CEDE: Collaborative Egocentric Design Environment for CAVE

Sinan Kockara[1], **Tansel Halic**[2]

Computer Science, University of Central Arkansas[1], Rensselaer Polytechnic Institute[2]

Abstract- *The geographical separation of the users and insufficient input from pairs may cause undesired costs in interior/exterior design process. If immersive communication and interaction exist between the pairs, those undesired costs will be minimal. Lack of communication and understanding causes much of the design works to be lost as it is improperly passed on and/or incompletely understood. The combined immersive collaboration can solve these existing problems so that the undesired costs can be eliminated. We present a framework for 3D design of interior/exterior spaces using a CAVE-based immersive virtual environment in that better and intuitive communication among the designers and users would lead to better integration of overall design such that designer would cede some of his/her work to the user. We have developed several user interface and intuitive spatial egocentric/exocentric perception and perceptual constancy techniques to build a fully interactive environment for novice users to help design process of interior or exterior spaces. Efficiencies of these spatial intelligence techniques' are investigated with several experimental cases.*

Keywords: CAVE, immersive environment, interior/exterior design, spatial perception, egocentric, exocentric, perceptual constancy.

1 Introduction

Immersive virtual reality (IVR) is usually a three-dimensional computer generated representation of the real world where user is surrounded by the environment and ideally, can interact with it. This is realized by a viewer centered view by tracking the position of the user. With IVR, complex 3D databases can be rendered and investigated real-time that enables exploration of products. One of the most important uses of virtual environments is that it provides a way where designs can be experimented and tested without actually building them. This would reduce development time, lower total costs of developments and in turn improve efficiency. VR has traditionally been used in planning, design, construction tasks in civil, structural, and mechanical engineering problems. A relatively less explored area is the collaborative design of interior/exterior spaces [28]. This requires an intuitive approach for user interactions as many of the users of the virtual environment will not be expert computer users. Our goal in this project is to provide easy-to-use yet powerful collaborative user interactivity and egocentric/exocentric space perception technique for a CAVE-based [1] virtual environment that will allow users to design interior or exterior spaces effectively and perceive spatial data accurately. Our framework is a tool that overcomes the limitations of communication problems and provides an interactive immersive design environment among the people. Through the methods discussed here, we envision that the users will not only be able to design environments dynamically, but also make distance and size judgments accurately by perceiving spatial data from intuitive egocentric perspective channels.

2 Motivation and Background

The exchange of ideas on a real scene objects rather than relying solely on mutual communications would greatly further the conception of 3D scene between designers and users. If we consider the case of custom designed aircrafts as an example, the necessity of collaborative, immersive, and interactive design environment can be easily understood. Planning, completing complex drawings of aircraft interior furnishings and secondary structural attachments, design analysis, testing of various aircraft interior elements and installations are crucial components of the design process. To ensure responses to customer inquiries are complete and accurate; traditional methods provide immediate technical assistance (on site/phone/fax) and perform routine visits to make customer satisfied. This is tedious, time consuming, and very vulnerable approach to costly errors. To overcome these problems in any design patterns (e.g. architecture design, custom aircraft interior design etc.), we develop egocentric, collaborative, interactive design interface in that customer/user can have hands-on participation on development and installation phases. Working on one-to-one same scale virtual objects corresponding to the real world counterparts would provide real sensation. In addition, user can see his/her interactions through different virtual cameras located into the scene. Cameras' views map to the several view ports as shown in Figure 4. This provides egocentric real sense of size, orientation, and distance perception comparable with user himself/herself since a virtual avatar corresponding to the

201 Donaghey Ave Conway, AR 72035
skockara@uca.edu

user with approximately the same height will be present at the scene.

According to Thurstone [23] spatial aptitude consists of three components: first, capability of distinguishing an object from different perspective views, second, ability to recognize displacements in measure, and third, ability to recognize spatial relations when the body of the observer is an essential part of the scene. Our framework complements all three of these spatial ability components at the same time by displaying different perspective views from virtual cameras, adding reference object with already known distance in the scene, and creating the same height corresponding avatar of the user respectively. CAVE (CAVE Automatic Virtual Environment) is an example system that provides a platform for virtual reality applications. It was first introduced in 1991 by Carolina Cruz-Neira at the University of Illinois at Chicago [1]. The CAVE is composed of three to six projection screens driven by a set of coordinated image-generation systems. It is assisted by a head and hand tracking devices that produces stereo perspective.

