A Multi-User 3-D Virtual Environment with Interactive Collaboration and Shared Whiteboard Technologies

Wing Ho Leung and Tsuhan Chen
Carnegie Mellon University

Abstract

A multi-user 3-D virtual environment allows remote participants to have a transparent communication as if they are communicating face-to-face. The sense of presence in such an environment can be established by representing each participant with a vivid human-like character called an avatar. We review several immersive technologies, including directional sound, eye gaze, hand gestures, lip synchronization and facial expressions, that facilitates multimodal interaction among participants in the virtual environment using speech processing and animation techniques. Interactive collaboration can be further encouraged with the ability to share and manipulate 3-D objects in the virtual environment. A shared whiteboard makes it easy for participants in the virtual environment to convey their ideas graphically. We survey various kinds of capture devices used for providing the input for the shared whiteboard. Efficient storage of the whiteboard session and precise archival at a later time bring up interesting research topics in information retrieval.

Keywords

3-D virtual environment, shared whiteboard, avatar, transparent communication, interactive collaboration

Introduction

In the past online text chat such as ICQ [1] has been used as a means of multi-user interaction. As higher bandwidth is available, multipoint videoconferencing systems such as Microsoft NetMeeting [2] or CU-SeeMe [3] that allow people from different geographical locations to see and talk to each other are becoming popular. In recent years, together with the fact that more and more

people are getting connected through the internet with faster and faster speed, there is a growing interest on developing an even more immersive way of multi-user communication by connecting the users in a 3D virtual environment so that people feel that they are communicating face-to-face. Some surveys about existing multi-user virtual environments are given in [4].

On the hardware side, pen-based devices are also becoming more and more popular since they allow a more natural way of providing input. This leads to the development of new applications to compliment these equipments. At the same time, it also brings up new problems that evolve into interesting research topics. For example, it is much easier to capture the data from a shared whiteboard application in electronic format with those pen-based devices. With more data captured at a faster rate, can we store the data more efficiently? How can we archive the data fast and accurately?

We have been developing a multi-user 3-D virtual environment called NetICE, Networked Intelligent Collaborative System, and the description of our prior work on NetICE can be found in [5][6]. In this article, we will first describe how each individual participant is represented in the virtual environment. Then we will review immersive interactive technologies used in our multi-user 3D virtual environments towards transparent communication between participants to make them feel that they are meeting in the same place despite their physical locations. We will demonstrate the interactive collaboration in the sharing and manipulation of 3-D objects in NetICE. Afterwards, we will provide a survey about the capture devices that can be used with our shared whiteboard in NetICE. Lastly we will discuss the storage and archival issues for the whiteboard content.

1. Avatar Representation

In NetICE, each participant is represented by an avatar that is an animated human character. By glancing at the avatars, the participant knows who and where they are in the current 3-D world. Analogous to the situation of a cocktail party where new incoming people look around to find friends, the avatar representation of a participant in the virtual environment gives other participants the sense of the presence.

The avatar's head can be represented by either a synthetic face model or a realistic face model. A synthetic face model, as shown in Figure 1(a), usually consists of relatively small number of meshes that are described by a limited set of colors. On the other hand, a realistic face model, as shown in Figure 1(b), requires a much larger number of meshes and it uses a texture map generated by the photo(s) captured for a particular person in order to make the face model realistic. It is easier to animate the synthetic face and the bandwidth required to transmit the animation parameters is much lower compared with the realistic face model. However, a customized realistic face model provides a more unique representation for the participant and it can give other participants better visual cue to associate with his/her identity. Realistic face model can be pre-generated by several methods such as a laser scanner [7], from multiple view photographs [8][9], or from a video camera [10][11].
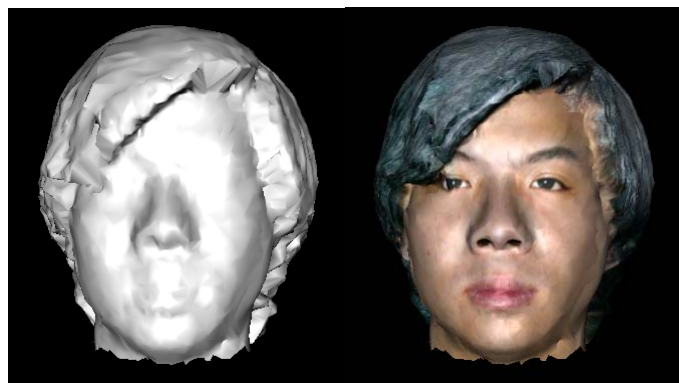
(a) Synthetic face model   (b) Realistic face model before and after texture mapping

Figure 1 Synthetic face model versus realistic face model

During a welcoming party or a reception dinner when many people meet each other for the first time, sometimes people are asked to wear a name tag to facilitate people to remember each other by associating their faces with their names. This idea can also be useful in the virtual environment. In NetICE, a participant can label his/her avatar by typing in his/her name that will show up on the avatar's front and back shirt as shown in Figure 2.

