The eXperience Induction Machine

Jamie Miller
Division of Science and Mathematics
University of Minnesota, Morris
mill5577@morris.umn.edu

ABSTRACT

Virtual reality is a useful tool that continues to be utilized across many fields. However there are inherent limitations because of the fact that all experience and interaction are purely virtual. A single user is immersed within virtual reality to gain experience alone, unable to be observed by another unless they have access to the virtual world as well. The XIM is a machine that is in mixed reality and therefore not bound by these limitations. This medium allows interaction to one or more users and observation by outside parties, opening up new possibilities that were not possible in virtual reality alone.

Keywords

XIM, Virtual Reality, Mixed Reality, Human-Computer Interaction, Telepresence, Client, Server, Avatar

1. INTRODUCTION

In 1962 the world saw its first immersive, multi-sensory system when Morton Heilig presented his prototype of a sensory simulator, the Sensorama. It displayed pre-recorded footage with other small touches, such as a fan to generate wind, to simulate a motorcycle ride through New York. Some years later in the 1970s, these multi-sensory systems had advanced significantly into true virtual reality systems, immersive systems that are computer generated, with some of the more sophisticated ones being that of flight simulators in the military. These flight simulators had computer-generated graphics in order to create the virtual environment for a pilot to navigate. By the 1980s, virtual systems had advanced to the point where a virtual workspace could be interacted with by a user moving their hand position. By the 1990s, computational ability had increased while the size of computers continued to decrease to the point where virtual interaction became widespread for practical means in areas such as education, medicine, industry, military training and entertainment. Since the beginning of virtual reality, with Heilig creating his Sensorama, to now, virtual reality has advanced tremendously with modern applications being widespread and largely beneficial.

One drawback of virtual reality in the modern day is that it requires immersion of one's senses within a virtual system, during which the user can only be observed by an outside party either in the real world or digital world at any one time. This also limits the interaction to only a single user per set of equipment, as any additional users would have to have a separate means of viewing and interaction. For the use by multiple parties, or to properly observe the user(s), a system would have to retain the element of direct interaction while changing how the virtual world is represented and viewed. This would dictate that the system must be a mixed reality system, a "hybrid" of the real world and virtual reality, that exhibits properties of each. A mixed reality system would allow for users to have a digitally immersed perception, be able to be observed in the virtual environment in which they are immersed in, and still retain ability to interact directly with the digital world.

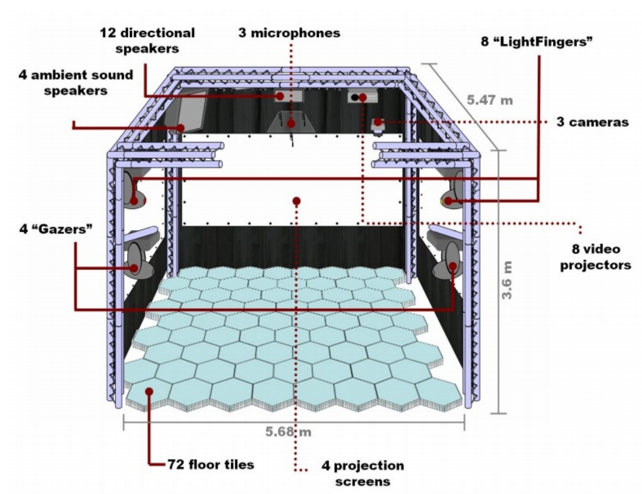

Figure 1: This view of the eXperience Induction Machine (XIM) shows where each component of the machine resides. The sensors are all shown as the methods of input whereas the output devices consist of the microphones, floor tiles, and motion sensors.

As one of the most advanced mixed-reality spaces that currently exists, the eXperience Induction Machine (XIM) does just this. The XIM is an immersive room in Barcelona,

UMM CSci Senior Seminar Conference, April 2015 Morris, MN.

Spain equipped with various sensors that was initially constructed for psychological experimentation in mixed reality. The combination of multiple sensors packed into as small a space as possible creates a unique combination of hand gestures, voice commands, and the tracking of the user's physical movement throughout the room itself as a means of user input with an output given through visualization and sonification. Measurements create a virtual version of the user, known as an "Avatar", on the XIM's server, which mirrors their measured real position and actions. Users are able to experience the digital world while still exhibiting observable behavior in an environment that can be varied in the exact same real environment. The XIM is 4 meters tall and the floor of pressure sensors covers an area of $5.5 \times 5.5\ m^2$.

