Collaborative 3D Visualization on Large Screen Displays

Daniel Mendes
INESC-ID / IST-UTL
danielmendes@ist.utl.pt
Maurício Sousa
INESC-ID / IST-UTL
antonio.sousa@ist.utl.pt
Bruno Araújo
INESC-ID / IST-UTL
brar@vimmi.inesc-id.pt
Alfredo Ferreira
INESC-ID / IST-UTL
alfredo.ferreira@ist.utl.pt
Hildegardo Noronha
University of Madeira
hildnoronha@gmail.com

Pedro Campos
UMA / INESC-ID
pcampos@uma.pt
Luciano Soares
Tecgraf / PUC-Rio
lpsoares@tecgraf.puc-rio.br
Alberto Raposo
Tecgraf / PUC-Rio
abraposo@tecgraf.puc-rio.br
Joaquim Jorge
INESC-ID / IST-UTL
jaj@inesc-id.pt

Figure 1: Our collaborative 3D virtual environment mixing large scale displays with mobile devices.

CHI 2013 Extended Abstracts, April 27 - May 2, 2013, Paris, France. ACM 978-1-4503-1952-2/13/04.

Abstract

Large Scale Displays, besides their visualization capabilities, can provide a great sense of immersion to a geographically distributed group of people engaging in collaborative work. This paper presents a system that uses remotely located wall sized displays, to offer immersive, interactive collaborative visualization and review of 3D CAD models for engineering applications.

Author Keywords

Virtual Environments; Large Scale Displays; Mobile Devices; 3D User Interfaces; Collaboration

ACM Classification Keywords

H.5.2 [User Interfaces]: Graphical user interfaces (GUI). Input devices and strategies (e.g., mouse, touch-screen), Interaction styles (e.g. commands, menus, forms, direct manipulation)

Introduction

Large Scale Displays offer a unique visualization environment favorable to both individual and collaborative design tasks. During the last decade they became both affordable and easier to setup, providing highly immersive virtual environments with higher resolution and support for stereoscopic images. Combined with emerging input devices, these provide natural ways to interact with virtual

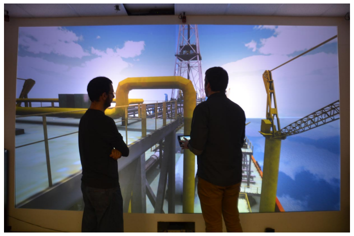

Figure 2: Powerwall with 4.0x2.3 m, using four stereoscopic projectors in Lisbon, Portugal.

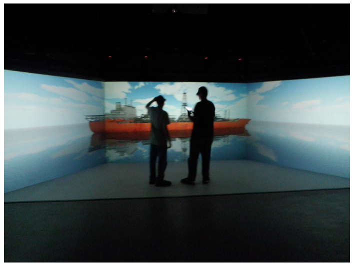

Figure 3: Cave-like system in Rio de Janeiro, Brazil. Each screen has 4.1x2.3 m.

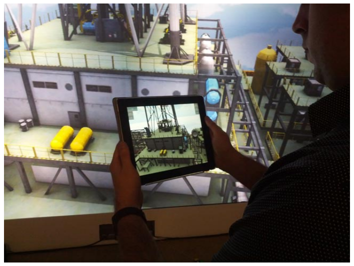

Figure 4: iPad user navigating and sharing his camera point-of-view with the Wall.

content as well as enabling new applications to support collaborative engineering tasks. Following seminal work in multimodal interfaces [2], new devices have become available to enhance interaction including multi-touch tablets combining several sensors, non-intrusive tracking solutions based on depth cameras and affordable brain computer interfaces. We present a virtual reality (VR) system designed to support collaborative visualization of 3D environments, applied to engineering (CAD) design review activities. Our approach combines stereoscopic Large Screen Displays with multiple input devices to provide natural interaction techniques for 3D visualization, while fostering collaboration and remote communication between users across different VR facilities simultaneously.