Figure 2 Avatars with names showing up on the front and back shirt

The avatar representation can be made more vivid by animating the avatar body. For example, in NetICE, there is a guide person who will first bow to the audience and then introduce the virtual environment with hand gestures (as shown in Figure 3). In the same room, there is a band whose corresponding avatars can be animated to play virtual musical instruments (as shown in Figure 4) while real music is being played.
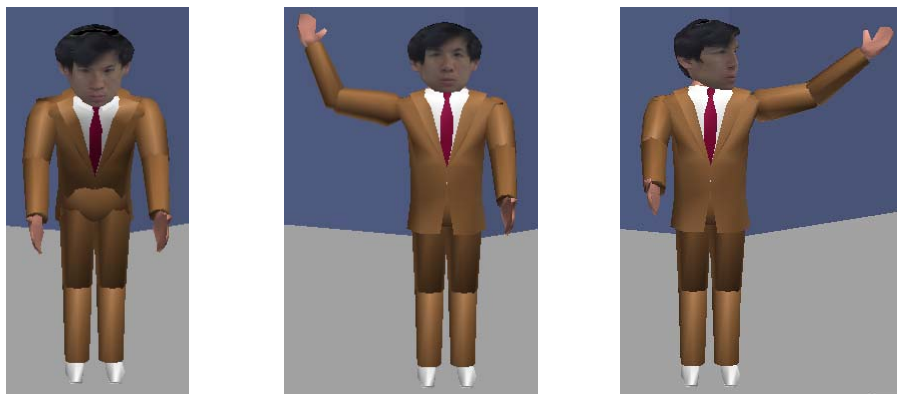

Figure 3 The guide person with body animation in NetICE

Figure 4 The band avatars playing virtual musical instruments

2. Transparent Communication

When people talk by phone, they do not have the same feelings as if they are talking face-to-face. This is because face-to-face communication is not merely verbal since many visual cues are missing when people talk by phone. In order to facilitate a transparent face-to-face communication between participants even though they are not in the same physical location, we make use of the animation, speech and image processing techniques for improving the quality of audiovisual communication. In NetICE, you can tell who is talking to whom, from where the speaker is talking, where he/she is in relation with others, and what expression he/she has on his/her face. These techniques for enhancing the immersive communication in the virtual environment include directional sound, lip synchronization, eye contact, facial expression and hand gesture.

2.1 Directional sound

The directional sound technology (sometimes referred to as the 3-D sound) provides a realistic and immersive sound environment [12][13]. Thus, a user can hear an increasingly louder voice while moving towards the speaker. Similarly, the voice fades away when the user moves away from the speaker. Moreover, different proportions of sound are steered among the left and right sound channels according to the listener's and speaker's relative positions so that the listener can feel

which person is talking by listening to the direction of the sound. For example, if the speaker is located at the left of the user, then the sound gets steered more to the left channel so that the user feels that the sound comes from the left.

The earlier prototype of NetICE achieves this simply by weighting the output from the left and right sound output channels. We have modified our approach to make use of the head-related transfer functions (HRTF) [14]. Each of these FIR filters has 128 coefficients. Seventy-four filters are used in NetICE. Half of these filters are the left-channel filters, and the other half are the right-channel filters, so there are 37 sets of left-right filters. Each set of the left-right filters corresponds to the virtual source at each angle, from the front to the side, then to the back, at 5˚ increment. The filter is chosen according to the relative position (angle) of the speaker with respect to the listener. When the relative angle is determined, both left-channel filter and right-channel filter of that angle are selected to filter the speech signal. For example, if a listener knows that the speaker is five degrees to the left of his/her front, then that user can select a left-channel filter and a right-channel filter corresponding to that five degrees angle. Figure 5 demonstrates this process. To add a perception of depth, the signal is divided by a factor determined by the square of the distance between the listener and the speaker before the filtering process.
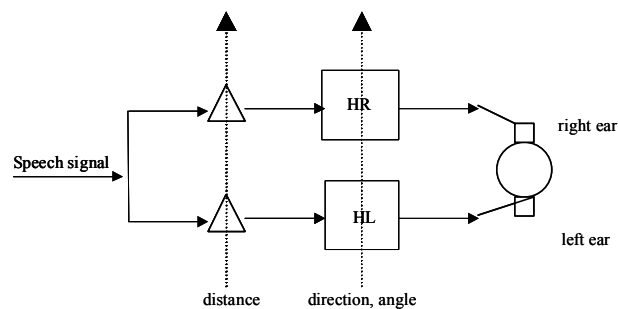
Figure 5 The process of left and right channel filtering