Figure 1. Three-wall CAVE used for the development of a design environment

Research in IVR applications has produced control menu techniques that take advantage of the user's inherent knowledge of the natural interactions. These techniques enhance the effectiveness by providing extra interactivity features through various types of menu systems used in such environments [2], [3]. Unfortunately, most of these techniques have not been designed to support full interaction and collaboration among geographically apart pairs as in [4]. Several file formats and Application Program Interfaces (API) or libraries (together or separately) can be used to develop 3D IVR applications. VRML, CAVELib[5], VRJuggler[6], OpenGL Performer[7], OpenGL, CAVERNsoft[8], and Open Inventor are examples to name a few. Germans et al. [9] summarizes some of those IVR libraries and applications which are extensively represented in Table 1 by their features. Among these libraries, we use OpenGL Performer, CAVELib, and OpenGL to implement our framework. OpenGL is used to implement some necessary features which are otherwise nonexistent in OpenGL Performer. CAVELib is used to implement necessary base for collaboration and visualization in CAVE.

Table 1. Some Available libraries and tools

Tools	Interaction	Scene Graph	Data Visualization	VR Hardware Support	Collaboration	Multi Platform
CAVELib	Low-level	No	No	Yes	Low-level	Yes
OpenGL	No	No	No	No	No	Yes
Performer	Low-level	Yes	Yes	Yes	No	Yes
Inventor	Low-level	Yes	No	No	No	Yes
VRML	Low-level	Yes	No	No	No	Yes
VRJuggler	Low-level	no	No	Yes	Low-level	Yes
CAVERNsoft	No	No	No	Yes	Yes	yes

Despite many obvious advantages virtual reality-based 3D space design techniques seem to offer, advanced research in this area is hindered by the lack of effective user interaction. We provide enhanced interactivity features in the immersive environment using our user interface for spatial scene representation, a sample scene of which is shown in Figure 2. In our interactive design environment, users are able to position, select, scale, add, remove, orient the models and change textures. With increased flexibility by our implementation, the users can move and interact in the environment freely. It is possible for the users to change the navigation speed in the 3D scene using user interface, for instance, as shown in Figure 2.

Figure 2. A sample interior scene with menu items (Navigation speed by slider)

3 Collaboration

Presence in the virtual world is generally maintained using avatars, or a computer generated representations of participants such as shown in [16]. Transmitting a sufficient amount of body language, furthermore; seeing real time high quality video of other collaborator can improve negotiation for tele-immersive applications. However, in the situation that users can change environment dynamically, face-to-face communication is not strictly required. Seeing what other participant is manipulating into the scene is a considerable issue rather than what he/she is looking at, pointing at, or what his/her head, hand position and orientations are. The most important advantage of doing design visualization in an immersive and collaborative environment is the ability to have geographically distributed participants sharing virtual scene with each other and the 3D objects into the scene can be manipulated reciprocally. This allows the participants to modify specific objects in the scene or set the parameters of these objects e.g. translation values. Collaboration gives the users a common context for their design.

Collaboration in VR applications is supported through numerous libraries that provide network functionalities [17]-[22]. In our implementation, we use simple network functionalities provided by CAVELib. Our purpose is not to produce a real face-to-face meeting between collaborators during dynamic scene creation, but to provide interface for them to work together in geographically apart virtual environments. By means of our collaboration theme, users immerse into the virtual environments to check on the state of the virtual scene and only one of them (primary user) makes alterations to the scene at the same time. First participant is the initiator of the virtual scene. Only one of the partners playing a leadership role (lead person) can adjust the scene, but by giving permission to other peer makes other participant leading person to make alterations.

3.1 System Architecture

The architecture of the proposed system is a client/server model showed in Figure 3, in which CAVE renders lists of 3D graphics objects created by user in parallel to the other collaboration side from where changes are sent to a server as text. Passing 3D databases to other client when modified by any lead participant in the virtual space requires a fair amount of infrastructure to be established and maintained such as higher network bandwidth even if network connection backbone is internet2. Instead, passing text messages to the other client that includes which part of the scene changed and what changes being done are more appropriate and yet more useful. To accomplish this, each participant should have the same scene with the same 3D databases and texture images initially.