In the rest of this paper, we will start by providing background to build an understanding of the terminology in this paper. We will discuss some historical precursors and technology that led up to the XIM to build a better understanding of its context, followed by a description of how the machine itself works. Some research cases will examine some possible XIM applications along with potential adaptations for further implementations.

2. BACKGROUND

This section will discuss concepts that are important to the ideas in this paper. We will address different levels of reality as well as other central ideas important to the XIM and interaction.

2.1 Virtual Reality and Mixed Reality

Virtual reality refers to a simulation of presence in an entirely digital world that may represent actual or artificially created places. This refers to any technology that aims to represent a world while using sensory inputs which include virtual representations of the senses. True virtual reality aims for the real world to be completely blocked out by being given stimuli that are purely digital, information that is directly from the virtual world. The XIM does not exclusively take place in virtual reality, though this idea remains relevant as it has components of virtual reality.

As the name implies, mixed reality is a combination of the real and virtual worlds in order to produce a sort of mixture of realities that exhibits some properties of each. Therefore, to interact in this medium, there are components of both the digital and physical worlds. While the line between virtual and mixed reality is not always well defined, mixed reality involves the augmentation of, or addition into, the already existing world as opposed to replacing it entirely, such as the case is with virtual reality.

2.2 Human-Computer Interaction

The field of research that focuses on design and use of interfaces between people and computers is human-computer interaction (HCI). The goal of this is to let humans interact with computers in novel ways, sometimes combining various interfaces in order to combine interactions unique to the technology. It is important that technology strives more toward good human-computer interaction. Human-computer interaction being better means that technology will be used more quickly and effectively in addition to providing the user with a better sense of satisfaction, therefore promoting the medium of usage. At the heart of HCI is improving the usability of a device by more effectively communicating with it and interpreting the actions of humans.

2.3 Telepresence

Telepresence can be simply defined as the state of one feeling as if they were present at a place other than their actual location through the usage of technology. Typical usage of telepresence is that of information travel through a user's voice, movements, or other possible actions being transmitted and duplicated. This allows a user to achieve two-way communication by also receiving information through cameras and microphones. Some practical usage of telepresence is for meetings or other personal use.

2.4 Client-Server Model

A model of client-server interaction in computing is a structure which includes a server, which provides the resources or services in question, and a client, the requester of these resources or services. A server is specified to the type of client it will be serve; a web server is for web pages and can provide this service or function for one or many clients. This model is used for a server to host a program which can be accessed and interpreted by the client.

3. PRECURSORS

Before XIM existed in its entirety, several other machines made important steps in mixed reality interaction. One such machine, RoBoser, is even a part of the XIM.

3.1 RoBoser and Ada the Intelligent

As an important factor in the design behind XIM, RoBoser is one of the preceding machines that led to the interpretation and sonification [2]. RoBoser can be equipped with a variety of sensors, including a camera and microphone, and has the ability to generate sonic structures without a need for human intervention. It maps a real-world system, namely whichever one it is present in, by employing its sensors, then communicates this information to its composition engine. This information is processed and used by its sonic generator to provide sound to reflect its experience; a lighthearted atmosphere may lead to more upbeat sound. This is important to the XIM because of the technology used to interpret signals and communicate through sonification, both of which XIM also possesses. For most demonstrations so far, RoBoser has been mobilized so that it can move around rooms and compose music in real time in response to its environment. The feature of roaming around is not required though and in the XIM, the same abilities take place throughout the entirety of the room.

Ada the Intelligent was an interactive environment that was built to advance research on conscious machines as well as to initiate public debate on brain-based technology [2]. Ada was equipped with mechanisms for visual, audio, and tactile input and was given the capability to learn from experience. For 6 months, the period during which it was displayed, it could interact with visitors and had graphics, light and sound to communicate. Ada's exhibit was not unlike that of XIM as a large room that was equipped with sensors to interact with users.