2.2 Lip Synchronization

Interaction between audio and visual information in human speech perception has been a fascinating research field for many years. The well-known "McGurk effect" [15] demonstrates the bimodality of speech perception by showing that, when given conflicting audio and visual cues, a person may perceive a sound that never exists in either modality. There are two types of lip-synchronization. The first is simple energy detection over the incoming samples that directly translates to an angle for the mouth to be open as shown in Figure 6. The second is to analyze the speech to determine which phonemes are present and map them to the basic subset of all visemes, i.e., mouth configurations, that the facial model is capable of generating.
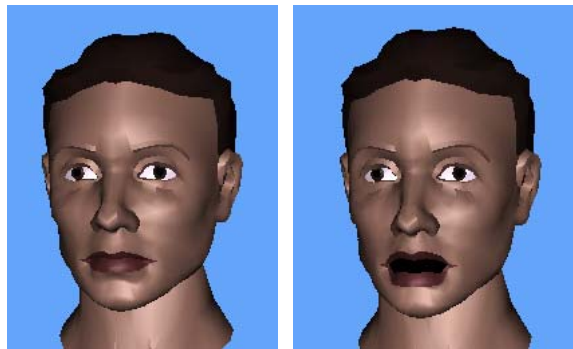

Figure 6 Mouth synchronized with energy of speech

2.3 Eye Contact

Current existing multipoint video conferencing systems such as Microsoft NetMeeting [2] or CU-SeeMe [3] lack the sense of immersion because each participant sees other participants in separate windows, thus it is often difficult to tell who is talking to whom. Tang and Isaacs [16] found that gaze awareness was important because participants could use the visual cue to determine when a remote participant was paying attention to them. In NetICE, it is possible to have eye contact between participants because their avatars are in the same virtual world. The participant simply needs to rotate the avatar's head to maintain eye contacts with others.

2.4 Facial Expression

We allow the participant to communicate basic facial expressions through preset deformations of the face. The face in Figure 7 is composed of triangle mesh of polygons pointing into a vertex list [17] which is manipulated using a combination of linear and sphincter muscles [18][19]. The linear muscles deform the face to produce a fixed set of six rudimentary facial expressions: Happiness, Anger, Surprise, Sadness, Fear and Disgust, as shown in Figure 8. The participant can select one of these expressions and their avatar's face is updated in each of the connected clients.
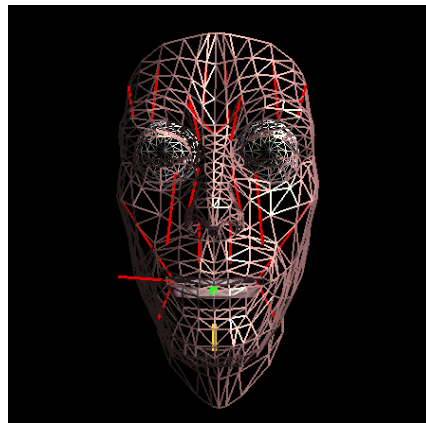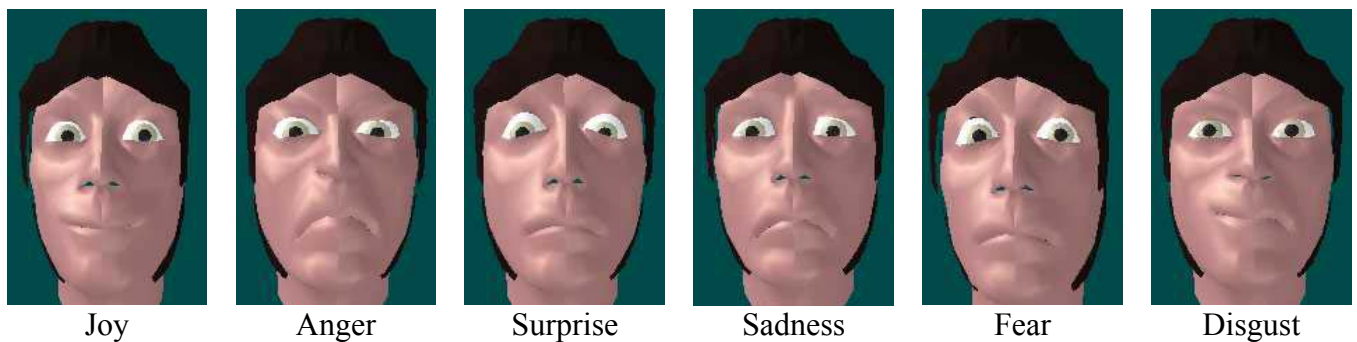

Figure 7 Mesh of the face model

| Joy | Anger | Surprise | Sadness | Fear | Disgust |

Figure 8 The six MPEG-4 compliant rudimentary facial expressions

2.5 Hand Gesture

In NetICE, the avatar can make hand gestures to allow a greater degree of interaction. Before speaking, the avatar may raise a hand in order to get the attention of other users. The hand

movements can also be driven by the energy of the user's speech to enhance the speech communication. When the avatar is speaking at normal voice levels, the avatar can make small hand gestures like any human making a regular speech. Once the avatar's speech emphasizes on certain words by raising the voice or by making sharp bursts of tones, the avatar can make sharp and brisk hand gestures to show the emphasis on the word. After the avatar completes his/her speech, his/her arms will naturally move down to the rest position. Figure 9 demonstrates that the left avatar has raised his right hand while the right avatar is making hand gesture while he is talking.