Related Work

There is considerable research on combining several modalities to interact with Large Screen Displays. Ni et al. [13] surveyed several interaction techniques using mid-air spacial gestures, handheld controllers or touch surfaces to navigate in 3D space. As shown by Jota et al. [8], gestures provide a hands–free solution and can be used precisely to acquire targets on the screen, while allowing mobility around the space in front of the display. Techniques developed for horizontal surfaces can also be applied to mobile touch devices. Boring et al. [3] proposes manipulating content displayed on a vertical display, using a mobile touch device. Nancel, studied [12] uni- and bi-manual interaction techniques to conclude that navigational activities, such as pan and zoom, using touch devices are both more efficient and cause less fatigue than those in mid-air.

Different navigation methods have been proposed for touch–enabled mobile devices. Navidget [7] presents a 2D navigation widget for 3D navigation using a lasso to define the point of interest and rotate around it. Other works, such as Dabr [6], ScrutiCam [5] and Drag'n Go [11] try to propose easier navigation techniques. However these methods do not fully explore the mobility offered by mobile devices. The Tether project [1] mimics the moving window paradigm using iPad devices. However, this approach requires expensive tracking solutions. On a complementary note, brain-computer interfaces (BCI) can enable users to effectively navigate through virtual environments. Larrue et al. [9] concluded that BCI combined with visual information, eliminates intrusive devices for navigation while promoting user experiences closer to reality. However BCIs are still far from being reliable devices.

Virtual environments enhance spatial perception of 3D content during design activities. Additionally, they can provide valuable insights to better decision support with risk mitigation when used in conjunction with collaboration techniques. Following Daily et al [4] work on virtual reality for distributed design review, different on-line collaborative CAD systems [16, 15] have been developed. However their limited interaction capabilities prevented wide acceptance [10]. Human-to-human interaction techniques to share design ideas instantly are thus crucial to make these systems useful. We believe that technologies such as immersive visualization can significantly improve collaborative CAD by providing new possibilities to both understanding information and interacting with it.

CEDAR System

CEDAR proposes a design review tool to support collaborative tasks from the oil industry field. We followed a user-centered methodology starting with a task analysis involving engineers operating traditional design review

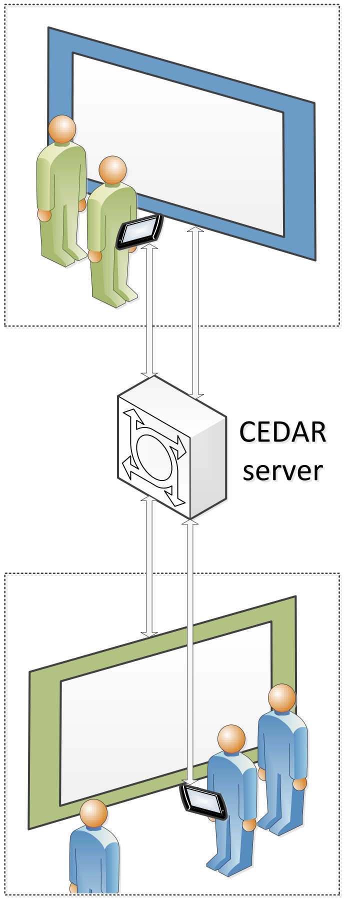

Figure 5: CEDAR System architecture overview.

CAD tools on a daily basis. Via both structured interviews and informal talks, we were able to clearly identify both challenges to be tackled and real usage scenarios.

In oil industry settings, large teams usually gather to review, manipulate and discuss around large CAD models, which are often difficult to visualize and to navigate owing to their complexity. These teams typically include field engineers working at remote locations and technical staff operation at a central location. Maintenance in specific infrastructures inside an oil platform is very frequent and technical staff have to plan how to access them beforehand. Furthermore, changes to these installations are frequent, removing and installing new components. One key goal of ours is to support engineers' field work by allowing them to visualizing in loco the procedures needed to repair devices while assessing possible interferences that are not reflected in the 3D model but exist in the real scene in close cooperation with headquarters.

Setup Overview
Our system was developed and evaluated using two geographically-distinct facilities based using large scale displays with different specifications.