When Ada the Intelligent was presented to the world, it was in conjunction with RoBoser. The two machines worked in harmonious cooperation with each other with Ada providing the environment while RoBoser provided the interactive sonification. Together they created an experience that is not dissimilar from XIM today. The original purpose for each is even similar with Ada being purposed for stimulating debate on brain-based technology while XIM was purposed for psychological experimentation.

3.2 CAVE

The Cave Automatic Virtual Environment (CAVE) is a 3D immersive virtual, cube-shaped room in which visualization is projected on three to six of its interior walls. The user must wear 3D glasses when in the room for a proper view of the graphics, but doing so allows for visualization of simulated virtual objects in complete 3D. The CAVE also has motion sensors to track user movement and multiple speakers to provide a sense of 3D sound. The motion sensors are mainly used to track the user and readjust the location of the graphics so that the user may continue viewing them as intended. Some modern applications that CAVE provides are for academic use and machine part prototyping. Engineering companies use CAVE in order to examine how a single part may behave or as much as simulating a factory layout in order to develop interfaces. The virtual representation of physical elements is helpful in having a simulation before creating the actual physical component [5].

CAVE is similar to XIM in that it is an environment in which the user is able to be detected and have that input directly affect the virtual environment. However whereas the CAVE does so mainly to update graphics, the XIM allows for much more diversity of interactions. Despite being more limited in most ways however, CAVE was an important step in interaction with mixed reality and therefore a precursor to the various subsequent environments that have come after.

4. XIM INFRASTRUCTURE

This section details the infrastructure of the XIM with regards to physical components and the organizational aspects of the system itself. The equipment mainly consists of devices that allow the taking in of input and projection of output in the form of displaying the virtual environment. These are done by various parts of the system communicating back and forth in order to work in real time.

4.1 Equipment

The XIM is made up of various pieces of equipment including the sensors and displays as shown in Figure 1. The input consists of pressure sensors on the floor to detect location, motion sensors to detect gestures, microphones to detect sound, and other various cameras and sensors to detect appearance and other motions. The output includes video projectors which display images on multiple screens and speakers to project sound. Together, this equipment creates a fully immersive experience which allows all motion, movement, and sound to be detected in tandem with reacting to these inputs to affect and modify the visuals and sound for whatever program may be running.

4.2 VR Server and Avatars

The virtual reality server is a part of the system architecture that embodies inputs through interactions and creates the virtual counterparts. This server is the actual implementation of the virtual world which includes users and even a virtual version of the physical floor, full of interactive floor sensors. A user will have information about their position and identity sent to the VR server by the multimodal tracking system. Outside users are also able to connect to the XIM VR server remotely, which returns a virtual representation of the user via an "Avatar" in the virtual environment. This Avatar is created regardless of the user's method of connection but stores more information in the case of actually present users, mainly what the sensors are picking up. This visitor information is sent to the database VisitorDB, which stores it also regardless of whether the user is connected remotely or is immediately present on the local server. The VR clients are connected to two video projectors within the physical machine and render different viewing angles to display these user Avatars in the virtual environment [2].

5. METHODS

To illustrate some of the more practical applications of the XIM, this section will cover some studies that have been done with it. The original purposing of it was for scientific investigation of mental processes and behavior, but the system has the potential for other applications.

5.1 Spatial Learning

One such experiment in which mental processes were empirically investigated was a spatial learning study, conducted by using the XIM [3]. The goal of the study was to assess how guided versus free movement navigation affects spatial memory, or memory pertaining to one's environment. The XIM is a suitable means for which this type of experiment can be done as it can replicate environments in which a subject may interact with a chosen setting.

10 subjects participated in the experiment and split up into 2 equal groups, 1 group to use guided navigation and the other to use free movement. There were 4 simple rooms, labeled A, B, C and D, belonging to a single house as the space to be navigated through. Each room had various details including lamps, tables, chairs, and other various household items as well as a unique wallpaper. These are what the subject would be tested on, remembering them once navigating through the virtual house shown in Figure 2. The 4 rooms are all connected in a square with a blue light on a single tile below where each door would be, stepping on this illuminated tile triggers the door and consequently changes the scenery to the corresponding room. For free navigation, the subject may roam freely within the room and traverse to another virtual room by triggering the door tile. For guided navigation, the subject is automatically guided through each room, beginning in room A and ending in room D, without any choice by the room itself changing to the next after a set amount of time. Each option lasted for 400 seconds with guided navigation taking place for 90 seconds in each room with 10 second transitions after each. After the experiment, each subject is given a printed 2D map of the house and separate printed 2D versions of each of the features contained within the house. The subject was then asked to place these features on the map in the correct places.