Figure 9 Hand gesture

3. Sharing and Manipulation of 3-D Objects

3-D objects can also be shared in a multi-user virtual environment. For example, in the multi-user networked game Quake, the weapons and the ammunitions (ammos) are examples of 3-D objects shared by multiple players. In the virtual world of the game, the players see the same set of weapons and ammos and a player can pick it up by passing through the item. Other multi-user virtual environment such as Interspace [20] allows users to build their own virtual worlds and let other users enter these worlds, by first sending the world information to the other users. Usually these systems contain 3-D objects that are predefined by the central server. It is more appealing, however, if a user is able to share a 3-D object from his/her side to the other users, without the server understanding what the object is. For example, during a discussion, a user wants to show different components of a

car, and then the user can select a 3-D car model stored in his/her computer to be sent to other users so that the 3-D car model appears in the virtual environment seen by all users. Then the user can explain different components of a car by simply referring to the 3-D model. Thus the sharing and manipulation of 3-D objects makes the communication more interactive.

Another issue for sharing and manipulating the 3-D objects is the transmission. The 3-D models usually have one or more attributed data such as normals, colors, or textures. A complex scene or a complex object can have a large size and it may take very long time to download. As a result, like video and audio, the 3-D model can be streamed so that a coarse model is seen first and the model is updated for refinement as more information is received. For example, one of the projects performed at our lab is Progressive 3-D Model Streaming [21] that sends the most important bits to represent the 3-D models and then the less important bits, allowing the users to stop the download at any time they want.

In NetICE, a participant can drag and drop a file in VRML (Virtual Reality Modeling Language) format that is a standard language for describing interactive 3D objects and worlds delivered across the Internet [22]. This 3-D object will be rendered in the NetICE virtual room. At the same time, this 3-D object will also be sent to the server that will update its record of the list of 3-D objects and will send this 3-D object to the other clients in the current room, as illustrated in Figure 10. When a new participant enters this room, the server will send out all the 3-D objects to this corresponding client after the current virtual room information is sent. In addition to 3-D object sharing, the participant is also able to move the 3-D object around the room by using the mouse to drag it to another position. If a collision is detected between this 3-D object with another 3-D object present in the room, then the two objects will be bounced away according to the angle of impact and the size of the objects.
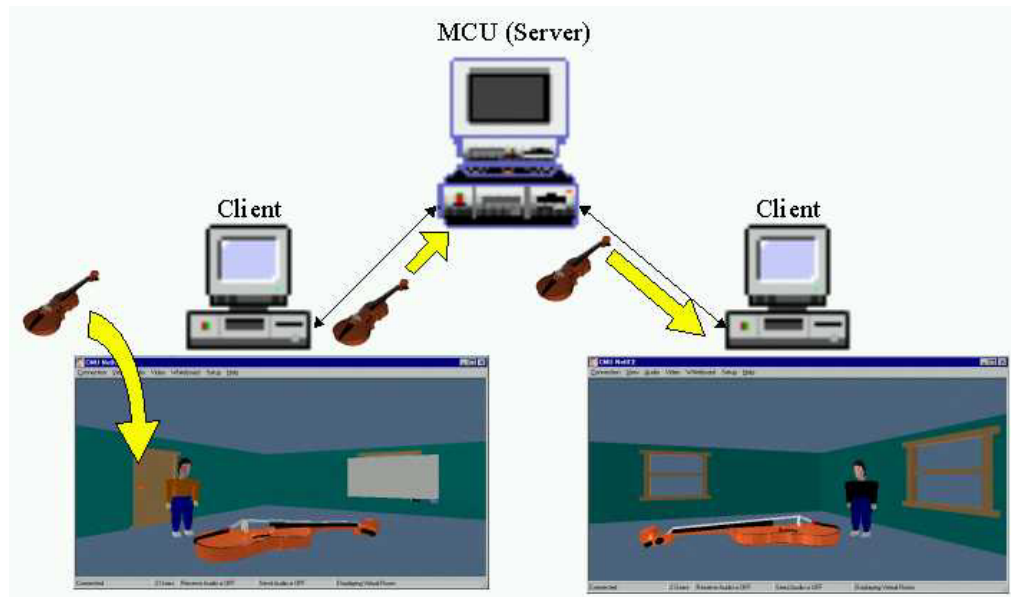
Figure 10 3-D object sharing

4. Shared Whiteboard

In an office, a whiteboard can be used to sketch a plan, write down reminders or illustrate an idea to visitors during an informal discussion. Some video conferencing applications such as NetMeeting [2] or CU-SeeMe [3] incorporate an electronic shared whiteboard that facilitates information sharing among users. The shared whiteboard allows multiple users from different locations to do collaborative work. Traditionally, a whiteboard (or blackboard) is used in a classroom for the teacher to write down course materials for the students to learn. As a result, a shared whiteboard can be included in a multi-user virtual environment so that it is suitable for distance learning. In addition, more than one shared whiteboard can appear in a virtual environment and each whiteboard may be used for a specific function that can be associated with the surroundings. For example, in a virtual office building, users may use the whiteboard for checking important phone numbers or appointment times in their offices or they may leave a message for other users on the whiteboard in the corridor.