One is a 2x2 tiled display system using four Optoma EW775 projectors in rear-projection mode located in Portugal as depicted in Figure 2. Each projector has 4500 lumens and displays 720p @ 120 Hz. The mosaic generates 3.6 megapixel stereoscopic images thanks to a small overlap (64 px horizontal and 19 px vertical) blending at each projector. The visualization system is driven by a HP Z820 Workstation coupled with two NVIDIA Quadro 5000 behaving as a single desktop and thus simplifying the development and deployment of applications. The stereo setting uses NVIDIA 3D Vision 2 shutter glasses and an IR emitter located behind the screen.

The second visualization facility is located in Rio de Janeiro (Brazil). It is composed of four Barco NH-12 projectors in a trapezoidal arrangement creating a cave-like system as shown by Figure 3. While lateral screens use rear-projection, the floor uses front-projection. Each projector has 12000 lumens and is capable of rendering 1080p stereoscopic content using Infitec active technology. The projectors are connected to a HP Z800 Workstation through one NVIDIA Quadro Plex D2. The desktop uses the mosaic feature and each screen is oriented side-by-side. Unlike traditional cave-type systems, we prefer a 60° lateral glass screen configuration, allowing either a visualizations by a sizable audience or several users to operate within the environment.

System Architecture
The CEDAR system follows a star network topology connecting several clients to a central server as depicted by Figure 5. We consider two different clients to interact with the CAD model under revision during a review session. A client running on a desktop computer connected to a Large Scale Display provides a unique visualization environment for several users discussing engineering problems. Another client, running on an iPad, offers a personal window displaying its own view of the scene. This view affords multi-touch 3D navigation and can be synchronized (or not) to other clients, namely the Large Screen Display.

The main server coordinates the different clients and manages the engineering review session, by forwarding both local and remote communications. By using the client application on each visualization facility, users can connect to the main server either locally, or using the Internet and start sharing data between them.

Figure 6: Screenshot of user interface in mobile device. Virtual joystick on the right to control the navigation with built-in sensors.

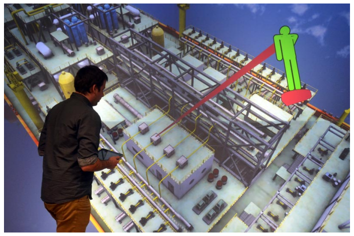

Figure 7: Remote user avatar representation highlighting a target viewed by another user in the large screen display.

3D Navigation

The CEDAR system offers two natural and easy-to-use alternatives to navigate the virtual environment and manipulate the visualization. Users can either use gestures combined with a BCI device to interact directly with the projection or use an iPad device synchronizing its view with the large screen's as depicted in Figure 4.

Users in front of the Large Screen Display are tracked using a Microsoft Kinect depth camera, enabling them to point at specific locations on the 3D scene. Pointing gestures use the Kinect-computed skeleton. A ray is cast from the user head position to their hand position, over the screen to define a 2D cursor. Currently, only one active user is supported within the range of the Kinect camera. Using an Emotiv Epoc device, the user can trigger via a pre-defined brain pattern to move to a specific location on the 3D scene or issue other commands to the graphical user interface, such as (un/)sharing the view with remote sites.

The iPad navigation uses the device built-in inertial sensors to aim the camera in a first-person view mode, as if the user were holding a real camera to visualize the scene. Combined with an on-screen joystick, users can move around the scene as illustrated in Figure 6. This joystick is exponentially scaled to allow both fine movements on the center of its working area and rapid ones at its outer edges. This allows for fine navigation while examining small structures as well as very fast navigation through large models. We also provide a "camera-freeze" mode allowing users to freeze the current view and move around the tablet without their movements changing the camera position on the large screen and without assuming stressful body postures. Further details regarding such technique can be found in [14].

Collaborative Features

Besides providing 3D model visualization and navigation, the CEDAR system supports several collaborative features. The following aim to improve the collaboration between users by providing both direct and indirect communication, visibility and awareness.