After completing this experiment, it was found that subjects who underwent the guided navigation received a higher spatial score. The guided navigation testing had a mean score only slightly higher than the mean score of the free navigation testing. Although this shows that guided navigation seems to be the favorable condition for spatial memory, the results were found to not be statistically significant overall. This suggests that despite guided navigation seeming preferable, the type of navigation does not significantly affect spatial memory.

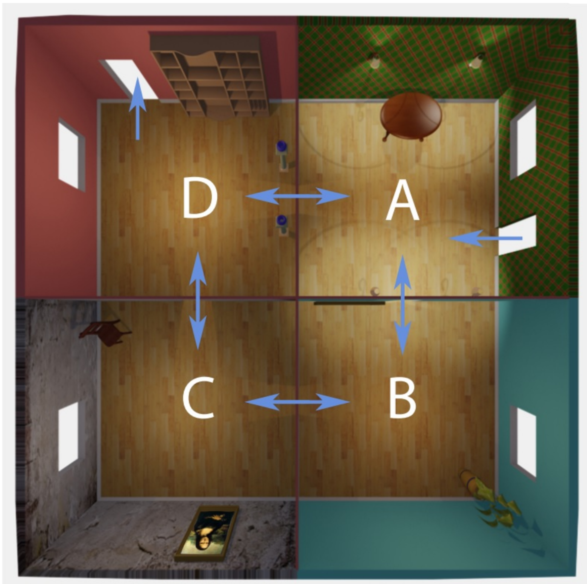

Figure 2: A view of the house from above in virtual reality. In real life, only one of these rooms is portrayed in its entirety at a given time.

5.2 Data Representation

There has been some experimentation done with the XIM as a means to represent complicated or large sets of data such as those of neural networks. In one such experiment, it was compared to a more standard software to provide visualization of large amounts of neural networks [4].

The experiment was conducted on 20 graduate students to test the XIM against a state of the art software, the Connectome Viewer. This software is capable of multi-scale neuroimaging and provides network datasets to develop a better understanding of neural maps. All students participating were selected due to successfully passing the course "Systems Design, Integration and Control" from which they gained first-hand experience in manipulating neural networks. Participants were divided into 2 groups of 10, one for the XIM and one for the Connectome Viewer, and were asked to explore the human cortex dataset provided. Structural understanding of the dataset presented by each of the two methods was subsequently tested via administration of a questionnaire as well as the subjects' visual memory of the dataset by drawing the structure. The questionnaire consisted of 6 questions with the scoring being 1 point scored for a correct answer and 0 points scored for an incorrect answer.

For the structural understanding portion, the participants who used the Connectome Viewer had a mean score of 2.80 $\pm$ 1.62 standard deviation while the XIM had a mean score of 4.30 $\pm$ 0.95 as shown in Figure 3. The p-value for this data is 0.024 which means that we can reject the null hypothesis. A null hypothesis in general refers to a general statement in which there is no relationship between the two sets of data. If a p-value is less than 0.05, or a 5% significance level, it suggests that the observed data are inconsistent with the assumption that the null hypothesis is true. This hypothesis must therefore be rejected and we know that the difference between the data is more extreme, or a significant difference.

The visual understanding test, in which the subjects were asked to duplicate the structure via drawing, was scored by giving participants a score from 1 to 5 based on sketching the structure they saw with as much detail and information as they could remember. The visual test concluded with participants using the Connectome Viewer scoring a mean of 1.5 $\pm$ 0.97 standard deviation while the participants using the XIM scored a mean of 2.5 $\pm$ 0.7 standard deviation. The results of this conclusively favored the XIM system as it allowed for users to have a 3D representation of the data. In addition, the XIM also made interaction and manipulation of the presented data much more manageable and intuitive for the user, as shown in Figure 4.