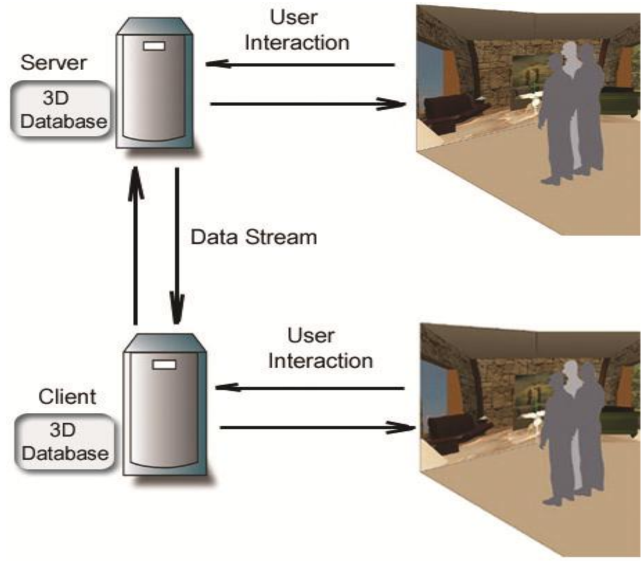

Figure 3 Architecture of the collaborative

We make collaboration synchronous meaning that each participant can see what modifications being done by lead person in the virtual scene. Not to increase network traffic and not to make implementation intricate, we do not provide audio communication between collaborators. Moreover, there is not much advantageous to provide aural communication among users for creating and adjusting virtual scene because every action they can do is implemented in user interface. Still there is no possibility to take user's text input from keyboard like device into the immersive environment except implementing a virtual keyboard. Therefore, we cannot provide users to send each other dynamically created text messages.

Changes made by a lead person in the virtual scene can be recorded as a text message and pass to the other user for informing what modifications being done to prevent confusions. In any case of network connections failure each modification in the 3D world is saved in both sides of the systems. We achieve this by only saving changes into the configuration file that is nothing but text file. This text file consists of all the objects in the scene with their names/IDs, translations, textures, scale factors, color (RGB), navigation speed, and pickable property etc. When collaboration is set up again in both systems, comparison between configuration files is done to detect whether there is any difference between last updated times. If there is difference in time, the most recent one will be transmitted to the other part. This is achieved because each pair has time that last save operation is made into the configuration file.

So far, we have introduced collaborative interactive design environment to found base for egocentric spatial perception framework. Next section introduces human visual perception metrics and their usage in immersive design environment.

4 Egocentric and Exocentric Judgment by Virtual Channels

In an immersive design environment, human's spatial perception plays a crucial role. Egocentric and exocentric judgments of a space are the most prominent contributors to this perception. Humans can either use egocentric or exocentric distance judgment for determination of absolute scales, positions, and orientations of objects in the environment [10]. Egocentric localization is a process of determination of the objects' spatial positions in the environment relative to one's body. Contrary to egocentric localization, in exocentric localization; instead of using one's body, humans use some objects in the environment as a reference for effectively employing positions, scales, and orientations. Humans usually correlate egocentric or exocentric distance judgments interchangeably and repeatedly to eliminate misperceptions [11], [12]. To circumvent these misperceptions in the virtual domain, we provide both strategies at the same time by tracking the user with virtual cameras so that different view ports from cameras provide exocentric perceived space judgment in egocentric terms [10] (see Figure 4 G). Another factor that contributes to accurate spatial perception is perceptual constancy. Perceptual constancy implies the tendency of humans to see already known objects as having standard shape and size regardless of changes in the viewing perspective, distance, and lighting etc. We take perceptual constancy into consideration in our framework by selecting some objects which have common perception in size for all humans e.g. soccer ball.

4.1 Human Perception Metrics

According to Pelz et. al [13], there are some linear combinations of locations specified by egocentric and exocentric reference frames. In our case, both of those frames are provided to the user via cameras' viewing frustums. To make viewing frustums compatible to human's binocular vision field we select binocular vision field of 200° width and 135° height. User can locate him/herself via his/her corresponding avatar in the virtual scene. Since judgments are made with respect to the position of the eye in the head, we resize and position corresponding avatar to the user's eye position and direction which are taken from the tracking device. This also makes the avatar's height almost equal to observer's height. With this help, insufficient proprioceptive information in the IVR system, which is the reception of information about body position and movement by the sensory systems, is tried to be eliminated as much as possible.