Speech Communication. This allows remote, natural communication between the users, which is crucial to the success of the collaborative tasks;

Viewport Sharing. The point of view presented on a display can be shared with other users, allowing geographically apart teams to see the same. Suggested uses are: Displaying the location of small objects, guided tours and simply knowing were the other colleagues are;

Avatars. An avatar representing the current position and orientation is presented on the environment. This affords faster recognition of where other users are located and their look–at direction;

Target Identification. When users point to specific locations in the scene, the system highlights target objects at that position in addition to displaying an avatar (see Figure 7). Thus, other users can know which object is being referred to, without requiring verbal descriptions;

Annotations. During reviewing sessions, it is often necessary to take notes. Instead of taking these on a notebook or using a different application, CEDAR offers both sketched (using the iPad) and speech annotations, that can be viewed or replayed later by users.

Current prototype

Both Large Screen Display and iPad clients implement specific components to manage hardware, communications and input from users as depicted by the software architecture presented in Figure 8. For each client, the User Interface is responsible for providing interaction support and visual feedback. Both server and

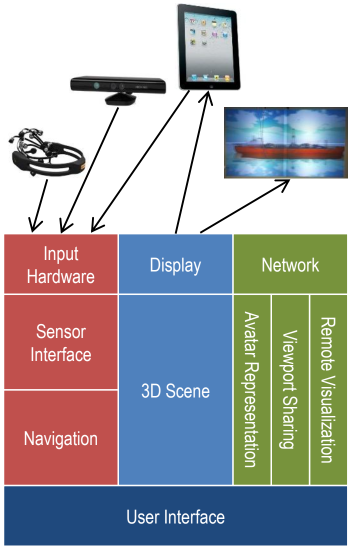

Figure 8: The CEDAR clients framework: software components used based on Unity3D Game Engine.

clients were developed on Unity3D game engine. This provides a common framework for the client applications while reducing software integration efforts on different platforms. Unity3D also provides a fast and reliable support to rendering complex 3D scenes. 3D Scene Unity components are created representing the 3D scene repository, virtual camera and avatar representations.

Three separate components are responsible for sharing virtual representation across the network using Unity remote procedure calls to support collaborative tasks. The *Avatar Representation* synchronizes the point of view among all clients on every frame, i.e. the position and the orientation of all connected users are sent to the server and are distributed among other clients. This component works together with the *Remote Visualization*, which presents remote client's point-of-view either as a preview or a full–display visualization. The *Viewport Sharing* is responsible for broadcasting the current viewport of the client to connected clients over the network.

The *Navigation* component implements different travel methods on the VE depending on the sensor interface provided. On the iPad client side, the *Sensor Interface* uses the iPad built-in sensors to support multi-modal input. On the Wall client side, the Sensor Interface provides different input modalities. Using Open Sound Control as a network protocol for communication among different and independent input applications, the system is scalable enough to handle navigational traffic originating from motion sensing input devices (Microsoft Kinect), brain-computer interfaces (Emotiv Epoc) and the more conventional WIMP input devices.

A preliminary evaluation of our prototype has already been carried out. We asked different pairs of participants to guide each other through the model using the iPad navigation method. While a large formal user evaluation is yet to be done, we gathered valuable qualitative information. Notably, *Viewport Sharing* was preferred over the *Avatar* when guiding or following another user. The probable reason for this is that the model has a lot of obstacles which occlude the *Avatar* most of the time. It is only when users are near to each other that they can take advantage of the *Avatar*. To overcome this and get more visual cues and landmarks, some users navigated to a position high enough that they could see the whole model. On the other hand, *Viewport Sharing* can be disorienting since fast rotational movements performed by the user in control can be hard to follow by other users who are not controlling the navigation.

Some users resisted to using the sensor–based iPad navigation, maybe owing to its novelty. However with practice, they did find it both easier and more fun to use than traditional touch-based interfaces. The main drawbacks were the thumb-stick metaphor on the screen which caused some problems to users less experienced with game-inputs-devices and how tiring the handling position could prove after prolonged use. To address the latter, we proposed to relax the virtual axis and to use the camera freezing feature to work as a clutch, which was also useful to prevent awkward user body postures.