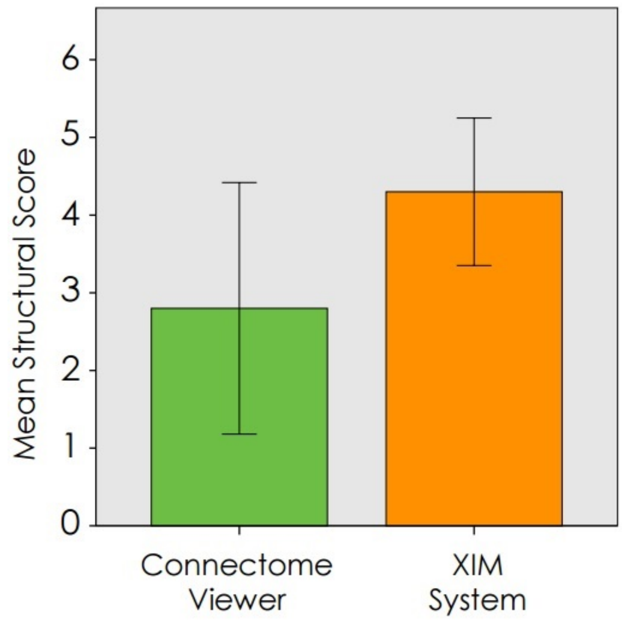

Figure 3: This bar chart represents the difference in structural score between the XIM system and the Connectome Viewer Toolkit. The p-value for this chart is 0.024, rejecting the null hypothesis to show the difference is statistically significant.

5.3 Presence and Experience

As a part of virtual and mixed reality environments, there is the question of user experience and their sense of presence. An experiment [1] attempted to break ground on these concepts with regards to the XIM to gain insight on how to quantify presence in a virtual system, as is the general goal of telepresence research. The researchers' goal was to determine whether employing more objective measures would allow them to externalize their internal states, thereby giving a better idea of presence within the XIM.

There were 18 subjects that took part in the experiment, 6 females and 12 males of mean age 30 $\pm$ 5 standard deviation. This experiment employed the use of Autodemo - a virtual, educational tour guide within the XIM that explains the components and properties of the XIM. The Autodemo is 9min 30s and consists of four stages: "sleep", "welcome", "inside story" and "outside story". An Avatar acts as the virtual guide and leads participants through the story by giving factual information about the XIM with a pre-recorded voice. The process begins with subjects entering during the "sleep" step during which the floor displays a blue wave animation, the walls of the XIM are black, and there is a sleep composition being played. Once a subject enters, this ends and the "welcome" stage initiates with a brief welcome ceremony. The virtual guide then appears, marking the start of the "inside story" step and points to every instrument in the space, explaining each one and what it does while they are respectively illuminated. Then a brief game is played called "Energy" in which the movement speed of the subjects is rewarded with a corresponding level of sound and light compositions. The point of this game is to encourage interaction within the space. Next comes the "outside story"; the walls light up and each user is represented in the virtual space by an avatar. The whole space moves through the virtual world to suggest a sense of space and depth to the subjects. The end is signaled by the users playing "soccer" which is almost identical to one of the earliest arcade video games, "pong". The position of the ball and paddles are represented by floor panels lighting up with a paddle being controlled by an individual standing on a panel.

Tests were administered afterwards for user experience and Autodemo recall. User experience was assessed using a questionnaire that measured the sense of physical space, engagement, ecological validity, and negative effects. This questionnaire was given as 44 statements that each subject was asked to respond to with a 5-point Likert scale, 1 being "strongly disagree" and 5 being "strongly agree". The Autodemo recall test was administered in the form of 10 questions with 3 choices, 2 quantitative open questions from the "inside story" stage, a question on how they perceived the emotional state of the avatar guide, a question asking the guessed duration of the Autodemo, and a request to draw the position of each XIM instrument within a cube. The subjects scored a mean of 6 ± 2 out of 11 factual questions correct with one participant's results being excluded after he got only 1 question right and was assumed to not have actively participated in the experiment. The mean ratings for the user experience test were as follows: 2.8 for physical space, 3.3 for engagement, 2.3 for ecological validity, and 1.7 for negative effects. These results for user experience were compared to the results taken from other media experience experiments that employed the same test and it was found that the sense of physical space was similar to that of IMAX 2D and computer game environments but less than IMAX 3D [6]. It was also found that there was a positive correlation between presence engagement and factual recall.