People usually rely on previous visual information for accomplishment of distance and size estimations. Brain compares the sensed size of an object to its known real size. A human's all the vision cues (binocular disparity, convergence etc.) are effective within 2 meters [25]. Arditi in [26] indicates that accommodation is effective within 2 meters. Foley in [27] shows that convergence can be effective for distances as great as 8 meters. Binocular disparity occurs when two eyes look at the same thing at slightly different angles those result in two slightly different images. Eyes convergence is the difference in the direction of the eyes. Eyes convergence is only effective smaller than 10 meters distances.

In our experiments, the soccer balls are viewed stereoscopically so that binocular disparity provides the order of objects in depth. However, it does not provide the exact perceived depth of the objects. For a correct depth perception disparity must be fed with information regarding egocentric distance. Therefore, we provide more visual disparity feedback cues to the user via exocentric cameras in egocentric terms in order to devise as a source of information. Depth perception is most effective on short distances (less than 10 meters). Human depth perception is dependent upon stereoscopic depth cue which consists of relative differences between parts of the images of the two eyes. In normal situations, this depth cueing is effective only the distances less than 25 meters or sometimes up to 30 meters (if we assume human eyes are placed 6.3 centimeters inter pupillary distance).

4.2 Experiments

We investigated the question of whether viewing some common sensed objects in the scene and some virtual cameras' viewing frustum that are continuously tracking the user can provide the information necessary to the user in order to successfully navigate, locate and accurately estimate the distances and the sizes in the virtual environment. To that end, the user is permitted to locate and move the virtual cameras freely (see Figure 4). In addition, experiments are developed in compliance with abovementioned human perceptual specifications. Error rates in predictions of distances and sizes are measured. Observer's estimations in size and distance are compared according to with/without egocentric viewing channels, with/without perceptual constancy objects, with/without corresponding avatar, and interior/exterior designs.

After spatial ability test [24] is applied to the participants, subjects are chosen with the score of 60 or higher. Experimental scenes are shown in Figure 4. Two of