4.1 Capture Device for Shared Whiteboard

In a classroom, blackboard and chalk have long been essential tools for teachers to convey the message to the students. In an office, the presence of a physical whiteboard and markers allows people to sketch their ideas. In these scenarios, people need to copy down the writings or drawings on the whiteboard to their notebook in order to keep a record. To remove the burden of the manual copying, a whiteboard session can be captured electronically by using some capture devices at the same time as the content is created. The capture devices usually consist of two components: an input device and a visual feedback device. The input device allows the user to write and draw naturally and its underlying electronic components capture the pen-down sample points. The visual feedback device allows the user to see instantly what is on the board so that they can continue to write or draw consistently. There exists various kinds of capture devices to achieve this purpose and we will discuss them in the following sections.

4.1.1 Mouse and Monitor

The most common way of capturing pen strokes is to use a mouse and then display the captured strokes on the monitor. Since mouse and monitor are standard components of a computer, no extra hardware is required. The shared whiteboard application simply captures mouse coordinates when the button is pressed. The problem with this approach is that it is difficult to draw with a mouse because the user needs to keep his/her hand steady. As a result, drawing or writing with a mouse is slow and the resulting drawings may contain a lot of jitterings.

4.1.2 Tablet and Monitor

The tablet is a pen-based device that can be attached to a computer and it allows users to draw naturally on it. While the user is drawing on the tablet, the visual feedback is provided on the

monitor. But since the visual feedback area (monitor) is different from the drawing area (tablet), it can be difficult to control the pen to start at the specific desired location after the user has a pause and then resumes drawing. Figure 11 illustrates an example of a tablet called graphire2 manufactured by WACOM [23].

Figure 11 Tablet manufactured by WACOM

4.1.3 Wireless Pen and Paper Pad

With this setup, the user can draw with a special pen on regular paper attached to a pad. This method is even more natural since the user is performing the same action as jotting notes on regular paper. However, the drawback in this approach is that if the user wants to erase something from the electronically captured content, he/she may need to choose special function from the application and the resulting change will not reflect on the regular paper. As a result, the content that is captured and the content that is actually drawn on the regular paper may be out-of-sync after the user does some editing. Figure 11 illustrates an example of this kind of commercial product called SmartPad manufactured by Seiko Instruments [24].

Figure 12 SmartPad manufactured by Seiko Instruments

4.1.4 Wireless Pen, Physical Whiteboard and Projector

In this setting, a capture bar is mounted on the side of a physical whiteboard to detect the coordinates of the wireless pen when the user uses it to draw or write on the whiteboard. The wireless pen is inkless meaning that it does not leave any physical mark on the whiteboard. Its tip has a switch and will be turned on while it is pressed towards the whiteboard during writing, thus sending the coordinates of the wireless pen to the capture bar. A projector is setup to project the resulting virtual ink back to the whiteboard to give the user instant visual feedback. Calibration is required at the beginning in order to map the cursor coordinates of the application to the physical location of the wireless pen such that the projection will overlap correctly with the designated drawing area of the whiteboard. Similar to the paper pad in the previous case, the physical whiteboard is served as both drawing area and the visual feedback area. The physical whiteboard provides a much bigger drawing area. However, this approach has the drawback that occlusion may occur because the projection may be blocked by the user. There exists several commercial products for this kind of capture devices such as mimio [25] or eBeam [26]. An example setup with mimio

and the projected whiteboard display is shown in Figure 13. A more detailed description about this particular setup used with NetICE can be found in [27].
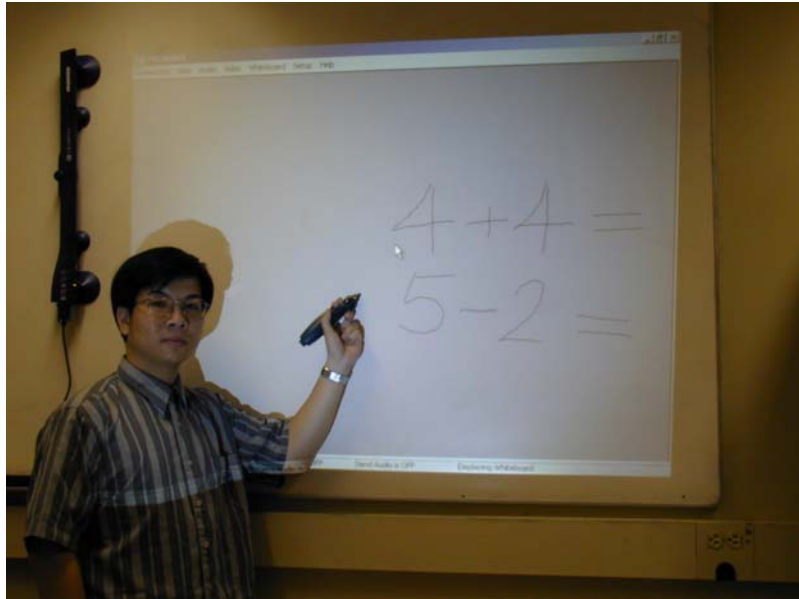