6. RESULTS

This section discusses how the results to these experiments gives us both hard evidence of what the XIM can do as well as an idea of how users interact with it.

6.1 Spatial Learning

Despite the lack of statistical significance between free and guided navigation in this experiment, there is still much to be observed. The XIM provides a customizable setting that researchers can modify to their specific needs. It would be possible to conduct this same study in a real house but it would be much less convenient and the nature of setting less volatile. This opens up many new possibilities for research and experimental opportunities that would otherwise be much less convenient, or even not possible. The sheer amount of equipment opens up the possibility of an immersive environment which is ideal for the immersion and examination of a test subject.

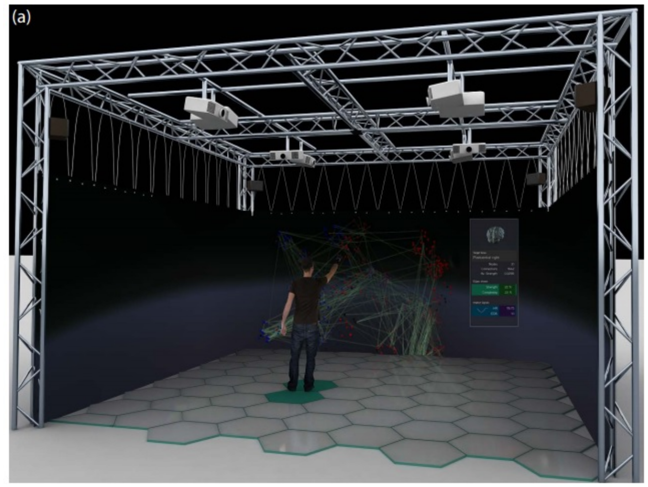

Figure 4: This illustration shows a subject in the XIM, manipulating neural data. Screens on each side give a 3D visualization of the selected dataset while hand gestures allow for interaction.

6.2 Data Representation

The data representation experiment shows promise for the XIM to be used for other applications including academic or for research. This is an example of how it can be used academically to develop a general understanding of vast or complicated sets of data. The usability and comprehension of presented data is strong and applied in a way that is generally convenient. If it is the case that the XIM is a stronger system for human usage than state of the art software in some fields, it definitely has foreseeable application. It also has distinct advantages as it does immerse the user to supply a much better perspective of information that is more sophisticated in visual display. If the 3D visualization of neural maps is better than that of a PC, it follows that it could also give a better display of other digital objects.

6.3 Presence and Experience

After the administration of various tests and measuring the results, there is some conclusive evidence on user experience and sense of telepresence within the XIM. The mean score of answering factual questions about the XIM being 6 out of 11 indicates that the presumably new experience was not too distracting for retention of information, perhaps maybe even the opposite [1]. The user experience score averaged 2.8 for physical space, similar to that of IMAX 2D and computer game environments, but the engagement score also averaged similar to that of a computer game environment [6]. We can conclude from this data that users are capable of feeling ample simulation, while not feeling as telepresent as for the IMAX 3D, but also feel engaged similar to a video game and retain information.

7. POTENTIAL ADAPTATIONS

This section discusses the XIM's potential for forward motion in application. There has been more recent experimentation done in order to use it for its intended design, psychological investigation of human behavior. However, the future holds more in store with a framework that gives a better ability to perform investigation.

7.1 XIM-Engine Framework

Being able to develop multimodal interpretation of human-machine interaction will become increasingly important with further exploration into the realm of mixed reality. The XIM-Engine is a software framework that has been developed as a result of this as it allows for the capturing and interpreting of human behavior for the XIM [7]. This was done by using the theory of Distributed Adaptive Control (DAC) cognitive architecture as the design foundation. DAC defines the inputs and outputs of a system to regulate decisions and infer unconscious process from signatures of implicit signals. A loop of communication is created in which the XIM tries to accelerate achieving the user's goal based on these signals. Some of the methods used include an eye tracker, sensing shirt, sensing glove, and depth sensor which take a variety of measurements such as pupil size, fixation, heart rate, and gestures.