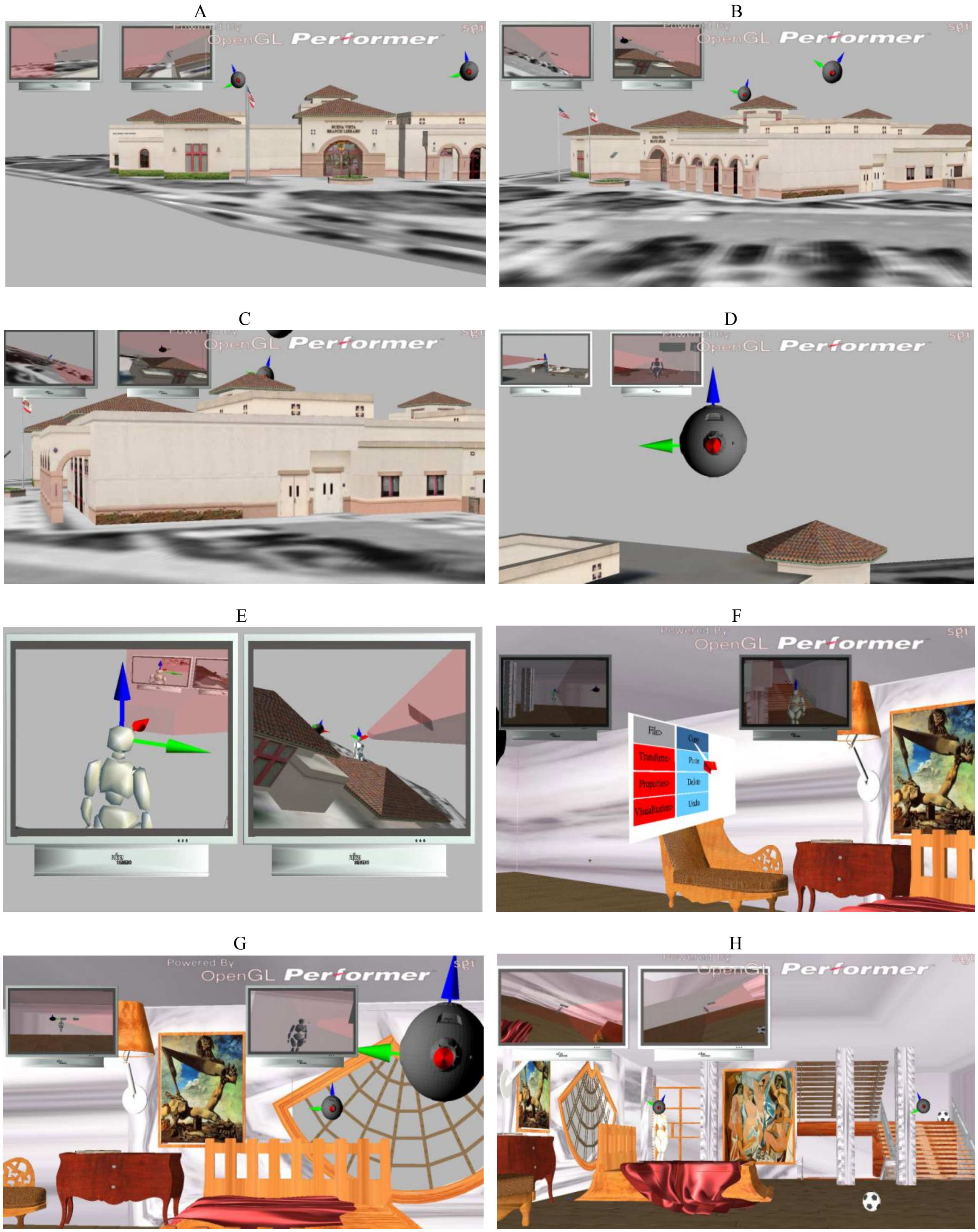

Figure 4. Implementations of spatial intelligence A) B) C) Exterior space with different views and cameras, D) One of the user tracking virtual cameras, E) Virtual cameras' viewing angles on the virtual displays with avatar, F) Object manipulation interface, G)Avatar from two dynamically located virtual cameras and their views on displays, H) Size estimation experiment

the subjects already had very little experience in immersive displays others didn't. Subjects were first guided through a practice session to have little experience and sensation of immersive environment. Interior design scene in Figure 2 is used for this environmental training session. When the subjects verbally declare that they learned how to navigate and interact with the scene, the experiment started.

The 3D interior and exterior scenes are created with exocentric head-up virtual cameras' view ports. Subjects have two different tests; one is depth discrimination, the other one is size discrimination. In the first, subjects are instructed to estimate the distance of a specified object. In the second, subjects are instructed to determine which is the larger of the two soccer balls. Soccer ball is selected because of perceptual constancy. To measure the error rates in depth and size discriminations, soccer balls with different distances (i.e.; 2 Meters, 2.7 meters, 5meters, 10 meters, 20 meters, and 30 meters) are located in front of the user. Extent of the scene is 40 meters in length and 30 meters in width (see Figure 4 H). Verbal indication is taken with questions of how far objects in meters and centimeters are and which one of the objects is larger. Subjects are not informed about their errors or accuracies until finishing all of the experiments. Another experiment was on the number of virtual cameras. We investigated the effect of increased number of virtual cameras would have on the human perception. In addition, shades and shadows are used to resolve some perceptual ambiguities. To that end, user can locate and change the intensity of the light freely.

4.3 Results

Average of all of the subjects' relative erroneous distance estimations with/out virtual cameras and/or the avatar, and the effect of increase on the number of cameras are shown in below graphs.