Figure 13 mimio and the projected whiteboard display

4.1.5 Stylus and Touch Screen

The user uses a stylus to write directly on top of a screen and whatever the user writes will be displayed instantly on the same screen. The screen is pressure sensitive in order to detect when and where the pen is down. The advantage of this setup is that the drawing area and the visual feedback area are integrated into the same device so that the inconsistency between drawing and displaying can be avoided. However, a touch screen is more expensive compared with other kinds of capture devices. On the other hand, as common to any typical new hardware, the cost is expected to decrease as time passes and this approach will be more and more common to the general public. The new product line Tablet PC that includes a pressure-sensitive tablet integrated into a laptop screen will make shared whiteboard applications much more convenient to use. An example Tablet PC manufactured by Acer is shown in Figure 14 [28].

Figure 14 Tablet PC manufactured by Acer

4.2 Storage Format for Shared Whiteboard

The user may want to save the content of the whiteboard after being captured by the input device so that he/she may review the content at a later time. There are two common formats that the whiteboard content can be stored: as an image or as strokes.

4.2.1 Image

The content on the whiteboard can be saved as an image. The background whiteboard may be represented by white pixels and the drawing or writing can be represented by black pixels. Using this format, the memory required to store the whiteboard content as a raw image is fixed since the total number of pixels on the whiteboard is constant. The image may be compressed to reduce the space required for storage. The image represents only the final outcome and the intermediate results are not stored. Timing information is not included in the image format. One can save a whiteboard session as a video, i.e., save the whiteboard image buffer at different time instants, to show the progress over time. However, it will require a lot of storage space or else the resolution will be poor if compressed with low bit rate by standard video compression techniques.

4.2.2 Strokes

An alternative solution to store the whiteboard content is to store them as strokes, i.e., save the sequence of the x and y coordinates captured from the input devices over time. Timing information is saved under this format therefore during playback the whiteboard session at any time instant can be easily displayed. Another advantage for the stroke format over the image format is the robustness to transformation. When an image is under rotation or scaling, the quality of the transformed image may be poor due to interpolation. However, since we are only transforming the stroke sample points while maintaining their connectivity, the resulting display can still have high resolution.

4.3 Archival for Shared Whiteboard Session

The reason for storing the whiteboard content is to be able to archive it at a later time. Here we classify the archival in two types: online archival and offline archival. The online archival refers to the scenario when the history of the whiteboard content needs to be archived during a live whiteboard session. The offline archival refers to the scenario when the relevant content needs to be archived according to what the user wants.

4.3.1 Online archival

In a virtual environment, participants may enter at any time. Therefore, it is necessary to keep the history of the whiteboard content and distribute it to the new incoming participant. In NetICE, when a participant first enters a virtual world, the NetICE server will send out the history of the whiteboard content to this participant in addition to other information about the virtual world. The participants can then draw or write simultaneously on the shared whiteboard and they may use any of the capture devices mentioned in Section 4.1 as the input device to capture the strokes on the shared

whiteboard in NetICE. All the captured strokes will be sent to the NetICE server to be distributed to the other NetICE clients and to update the whiteboard history of the NetICE server. Figure 15 illustrates the shared whiteboard in NetICE.
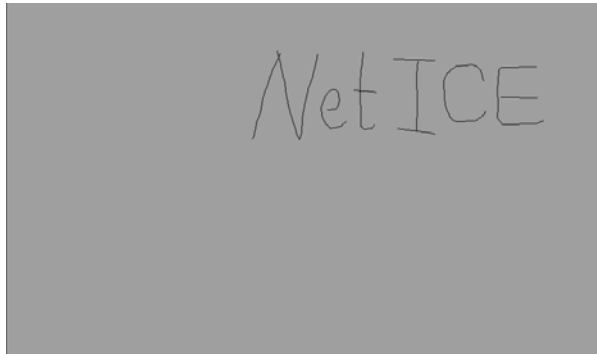

(a) Participant 1 is writing on the whiteboard   (b) The strokes are sent to be displayed on the shared whiteboard for the other participants