This is a significant component for the future of the XIM because it allows for the furthering of human interaction with mixed reality. Modern day analysis of complex datasets is becoming increasingly complicated as the volume of data increases which make a framework like this important to supplement it. As we are human, the nature of our own interaction should be an integral part of the design. With the XIM-Engine, we can use vast amounts of collected data which is then interpreted and integrated into the behavior of the XIM. Because this is a software framework, this also means that it opens up possibilities for new experimentation and research to be done in this field.

8. CONCLUSIONS

Examining the eXperience Induction Machine in its entirety reveals what it excels at and what it perhaps does not. As a mixed reality machine it has the capability to allow users to directly interact with the virtual world by performing physical actions. Although not complete virtual reality, XIM existing as a mixed reality machine is more practical for applications that are not possible through virtual immersion. Interaction can be achieved more easily, as it requires no additional hardware peripherals, and is more intuitive, as the controls requires less training.

As one of the original design purposes was to study human beings in an environment, various studies have been conducted since. The environment may lack physical components outside of the machine itself, but this can be used as an advantage since it gives researchers complete control to change the nature of the environment with the XIM. In psychological research, the XIM is an invaluable tool because of this; a subject can be observed in a customizable environment. In addition, the physical subject can then be observed in physical reality by all present parties despite being in a virtual setting, giving observers or researchers a much better perspective of the whole picture. This allows for easier testing of variables as well; the experiment takes place in a world that is subject to change through the simple manipulation of a computer peripheral.

Other experimentation and discussion suggests that the XIM may not be limited to its usefulness to research. Showing that visualization, usability, and comprehension of 3D objects is superior to alternative methods opens up many new possibilities. It also brings us back to concepts from human-computer interaction; the XIM is intuitive to use, versatile, and promotes better comprehension. The sensors that the machine has pick up simple hand motions, movements, and sound permit a high degree of interaction. This suggests that in addition to research, the interface design holds merit on its own. If this mixed reality environment is explicitly preferable to the traditional desktop computer for some applications, it could mean a paradigm shift for information technology.

9. ACKNOWLEDGEMENTS

I would like to thank my adviser, Elena Machkasova, and Scott Steffes for providing me with their feedback and time.

10. REFERENCES

[1] U. Bernadet, M. Inderbitzin, S. Wierenga, A. Väljamäe, A. Mura, and P. Verschure. Validating presence by relying on recollection: Human experience and performance in the mixed reality system XIM. In *Proceedings of the 11th Annual International Workshop on Presence*, Presence 2008, Padova, Italy, 2008. CLEUP Cooperativa Libraria Universitaria Padova.

[2] U. Bernadet, B. Badia, A. Duff, M. Inderbitzin, S. Le Groux, J. Manzolli, Z. Mathews, A. Mura, A. Väljamäe, and P. Verschure. *The eXperience Induction Machine: A new paradigm for mixed-reality interaction design and psychological experimentation*. Springer London, 2010.

[3] A. Betella, E. Bueno, U. Bernadet, and P. Verschure. The effect of guided and free navigation on spatial memory in mixed reality. In *Proceedings of the Virtual Reality International Conference: Laval Virtual*, VRIC '13, New York, NY, USA, 2013. ACM.

[4] A. Betella, E. Bueno, W. Kongsantad, R. Zucca, X. Arsiwalla, P. Omedas, and P. Verschure. Understanding large network datasets through embodied interaction in virtual reality. In *Proceedings of the 2014 Virtual Reality International Conference*, VRIC '14, New York, NY, USA, 2014. ACM.

[5] T. DeFanti, C. Cruz-Neira, and D. Sandin. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '93, New York, NY, USA, 1993. ACM.

[6] J. Lessiter, E. Keogh, J. Freeman, and J. Davidoff. A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence*, 10:282–297, June 2001.

[7] P. Omedas, A. Betella, R. Zucca, X. Arsiwalla, D. Pacheco, J. Wagner, F. Lingenfelser, E. Andre, D. Mazzei, A. Lanatá, A. Tognetti, D. de Rossi, A. Grau, A. Goldhoorn, E. Guerra, R. Alquezar, A. Sanfeliu, and P. Verschure. Xim-engine: a software framework to support the development of interactive applications that uses conscious and unconscious reactions in immersive mixed reality. In *Proceedings of the 2014 Virtual Reality International Conference*, VRIC '14, New York, NY, USA, 2014. ACM.