Conclusions and Future Work

This paper explores collaborative 3D visualization on large and small screen displays. We presented our setup composed of two (extensible to more) geographically disjoint large screen stereo displays. To navigate the virtual environment, we propose two alternatives: iPads, that besides offering multi-touch navigation can work as a second personal window on (or off) the scene while of depth cameras and brain-computer interfaces provide novel approaches to interaction. To support remote collaboration between users, we developed a set of specific groupware features. While our preliminary evaluation provided positive results and invaluable feedback from users, further extensive user tests of the system are in order. Our current and future work focuses on improving the prototype, both in terms of user collaboration and immersion.

Acknowledgements

This work was partially supported by FCT (Portuguese Foundation for Science and Technology) through the project CEDAR (PTDC/EIA-EIA/116070/2009) and through the PIDDAC via INESC-ID multi-annual funding.

References

[1] M. Blackshaw, D. Lakatos, H. Ishii, and K. Perlin. T(ether) a tool for spatial expression, 2012. http://kiwi.media.mit.edu/tether/.

[2] R. A. Bolt. ′put-that-there′: Voice and gesture at the graphics interface. In *Proceedings of SIGGRAPH '80*, pages 262–270. ACM, 1980.

[3] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch projector: mobile interaction through video. In *Proceedings of CHI '10*, pages 2287–2296. ACM, 2010.

[4] M. Daily, M. Howard, J. Jerald, C. Lee, K. Martin, D. McInnes, and P. Tinker. Distributed design review in virtual environments. In *Proceedings of CVE '00*, pages 57–63, NY, USA, 2000. ACM.

[5] F. Decle, M. Hachet, and P. Guitton. Tech-note: Scruticam: Camera manipulation technique for 3d objects inspection. In *Proceedings of 3DUI '09*, pages 19–22. IEEE Computer Society, 2009.

[6] J. Edelmann, A. Schilling, and S. Fleck. The dabr - a multitouch system for intuitive 3d scene navigation. In *3DTV Conference*, pages 1 –4, may 2009.

[7] M. Hachet, F. Decle, S. Knodel, and P. Guitton. Navidget for easy 3d camera positioning from 2d inputs. In *Proceedings of 3DUI '08*, pages 83–89. IEEE Computer Society, 2008.

[8] R. Jota, J. M. Pereira, and J. A. Jorge. A comparative study of interaction metaphors for large-scale displays. In *Proceedings of CHI EA '09*, pages 4135–4140. ACM, 2009.

[9] F. Larrue, H. Sauzéon, L. Aguilova, F. Lotte, M. Hachet, and B. NKaoua. Brain computer interface vs walking interface in vr: the impact of motor activity on spatial transfer. In *Proceedings of VRST '12*, pages 113–120. ACM, 2012.

[10] W. Li, W. Lu, J. Fuh, and Y. Wong. Collaborative computer-aided design: research and development status. *Computer-Aided Design*, 37(9):931–940, 2005.

[11] C. Moerman, D. Marchal, and L. Grisoni. Drag'n go: Simple and fast navigation in virtual environment. In *Proceedings of 3DUI '12*, pages 15–18. IEEE, 2012.

[12] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of CHI '11*, pages 177–186. ACM, 2011.

[13] T. Ni, G. Schmidt, O. Staadt, M. Livingston, R. Ball, and R. May. A survey of large high-resolution display technologies, techniques, and applications. In *Virtual Reality Conference*, pages 223 – 236, march 2006.

[14] H. Noronha, P. Campos, J. Jorge, B. d. Araujo, L. Soares, and A. Raposo. Designing a mobile collaborative system for navigating and reviewing oil industry cad models. In *Proceedings of NordiCHI 2012*. ACM, 2012.

[15] M. Tang, S.-C. Chou, and J.-X. Dong. Conflicts classification and solving for collaborative feature modeling. *Advanced Engineering Informatics*, 21(2):211–219, 2007.

[16] F. E. H. Tay and A. Roy. Cybercad: a collaborative approach in 3d-cad technology in a multimedia-supported environment. *Computers in Industry*, 52(2):127–145, 2003.