Figure 15 Shared whiteboard in NetICE

4.3.2 Offline archival

Given a lot of content from various whiteboard sessions, a database of sketches can be formed and it will be useful if we are able to search through this database for relevant information according to our input query. We have been developing a sketch retrieval system that can accept a rough sketch as a query and the sketch can consist of multiple components, as shown in Figure 16.
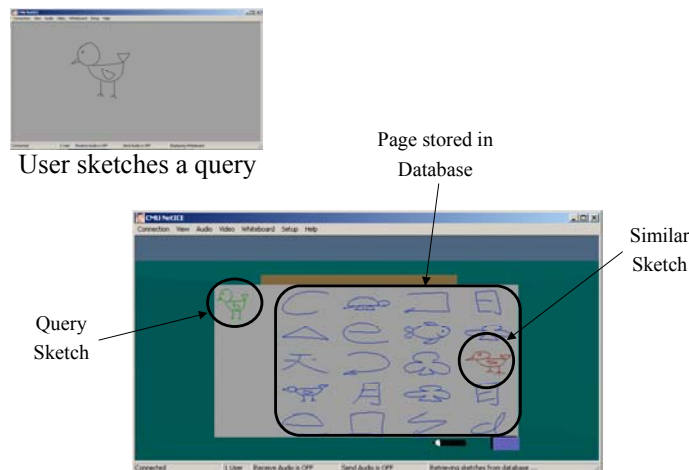
Figure 16 Prototype of free-form hand-drawn sketch retrieval system

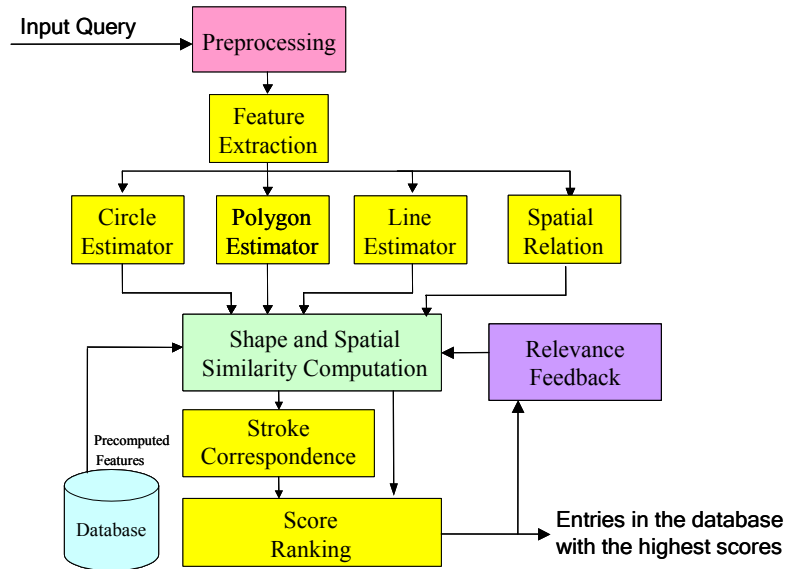The block diagram of our sketch retrieval system is shown in Figure 17.

Figure 17 Block diagram of our sketch retrieval system

Given an input query, some preprocessing is performed in order to obtain a good representation of the input data. The preprocessing modules include resampling the stroke into equidistant sample points and merging strokes based on proximity. Then the features are extracted from each stroke. These features are used for the shape estimators to determine the likelihood that each stroke falls in each primitive shape: line, circle and polygon. A likelihood measure that takes the value between 0 and 1 is assigned for each stroke with respect to each shape. In addition, the spatial relation between the stroke pairs is also considered. Based on the shape and spatial similarity between strokes, the stroke correspondence between two sketches is determined. Then the final similarity scores between the query sketch and each sketch in the database are computed. The sketches in the database with the highest similarity scores are retrieved. The user has the option to provide relevance feedback to the system by indicating positive and negative examples. The system will retrieve new results after learning from those examples. More details about our prior work on the sketch retrieval system can be found in [29][30][31].

5. Conclusions

In this article, we have reviewed various immersive technologies that enhance multimodal interaction between the participants in our multi-user 3D virtual environment system NetICE. A demo version of NetICE can be downloaded from *http://amp.ece.cmu.edu/*. We believe that besides speech, behaviors such as directional sound, eye gaze, hand gestures, lip synchronization, facial expressions are all important aspects of communication and should be performed by the computer using speech processing and animation techniques to increase the sense of immersion. The ability to share and manipulate 3-D objects in the virtual environment makes the communication more interactive and stimulates collaboration between remote sites so that the travel cost can be reduced. With the availability of better and better pen-based capture devices, the shared whiteboard provides a natural means of exchanging ideas in the virtual environment. It brings up a lot of interesting research issues about how to store the whiteboard session to facilitate the archival at a later time.