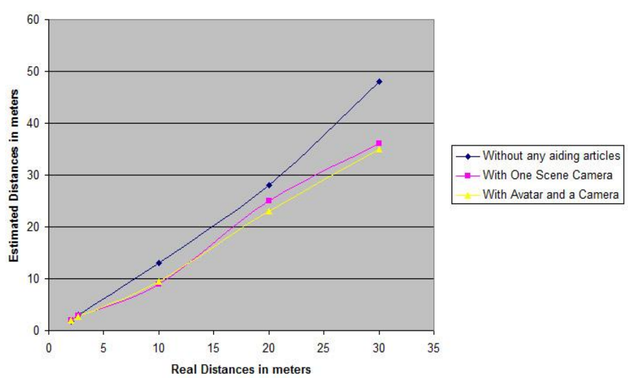

Figure 5. Erroneous Estimations with/out avatar and/or a camera

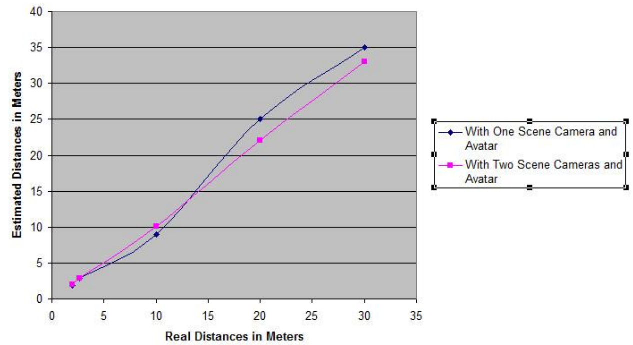

Figure 6. Effects of number of cameras

As seen in the graphs (Figure 5 and Figure 6), including the cameras and the avatar contributes to the spatial intelligence of the human perception and helps to reduce erroneous estimations. This contribution increases with increase in the number of cameras. The experiment showed significant reduction at erroneous distance and size perceptions by combining 3D egocentric view in exocentric virtual displays (which are presenting virtual world representations in an exocentric perspective). We call this intuitive egocentric spatial reading. Without a visual reference (camera's view ports in our case), erroneous judgments of visual directions and spatial perceptions are made by the users. It is also interesting to note that camera's viewpoints' views on the screen are not disorienting the user [13]-[15]. In addition, results showed that perceived size and distance are proportional to perceptual distance and size in immersive displays. Therefore, supporting the proposition of size constancy and distance discrimination is achieved by adding familiar objects to the scene.

5 Conclusion and Future Work

An efficient interactive human centric collaborative design framework is developed. A way of improving human's spatial intelligence in collaborative immersive displays considering human spatial perception metrics is investigated. To that end, egocentric perception from virtual cameras is used as a reference frame. Virtual Cameras with tracking capabilities of the user indeed provide important information for successful navigation and spatial perception through the environment and even affects normal navigational behaviour and perception of the space. Combination of the cameras and the same height avatar with the observer is even further the human perception in space. The results showed that in immersive environments 3D exocentric displays with egocentric frame of reference is far more efficient than only egocentric view.

As a future work we can add some security options to our virtual scene to prevent from fraudulent attacks. In future phase, system can be converted to asynchronous one to make system independent not only from geographical restrictions but also from time zone restrictions. Also, experiments will be pervaded to all subjects including lower scored ones in spatial ability test. Furthermore, gender factors in spatial intelligence for immersive displays will be investigated. Even though effects of increase in the number of cameras are explored, the supportive contributions of more than two cameras on spatial egocentric intelligence are subject to our future study.

7 References

[1] C. Cruz-Neira, D. Sandin, and T. DeFanti. Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. Computer Graphics, pages 135-142. ACM SIGGRAPH, 1993.

[2] S. Kockara, E. Sarioglu, and S. Dagtas. A CAVE-based 3D Interactive System to Design Environments for Performing Arts. ACM SIGGRAPH 2005 Symposium on Interactive 3D Graphics and Games, I3D 05, Washington D. C, April 4–5, 2005.

[3] R.P. McMahan and D.A. Bowman, An Empirical Comparison of Task Sequences for Immersive Virtual Environments. IEEE Symposium on 3D User Interfaces, March 2007.

[4] T. Duval, A. Lecuyer, and S. Thomas. SkeweR: a 3D Interaction Technique for 2-User Collaborative Manipulation of Objects in Virtual Environments. IEEE Symposium on 3D User Interfaces, pages 69-72, March 2006.

[5] http://www.vrco.com, 2007.

[6] C. Just, A. Bierbaum, A. Baker, and C. Cruz-Neira. VR Juggler: A Framework for Virtual Reality Development. In 2nd Immersive Projection Technology Workshop (IPT98), 1998.

[7] J. Rohlf and J. Helman. IRIS performer: A high performance multiprocessing toolkit for real–Time 3D graphics. In Andrew Glassner, editor, SIGGRAPH '94, Computer Graphics Proceedings, Annual Conference Series, pages 381–395. ACM SIGGRAPH, ACM Press, July 1994.

