -the effective space that can be used as the ground of the virtual space- it can as well be implemented in an elaborate museum space, as in an underground parking garage with a ceiling at 2.5m.
- The content of exhibition, in this case, is not limited to a predefined form or collection. It can host any type of medium of "reproduction" (Benjamin, 2006) which can furthermore reach 4-dimensions: image, audio, video, 3D geometry, 3D animation. To offer an example, one can experience and walk inside a virtual Parthenon while the ancient greeks are having a religious ceremony -regardless of their physical location.
- Another emergent characteristic of this space is parallel heterotopy (Foucault and Miskowiec, 1986). Since each visitor is experiencing the virtual space from their individual viewport -HMD device- each virtual space can be different. Each visitor, can, concurrently, experience different content and different surroundings. While one has the Parthenon in front of them, another can have the Great Pyramid of Giza and yet another can be in the sculpture section of a virtual Louvre, and all of them can interact with each other. The museum can overlay multiple exhibitions floors in one.
- Subsequently and in relation to the dimension of time, the space is heterochronic. Hosting exhibits such as 3D animation and film, allows for them to be in a different state for each visitor: the film screening begins at the moment when the visitor enters the room, for each one individually. Therefore, the museum experience can be furthermore personalised and function in multiple dimensions of time at once.
- Since the visitors' motion is synchronised over the internet, these data are non-contextual. Overlaid heterotopy is the inherent property, in which multiple virtual museums of this kind, can overlay their visitors in the same virtual space, regardless of where they are located. A museum visitor in Japan, can be found in the same virtual space next to a visitor in London. The museum is a multiplicity (Deleuze 1994, p. 182).

CONCLUSIONS

Virtual Reality technology is already transforming architectural visualisation, by providing means to preview a design before its materialised. However, the actual benefit for architecture, that the author wants to stress, diverges from this use. The experience of Virtual Space in a full bodily immersion, is one that does not easily accept classification under the deterministic dipole of real and unreal. While not material, this space is definitely sensible and affective to the extend that it requires for itself a new spatial paradigm. As we have demonstrated, the characteristics and circumstances that collective, fully immersive, augmented reality space allows, open new unimaginable horizons for architecture, which we believe that can and will play an instrumental role in developing novel, previously unconceivable, spatial forms and experiences.

REFERENCES

Benali-Khoudja, M, Hafez, M, Alexandre, JM and Kheddar, A 2004 'Tactile interfaces: a state-of-the-art survey', *31st Symposium on Robotics*, Paris

Benjamin, W 2008, *The work of art in the age of mechanical reproduction*, Penguin UK

Brooks, Jr and Frederick, P 1999, 'What', *IEEE Computer Graphics and Applications*, 19(6), pp. 16-27

Cruz-Neira, C, Sandin, DJ and DeFanti, TA 1993 'Surround-screen projection-based virtual reality: the design and implementation of the cave', *SIGGRAPH*, pp. 135-142

Deleuze, J 1994, *Difference and Repetition*, Columbia University Press

Foucault, M and Miskowiec, J 1986, 'Of Other Spaces', *Diacritics*, Vol. 16(1), pp. 22-27

Hettinger, LJ, Berbaum, KS, Kennedy, RS, Dunlap, WP and Nolan, NP 1990, 'Vection and simulator sickness', *Military Psychology*, 2(3), pp. 171-181

Lange, B, Rizzo, A, Chang, C, Suma, EA and Bolas, M 2011 'Markerless full body tracking: Depth-sensing technology within virtual environments', *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*

Moss, JD and Muth, ER 2011, 'Characteristics of Head-Mounted Displays and Their Effects on Simulator Sickness', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(3), p. 308–319

Sutherland, IE 1965 'The ultimate display', *Proceedings of IFIP Congress*, p. 506–508

Sutherland, IE 1968 'A head-mounted three dimensional display', *Joint Computer Conferences*, pp. 757-764 (part I)

[1] https://www.oculus.com

[2] https://www.oculus.com/blog/first-look-at-the-rift-shipping-q1-2016/

[3] https://www.google.com/get/cardboard/

[4] http://www.samsung.com/global/microsite/gearvr/

[5] http://zeissvrone.tumblr.com

[6] http://tuebingen.mpg.de/en/cybernetics-news-2/detail/article/virtuelle-realitaet-braucht-platz/134.html

[7] http://somniacs.co

[8] http://www.htcvr.com

[9] http://thevoid.com

[10] http://www.studioany.com/#!projectanywhere/c1g1s

[11] https://www.google.com/events/io

[12] https://www.durovis.com/index.html

[13] https://www.durovis.com/opendive.html

[14] http://www.studioany.com/#!omni/c1cs0

[15] https://www.leapmotion.com

[16] http://myo.com

[17] http://www.studioany.com/#!inteliglove/cd97

[18] https://www.kickstarter.com/projects/mimu/mimu-glove-for-music?ref=discovery

[19] https://github.com/ptrbrtz/razor-9dof-ahrs/wiki/Tutorial

[20] http://unity3d.com

[21] http://xamarin.com

[22] https://www.exitgames.com/en/PUN

[23] http://www.museumofsciencefiction.org/winners/