We believe that in the future, with higher bandwidth, better compression and faster rendering techniques, more complex 3D models with textures can be used to render an even more realistic environment. In a few years, pen-based computing will become a ubiquitous platform such that note-taking will be completely inkless and traditional pen and paper are no longer needed. With the abundance of bandwidth, more and more PCs will be connected together to form communities of different interest groups. Each group will have its own virtual environment where community members can meet. Eventually these multi-user 3D virtual environments will evolve from a human-computer interface to a truly immersive human-to-human interface to provide transparent communication between people.

References

1. ICQ, online chat software, *http://www.icq.com/*.

2. Microsoft, NetMeeting 3, videoconferencing software, *http://www.microsoft.com/windows/netmeeting/*.

3. White Pine, CU-SeeMe, videoconferencing software, *http://www.wpine.com/Products/CU-SeeMe/*.

4. Leung W. H. and Chen T., "Creating a Multiuser 3-D Virtual Environment", *IEEE Signal Processing Magazine*, May 2001.

5. Leung W. H., Goudeaux K., Panichpapiboon S., Wang S.-B. and Chen T., "Networked Intelligent Collaborative Environment (NetICE)", *IEEE Intl. Conf. on Multimedia and Expo.,* New York, July 2000.

6. Leung W. H. and Chen T., "Networked Collaborative Environment with animated 3D avatar", *IEEE Workshop on Multimedia Signal Processing*, Los Angeles, California, Dec 1998.

7. Escher M., Panzic I., and Magnenat-Thalmann N., "Facial Deformation from MPEG-4." Proceedings on Computer Animation '98, 1998.

8. Kurihara T. and Arai K., "A Transformation Method for Modeling and Animation of the Human Face from Photographs." Proc. Computer Animation '91, Springer-Verlag, Tokyo, pp. 45-58, 1991.

9. Pighin F., Hecker J., Lischinski D., Szeliski R., and Salesin D., "Synthesizing Realistic Facial Expressions from Photographs." SIGGRAPH '98, 1998.

10. Leung W. H., Tseng B., Shae Z.-Y., Hendriks F. and Chen T. "Realistic Video Avatar", *IEEE Intl. Conf. on Multimedia and Expo.*, New York, July 2000.

11. Fua P., "Face Models from Uncalibrated Video Sequences." Proc. CAPTECH '98, pp. 214-228, 1998.

12. Gardner, W. G., "3-D audio and Acoustic Environment Modeling". Wave Arts, Inc., March 1999.

13. Gardner, W. G., "Transaural 3-D Audio", MIT Media Lab, July 1995.

14. Gardner, W. G. and Martin K. D., "HRTF measurements of a KEMAR", *J. Acoust. Soc.*, vol 97, 1995.

15. McGurk, H., and MacDonald, J., "Hearing lips and seeing voices," *Nature*, pp. 746-748, December 1976.

16. Tang, J. C. and Isaacs, E. A., "Why do users like video? Studies of multimedia-supported collaboration", *Computer Supported Cooperative Work (CSCW)*, 1, 163-193, 1992.

17. Foley, J. D., van Dam, A., Feiner, S. K., Hughes, J. F., "Computer Graphics: Principles and Practice". Addison-Wesley, 1990.

18. Parke, F. I., Waters, K., "Computer Facial Animation". A K Peters, 1996.

19. Waters, K. "A muscle model for animating three-dimensional facial expressions". Computer Graphics (SIGGRAPH '87), 21(4):17-24, July 1987.

20. NTT Software Corporation, Interspace, 3-D virtual environment, *http://www.ntts.com/ispace.html*.

21. Okuda M. and Chen T., "Joint Geometry/Texture Progressive Coding of 3-D Models", *IEEE Intl. Conf. on Image Processing*, Vancouver, Sep 2000.

22. Web3D Consortium, the VRML/Web3D Repository, *http://www.web3d.org/vrml/vrml.htm*.

23. WACOM Technology co., graphire2 tablet, *http://www.wacom.com/graphire/*.

24. Seiko Instruments USA Inc., SmartPad, *http://www.seikosmart.com/products/sp580.html*.

25. Virtual Ink Inc., Mimio, *http://www.mimio.com/.*

26. Electronics Imaging Inc., E-Beam, *http://www.e-beam.com/.*

27. Tseng B., Shae, Z.-Y., Leung W. H. and Chen, T., "Immersive Whiteboards in a Networked Collaborative Environment", *IEEE Intl. Conf. on Multimedia and Expo.*, Tokyo, August 2001.

28. Acer Inc., Tablet PC, *http://aac.acer.com/.*

29. Leung W. H. and Chen T. "User-Independent Retrieval of Free-Form Hand-Drawn Sketches". *ICASSP2002*, Orlando, Florida, May 2002.

30. Leung W. H. and Chen T. "Retrieval of Sketches Based on Spatial Relation Between Strokes". *ICIP2002*, Rochester, New York, September 2002.

31. Leung W. H. and Chen T. "Trademark Retrieval Using Contour-Skeleton Stroke Classification". *ICME2002*, Lausanne, Switzerland, August 2002.