[8] J. Leigh. Issues in the Design of a Flexible Distributed Architecture for Supporting Persistence and Interoperability in Collaborative Virtual Environments. In ACM, editor, SC'97: High Performance Networking and Computing, San Jose (CA) USA, 1997.

[9] D. Germans, H.J.W. Spoelder, L. Renambot, H. E. Bal. VIRPI: A High-Level Toolkit for Interactive Scientific Visualization in Virtual Reality. Proc. Immersive Projection Technology/Eurographics Virtual Environments Workshop, Stuttgart, May 2001.

[10] H. Frenz and M. Lappe. Visual Distance Estimation in Static Compared to Moving Virtual Scenes. The Spanish Journal of Psychology. Volume 9, no 2, pages 321-331, Nov. 2006.

[11] J. M. Foley, N. P. Ribeiro-Filho, and J. A. Da Silva. (2004). Visual perception of extent and the geometry of visual space. Vision Research, volume 44, pages 147-156, 2004.

[12] B. Wu, T. L. Ooi, and Z. J. He. Perceiving distance accurately by a directional process of integrating ground information. Nature, volume 428, pages 73-77, 2004.

[13] J. Pelz and M. Hayhoe. The role of exocentric reference frames in the perception of visual direction. Elsevier Science B.V., Vision Research, volume 35, no 16, pages 2267-2275, Aug. 1995.

[14] L. Matin, E. Picoult, J. K. Stevens, M. W. Jr. Edwards, D. Young, and R. MacArthuer. Oculoparalytic illusion: Visual-field dependent spatial mislocalizations by humans paralyzed with curare. Science, volume 26, pages 198-201, 1982.

[15] L. Stark, and B. Bridgeman. Role of corollary discharge in space constancy. Perception & Psychophysics, volume 34, pages 371-380, 1983.

[16] J. Mortensen, V. Vinayagamoorthy, M. Slater, A. Steed, B. Lok, and M. Whitton. Collaboration in Tele-Immersive Environments. Eurographics Workshop on Virtual Environments, Barcelona, pages 93-101, May 2002.

[17] Christer Carlsson and Olof Hagsand. DIVE — A Platform for Multi-User Virtual Environments. Computers and Graphics, 17(6):663–669, November–December 1993.

[18] Chris Greenhalgh and Steven Benford. MASSIVE: A collaborative virtual environment for teleconferencing. ACM Transactions on Computer-Human Interaction, 2(3):239–261, 1995.

[19] J. Leigh, A.E. Johnson, T.A. DeFanti, and M. Brown. A Review of Tele-Immersive Applications in the CAVE Research Network. In IEEE Virtual Reality'99, pages 180–187, 1999.

[20] M. R. Macedonia, M. J. Zyda, D. R. Pratt, P. T. Barham, and S. Zeswitz. NPSNET: A Network Software Architecture for Large-Scale Virtual Environment. Presence, 3(4):265– 287, 1994.

[21] G. Singh, L. Serra, W. Png, and Hern Ng. BrickNet: A Software Toolkit for Network-Based Virtual Worlds. Presence, volume 3, no. 1, pages 19–34, 1994.

[22] Jason Leigh. Issues in the Design of a Flexible Distributed Architecture for Supporting Persistence and Interoperability in Collaborative Virtual Environments. In ACM, editor, SC'97: High Performance Networking and Computing, San Jose (CA) USA., 1997.

[23] L. L. Thurstone. Primary mental abilities. Psychometric Monographs, 1., 1938.

[24] www.psychometric-success.com/practice-papers/Psychometric%20Success%20Spatial%20Ability%20-%20Practice%20Test%201.pdf, 2007.

[25] L. Kaufman, J. H. Kaufman, R. Noble, S. Edlund, S. Bai, and T. King. Perceptual distance and the constancy of size and stereoscopic depth. Spatial Vision, volume 19, no. 5, pages 439–457, 2006.

[26] A. Arditi. Binocular vision. In: Handbook of Perception and Human Performance, volume 1. Sensory Processes and Perception, Boff, K., Kaufman, L. and Thomas, J. (Eds), pages 23–41.Wiley-Interscience, New York, 1986.

[27] J. M. Foley. Binocular distance perception. Psychol. Rev. 87, pages 411–434, 1980.

[28] Lindsey, P. and McLain-Kark, J., "A Comparison of Real World and Virtual World Interior Environments." Journal of Interior Design, 1998. 24(1): pp. 27-39.